Going Digital Guide to Data Governance Policy Making

OECD

Going Digital Guide to Data Governance Policy Making

Please cite this publication as:
OECD (2022), *Going Digital Guide to Data Governance Policy Making*, OECD Publishing, Paris, https://doi.org/10.1787/40d53904-en.

ISBN 978-92-64-84416-2 (print)
ISBN 978-92-64-84995-2 (pdf)
ISBN 978-92-64-36908-5 (HTML)
ISBN 978-92-64-64415-1 (epub)

Photo credits: Cover © Sunward Art/Shutterstock

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

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 possibilities and pitfalls mean 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. Policy makers increasingly seek guidance on how to best design appropriate policies for data governance in a complex and evolving policy and technological landscape.

Drawing on fundamental commonalities for data governance across policy areas, the *Going Digital Guide to Data Governance Policy Making* (hereafter the *Guide*) helps policy- and decision-makers develop, revise and implement policies for data governance in the digital age. The *Guide* uncovers the policy tensions that underpin data governance and proposes policy options for balancing openness and control, the interests of different stakeholders, and the need for incentives to invest in data and their effective re-use.

Drawing on concrete examples from countries' policies and practices, as well as the guidance of OECD legal instruments, the *Guide* provides a checklist to orient data governance policymaking. In so doing, the *Guide* advances policy action to realise the benefits of data governance for growth and well-being.

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 *Going Digital Guide to Data Governance Policy Making* (hereafter the *Guide*) provides a practical tool to help policy makers design effective, technology-neutral, forward-looking and coherent data governance policies across sectors and jurisdictions. Complementing the report *Going Digital to Advance Data Governance for Growth and Well-being*, the *Guide* is another key outcome of 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.

The *Guide* was drafted by Francesca Casalini and Christian Reimsbach-Kounatze with input from Angela Attrey and Gallia Daor. It draws on the work of the project's co-leads from five OECD Directorates, including the Directorate for Science, Technology and Innovation (Angela Attrey, 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 *Guide* 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, Christian Biesmans, Sara Calligaris, Flavio Calvino, Alessandra Colecchia, David Gierten, Nicholas McSpedden-Brown, Simon Lange, Molly Lesher, 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), and the Public Governance Directorate (Miguel Amaral, Cecilia Emilsson, Marianna Karttunen, Jacob Arturo Rivera Perez and Barbara Ubaldi). 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 the *Going Digital Guide to Data Governance Policy Making* on 27 September 2022. The OECD Secretariat prepared the report for publication.

Table of contents

Figures

Boxes

Abbreviations and acronyms

ACCC	Australian Competition and Consumer Commission
AI	Artificial intelligence
AIAs	Accredited Integrating Authorities
CMA	Competition and Markets Authority
DGCCRF	Directorate-General for Competition, Consumer Affairs and Fraud Control
EASD	Enhancing access to and sharing of data
ERB	Ethics review body
EU	European Union
FAIR	Findability, accessibility, interoperability and re-use
FCA	Financial Conduct Authority
GDPR	General Data Protection Regulation
HCCI	Health Care Cost Institute
ICO	Information Commissioner's Office
ICT	Information communications technology
IoT	Internet of Things
IPR	Intellectual property right
NSO	National statistical organisation
OAIC	Office of Australian Information Commissioner
PEA	Privacy enforcement authority
PET	Privacy-enhancing technology
PIMS	Personal information management system
PPP	Public-private partnership
PSI	Public sector information
R&D	Research and development
SMEs	Small and medium-sized enterprises
SNA	System of National Accounts

Executive summary

Overview

The ubiquitous collection, use and sharing of data that power today's economies challenge existing governance frameworks and policy approaches. Drawing on the extensive research and analysis by the OECD on data governance, and the OECD legal instruments in this area, this *Going Digital Guide to Data Governance Policy Making* (hereafter the *Guide*) aims to help policy makers navigate three fundamental policy tensions and objectives that characterise most, if not all, efforts to develop, revise and implement policies for data governance in the digital age. The tensions and objectives relate to balancing data openness and control, while maximising trust; managing overlapping and potentially conflicting interests and regulations related to data governance; and incentivising investments in data and their effective re-use. For each, the *Guide* outlines underlying issues and presents promising approaches that can help address them. The *Guide* also contains a checklist of questions to orient policy makers as they develop and revise effective policies for data governance. Finally, it includes a number of policy approaches and real-life policies as examples.

Findings

Data openness brings both benefits and risks inherent to the economic properties of data

Data openness brings both benefits and risks. Policies need to reconcile how to: foster a culture of risk management and transparency across the data ecosystem; leverage the full spectrum of the data openness continuum to balance risks and benefits; enhance users' agency and control over data through legal means; support adoption of technological and organisational measures to enhance control; and enhance technical interoperability for data openness.

The interests of different stakeholders and policy communities can conflict

Multiple parties are involved in data-driven contexts with potentially conflicting interests at different phases of the data value cycle (data collection, analysis, use, deletion). The engagement of different policy communities in data governance thus gives rise to multiple, and sometimes overlapping, policy and regulatory frameworks that focus on concerns mainly relevant to those policy communities. These frameworks can be both sectoral or cross-sectoral, as well as national and international.

More incentives are needed to strengthen data governance

While the marginal costs of transmitting, copying and processing data can be close to zero, substantial investments are often required to generate and collect data and to enable data sharing and re-use. Investments may also be needed for data cleaning and data curation, often beyond the scope and timeframe of the activities for which the data were initially collected. In many cases, complementary investments are also needed in data-related

skills and competencies, as well as in information communication technologies. This includes investments in algorithms and software along the data value cycle (from generation and collection to processing and re-use). Indeed, evidence shows that firms are increasingly buying start-ups to secure access to data and other complementary assets that may be critical for the development of their data-driven business.

Recommendations

Balance data openness and control while maximising trust

Policies should seek to maximise the benefits from data access, sharing and re-use across organisational and national borders while addressing related risks, including the violation of the rights of individuals and organisations. To this end, data governance policies can:

- foster a culture of risk management and transparency across the data ecosystem
- leverage the full spectrum of the data openness continuum
- provide legal options and tools to enhance right-holders' agency and control over data
- support development and adoption of technological and organisational measures to enhance control of stakeholders over data
- enhance interoperability of data across organisations and sectors.

Manage overlapping and potentially conflicting interests and regulations related to data

Policies should balance the interests of different stakeholders, while ensuring consistency across different policy and regulatory frameworks. To this end, data governance policies can:

- identify and consider the contribution of different stakeholders throughout the data value cycle including by promoting multi-stakeholder engagement
- support cross-agency co-operation to help reconcile different domestic frameworks
- promote model contracts, contractual clauses, public procurement, codes of conduct and ethics frameworks to leverage contracts as means to clarify overlaps in data governance frameworks
- reconcile varying data governance frameworks across countries, promote international regulatory co-operation, including cross-border enforcement co-operation, to enable cross-border data flows with trust.

Incentivise investments in data and their effective re-use

Policies should incentivise investments in data and their effective re-use for data-driven solutions and a thriving data ecosystem, among others. To this end, data governance policies can:

- promote appropriate knowledge and skills for responsible data sharing and use
- encourage investments in and adoption of financially viable information and communication technology infrastructures for data openness
- foster competition in data-driven markets and address barriers to entry for new firms
- promote standardised approaches for evaluating the social and economic value of data.

1 Introduction

The *Going Digital Guide to Data Governance Policy Making* (hereafter the *Guide*) aims to advance the development, revision and implementation of policies for data governance, by helping to overcome key related policy tensions. Addressing the complexities arising from the nature of data as an intangible infrastructural resource of global strategic importance, the *Guide* helps policy makers design effective, technology-neutral, forward-looking and coherent data governance policies across sectors, policy domains and jurisdictions. It proposes a set of questions and highlights promising policy approaches based on three policy tensions and objectives that characterise data governance policy making: balancing data openness and control while maximising trust; managing overlapping and potentially conflicting interests and regulations related to data governance; and incentivising investments in data and their effective re-use.

Data are an intangible infrastructural resource that have significant potential spillover benefits across sectors and policy domains (OECD, 2015[1]; 2022[2]). Global economies are increasingly powered by digital technologies, and the ubiquitous collection, processing, use and sharing of data are challenging governance frameworks and policy approaches. For individuals, businesses and governments, these developments provoke fundamental questions. How can data be open while still controlled to maximise trust? How can overlapping and potentially conflicting interests and regulations related to data be managed? How can investments in data and their effective re-use be incentivised? In this context, a field of policy practice has emerged, broadly referred to as "data governance" (Box 1.1).

Box 1.1. What is data governance?

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 or institutional provisions, that affect data and their creation, collection, storage, use, protection, access, sharing and deletion 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.

Source: OECD (2022[3]), *Going Digital to Advance Data Governance for Growth and Well-being*, http://dx.doi.org/10.1787/e3d783b0-en.

Although policy and regulatory co-operation on data governance has been steadily improving, approaches remain largely siloed. Policy makers and regulators naturally tend to focus on their respective policy domains and country contexts. As such, they often fail to benefit from experiences in other policy domains. Furthermore, they face difficulties in fully accounting for the unintended consequences of their policies for stakeholders beyond their competence. In addition to benefiting from cross-disciplinary approaches, policy makers and regulators could also do more to manage overlaps and potential conflicts between multiple policy frameworks applying to the same data. All these challenges add to difficulties in defining a common ground and language on data governance at the national and international level.[1]

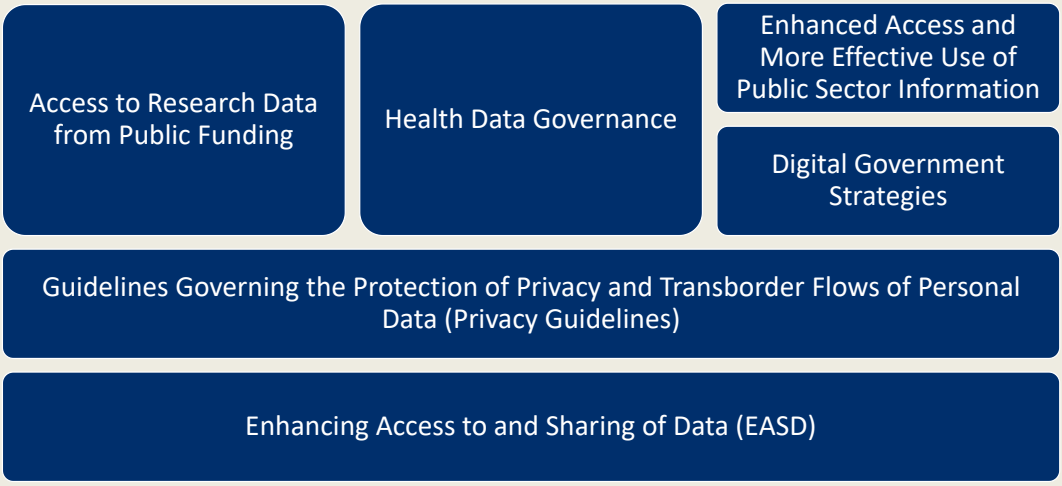
Governance challenges relating to data are particularly significant in a context of fast technological developments. Emerging technologies include cloud computing, big data analytics, artificial intelligence (AI), the Internet of Things (IoT), immersive environments and distributed ledger technologies. For example, the combination of big data analytics, AI and the IoT has enabled development of digital environments where personal data collection and processing are increasingly comprehensive and pervasive. These developments require innovative policy thinking in terms of privacy and personal data protection. This, in turn, raises questions about the role of informed consent or the applicability of principles such as purpose specification and use limitation, when not challenging the definition of personal data altogether (OECD, 2021[4]; 2019[5]; 2015[1]).[2]

Against this background, the *Guide* is conceived as a practical tool for data governance policy makers, complementing the OECD (2022[3]) report *Going Digital to Advance Data Governance for Growth and Well-being*. Chapter 2, which is key for using the *Guide*, presents a checklist for policy makers in different domains to review data governance policies.[3] Using the checklist, they can assess whether policies are fit for purpose to address the three fundamental tensions in data governance policy making and achieve related objectives. Chapter 3 elaborates the three cross-cutting policy tensions and objectives of data governance. For each one, the chapter presents possible governance approaches and relevant provisions of OECD Council Recommendations on data governance that support these approaches (Box 1.2). It also provides examples of how governments and the private sector have been implementing these approaches.[4]

Box 1.2. The OECD approach to data governance as reflected in its legal instruments

Throughout the *Guide*, the following OECD Recommendations are highlighted to provide direction in data governance policy making (Figure 1.1):

Figure 1.1. OECD Recommendations relating to data governance



- The OECD (2021[6]) *Recommendation of the Council on Enhancing Access to and Sharing of Data* provides the foundation of the OECD approach to data governance and is the most recent and overarching instrument in this field.
- The OECD *Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data* launched in 1980 (revised in 2013) was the first internationally agreed-upon set of privacy principles.

In addition, some OECD legal instruments provide domain-specific guidance on data governance:

- The OECD (2021[7]) *Recommendation of the Council concerning Access to Research Data from Public Funding* provides guidance on enhancing access to research data and other research-relevant digital objects from public funding.
- The OECD (2016[8]) *Recommendation of the Council on Health Data Governance* calls for national health data governance frameworks to encourage the availability and use of personal health data to serve health-related public interest purposes while promoting the protection of privacy, personal health data and data security.
- The OECD (2014[9]) *Recommendation of the Council on Digital Government Strategies* and the OECD (2008[10]) *Recommendation for Enhanced Access and More Effective Use of Public Sector Information* supports the development and implementation of digital government strategies and the access and use of public sector information, including data.

Source: OECD (2022[11]), "OECD Legal Instruments", https://legalinstruments.oecd.org (accessed 4 November 2022).

References

OECD (2022), "Data shaping firms and markets", *OECD Digital Economy Papers*, No. 344, OECD Publishing, Paris, http://dx.doi.org/10.1787/7b1a2d70-en. [17]

OECD (2022), "Fostering cross-border data flows with trust", *OECD Digital Economy Papers*, No. 343, OECD Publishing, Paris, http://dx.doi.org/10.1787/139b32ad-en. [15]

OECD (2022), *Going Digital to Advance Data Governance for Growth and Well-being*, OECD Publishing, Paris, http://dx.doi.org/10.1787/e3d783b0-en. [3]

OECD (2022), "Measuring the value of data and data flows", *OECD Digital Economy Papers*, No. 345, OECD Publishing, Paris, http://dx.doi.org/10.1787/923230a6-en. [2]

OECD (2022), *OECD Legal Instruments*, https://legalinstruments.oecd.org (accessed on 4 November 2022). [11]

OECD (2022), "Responding to societal challenges with data: Access, sharing, stewardship and control", *OECD Digital Economy Papers*, No. 342, OECD Publishing, Paris, http://dx.doi.org/10.1787/2182ce9f-en. [16]

OECD (2021), *Recommendation of the Council concerning Access to Research Data from Public Funding*, OECD/LEGAL/0347, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347. [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. [6]

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/official-document/C(2021)42/en. [4]

OECD (2019), "Challenges to consumer policy in the digital age", *Background Report G20 International Conference on Consumer Policy*, OECD, Paris, https://www.oecd.org/going-digital/topics/digital-consumers/challenges-to-consumer-policy-in-the-digital-age.pdf. [5]

OECD (2019), "Vectors of digital transformation", *OECD Digital Economy Papers*, No. 273, OECD Publishing, Paris, http://dx.doi.org/10.1787/5ade2bba-en. [12]

OECD (2018), *Digital Government Review of Sweden: Key Findings*, OECD Publishing, Paris, http://www.oecd.org/governance/digital-government/key-findings-digital-government-review-of-sweden-2018.htm (accessed on 13 July 2018). [13]

OECD (2016), *Recommendation of the Council on Health Data Governance*, OECD/LEGAL/0433, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433. [8]

OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264229358-en. [1]

OECD (2014), *Recommendation of the Council on Digital Government Strategies*, OECD/LEGAL/0406, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0406. [9]

OECD (2008), *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information*, OECD/LEGAL/0362, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0362. [10]

OECD (forthcoming), "Emerging privacy enhancing technologies: Maturity, opportunities and challenges", *OECD Digital Economy Papers*, OECD Publishing, Paris. [14]

Endnotes

[1] This is the case for instance for privacy-enhancing technologies that have long been promoted to address domestic challenges in the field of privacy and data protection (OECD, forthcoming[14]) and that are now also increasingly discussed as a means to address data governance challenges in other, international, contexts (OECD, 2022[16]; 2022[15]; 2022[17]). Similarly, sectoral governance arrangements such as in the domain of geodata could be scaled up for use of those data in other sectors, like the public sector (OECD, 2018[13]).

[2] Another example in the area of trade relates to the rules governing digital trade, which include issues related to cross-border data flows (OECD, 2022[15]).

[3] Compare with OECD (2019[12]) that presents the "vectors of digital transformation" including "a checklist against which existing and new policies can be reviewed to see if they are appropriate and fit-for-purpose in the digital era".

[4] Three complementary surveys by the OECD on data governance informed the collection of examples. These surveys were undertaken in context of respective work related to the OECD (2021[6]) *Recommendation of the Council on Enhancing Access to and Sharing of Data*, the OECD (2021[7]) *Recommendation of the Council concerning Access to Research Data from Public Funding*, and the OECD (2014[9]) *Recommendation of the Council on Digital Government Strategies*.

2 Checklist for assessing and designing data governance policies

This chapter presents a checklist for policy makers in any domain to review data governance policies and assess whether they address key policy tensions and achieve related objectives effectively. The policy tensions considered are the need to balance data openness and control while maximising trust; to manage overlapping and potentially conflicting interests and regulations related to data governance; and to incentivise investments in data and their effective re-use. Although these three tensions and objectives are common across most policies for data governance, not all questions and policy solutions will be relevant to all contexts.

In Table 2.1 the first column indicates the policy tension and related objective. The second column provides specific questions relating to that tension and links to the relevant subsection in Chapter 3. The last column highlights promising policy approaches that are explored further in Chapter 3. This includes how governments or the private sector have implemented these approaches, as well as references to relevant OECD Recommendations.

For example, to help government balance data openness and control while maximising trust, the checklist includes the following question: *Does the policy support the adoption of technological and organisational measures?* It then refers the reader to subsection 3.1.4 that elaborates on possible approaches to enable trusted data access and sharing through technological and organisational measures, such as data access control mechanisms, data intermediaries and privacy-enhancing technologies.

Table 2.1. Checklist for data governance policies

Cross-cutting policy tensions and objectives	Policy questions	Examples of policy approaches
Balancing data openness and control while maximising trust (section 3.1): *How policies can maximise the benefits from data access, sharing and re-use across organisational and national borders, while addressing related risks, including the protection of the rights of individuals and organisations.*	*Does the policy foster a culture of risk management and transparency across the data ecosystem?* (subsection 3.1.1)	Recommend the systematic implementation of risk management measures throughout the data value cycle. Promote transparency, considering the risk of information overload and other cognitive biases.
	Does the policy leverage the full spectrum of the data openness continuum? (subsection 3.1.2)	Design data governance arrangements that leverage different possible degrees of data openness, striving to be as open as possible and as closed as necessary.
	Does the policy provide options and tools to enhance users' agency and control over data? (subsection 3.1.3)	Formulate different consent models that allow individuals to exercise control while enabling business opportunities to benefit from data openness. Empower stakeholders through appropriate mechanisms such as data portability.
	Does the policy support adoption of technological and organisational measures to maximise trust? (subsection 3.1.4)	Provide data access control mechanisms, data intermediaries (e.g. data trusts, data commons, Personal Information Management Systems) and privacy-enhancing technologies.
	Does the policy enhance the interoperability of data across organisations, including within and across the public and private sectors? (subsection 3.1.5)	Provide data together with any required complementary resource, including metadata, documentation, data models and algorithms.
Managing overlapping and potentially conflicting interests and regulations related to data governance (section 3.2): *How policies can balance the overlapping and sometimes conflicting interests of stakeholders in data governance and clarify the relationship between different frameworks affecting data governance.*	*Does the policy identify and consider the contribution of different stakeholders in the data value cycle, including by promoting multi-stakeholder engagement?* (subsection 3.2.1)	Map impact on different stakeholders at different phases of the data value cycle to assess whether it reflects reasonable expectations and the public interest. Engage relevant stakeholders in the data ecosystem to identify their different interests and roles in data-driven value creation through open and inclusive processes.
	Does the policy support cross-agency co-operation to help reconcile different domestic frameworks affecting data governance? (subsection 3.2.2)	Encourage co-operation across the various regulatory and policy areas, including competition, privacy, consumer protection, as well as sector-specific regulators.
	Does the policy leverage contract to clarify and strengthen data governance? (subsection 3.2.3)	Collaborate with the private sector on voluntary guidance, codes of conduct, ethics frameworks and model contracts. Use public procurement to promote good data governance standards.
	Does the policy promote international regulatory co-operation to reconcile data governance across countries and enable cross-border data flows with trust? (subsection 3.2.4)	Promote interoperability of data governance frameworks to enhance cross-border data flows while protecting legitimate interests. Promote continued dialogue and international co-operation on ways to foster data access and sharing across countries.

Incentivise investments in data and complementary resources (section 3.3): *How policies can provide incentives for investments in data and their effective re-use.*	*Does the policy promote appropriate knowledge and skills for responsible data sharing and use?* (subsection 3.3.1)	Identify gaps and formulate strategies to develop and maintain the skills and infrastructures needed. Establish partnerships and data analytic support centres for development of data-related skills and the supply of data analytic expertise.
	Does the policy foster investments in and adoption of financially viable information and communication technology infrastructures for data openness? (subsection 3.3.2)	Promote adoption of data storage, processing and analytic services, especially for small and medium-sized enterprises. Promote adoption of new business and revenue models needed for data openness infrastructure. Promote adoption of shared data infrastructures. (e.g. interoperability buses) in the public and private sector and of the "once-only" principle in the public sector.
	Does the policy foster competition in data-driven markets and address barriers to entry for new firms? (subsection 3.3.3)	Assess the contribution of data to market power. Consider asymmetric approaches to ensure that competition measures address large incumbents and do not create barriers to entry for new firms.
	Does the policy promote standardised approaches for evaluating the social and economic value of data? (subsection 3.3.4)	Support promising approaches for valuing data, including in the context of the System of National Accounts and including efforts to measure their social value.

3 Cross-cutting policy tensions and objectives for data governance

This chapter elaborates three fundamental policy tensions and objectives common to data governance policy making across different domains: balancing data openness and control while maximising trust; managing overlapping and potentially conflicting interests and regulations related to data governance; and incentivising investments in data and their effective re-use. For each of these policy tensions and objectives, it outlines underlying issues and presents promising approaches that can help address them. These approaches are based on the OECD Horizontal Project on Data Governance for Growth and Well-being, on OECD Recommendations relating to data governance, and on relevant policy examples.

3.1. Balancing data openness and control while maximising trust

Balancing the social and economic benefits of "data openness" with its associated risks represents a fundamental policy challenge inherent to the economic properties of data. This section looks at key concepts associated with data openness, as well as associated benefits and risks. It then explores how to balance data openness and control while maximising trust, by fostering a culture of risk management and transparency across the data ecosystem; by leveraging the full spectrum of the data openness continuum to balance risks and benefits; by enhancing users' agency and control over data through legal means; by supporting adoption of technological and organisational measures to enhance control; and by enhancing technical interoperability for data openness.

Data openness exists along a continuum, with "open data" as one extreme. With each step towards openness, it becomes easier to access, share and re-use data, including across organisational and national borders. As such, the term "data openness" encompasses the ideas of "free flow of data" and "transborder data flows" (subsection 3.1.2).

The degree of data openness is determined by the legal, technical or financial requirements for data access, sharing and re-use.[1] Legal requirements are, in turn, defined by policy and regulatory frameworks applicable to data. These include privacy and data protection frameworks, intellectual property rights (IPRs), and national security and other sector-specific frameworks, and trade provisions also increasingly affect openness of cross-border data flows (OECD, 2022[1]). Contract law also helps determine rights and obligations related to data and thus data openness (subsection 3.2.3). Legal requirements affecting data openness may also be defined by *ex post* enforcement measures, such as where data portability is mandated as a competition enforcement mechanism (OECD, 2021[2]).

Technical requirements mostly refer to access control mechanisms, as well as specifications and standards required to access data. Finally, financial requirements, including pricing, such as determined by licensing agreements, can also affect the extent to which data users can afford data access and re-use.

A range of social and economic benefits are associated with greater data openness. Evidence suggests that as data openness increases, positive social and economic benefits also increase for data providers (direct impact), their suppliers and data users (indirect impact) and for the wider economy (induced impact) (OECD, 2019[3]). However, the magnitude of the relative effects will vary depending on sectors and context.[2]

The social and economic benefits of greater data openness include the following:

- enabling greater efficiency and productivity, transparency and accountability across society
- boosting sustainable consumption and growth, and enhancing social welfare and health care
- improving evidence-based policy making, as well as public service design and delivery
- enhancing consumer decision making and empowering users of digital goods and public and private services
- facilitating scientific discovery, enhancing its reproducibility, reducing duplication and enabling cross-disciplinary co-operation.

However, data openness can also pose risks. Data openness can generate significant risks for individuals and organisations related to three broad areas:

- **Violations of rights such as privacy and intellectual property rights (IPRs)**: the protection of privacy and to some extent of IPRs[3] is often considered as the biggest challenge associated with data openness. These risks include the privacy interests of individuals and the commercial interests of organisations (OECD, 2019[3]). They embody risks of data being used and re-used in ways that violate the applicable legal frameworks and contractual terms, or that deviate from the (reasonable) expectations of stakeholders and the law.

- **Data and digital security**: the risk of breaches or incidents that affect the availability, integrity or confidentiality of data (data security) and of information systems (digital security) also tend to increase with data openness. Opening information systems to access, share or transfer data may expose an organisation to digital security threats that affect data and information systems. The negative impacts of these breaches and incidents may have cascading effects along an entire supply chain. They may also undermine critical information systems, such as those in the health care, finance or energy sectors.
- **The unethical use of data**: more open approaches to data governance may also increase the risk of data use that violates ethical values and norms. These values could include fairness, human dignity, autonomy, self-determination, and the protection against undue bias and discrimination between individuals or social groups. Unethical use of data may still be legal under existing frameworks (OECD, 2021[4]; 2016[5]).

In this context policy approaches to reconcile the benefits and risks of data openness are needed. Balancing the social and economic benefits of data openness with associated risks represents a fundamental policy challenge inherent to the economic properties of data. On the one hand, data are non-rivalrous, which calls for maximum data openness to maximise the potential spillover benefits of data. However, data are also imperfectly excludable in that organisations and firms cannot generally prevent other people or institutions from re-using their data.

Coupled with their ability to be easily copied and re-shared, these characteristics can imply a risk of loss of control of data that increases with the degree of openness.[4] This risk is compounded as data are shared across multiple actors and machines, especially when these tiers cross multiple jurisdictions (OECD, 2022[1]). Consequently, the policy challenge lies in how data openness is positively correlated to both potential benefits and risks.

Many policies promote a "risk-based approach" to data governance to maximise the benefits of openness while minimising risks. Such an approach is helpful as it recognises that risk, like openness, is not binary but increases along a spectrum. This offers stakeholders the possibility to adjust the restrictedness of control and protection measures. In so doing, they would aim for the level of risk agreed upon or expected by stakeholders, while considering social and economic benefits and the public interest.

The rest of this section explores approaches to balance openness and control and materialise a "risk-based approach" to data governance. These approaches comprise fostering a culture of risk assessment and transparency across the data ecosystem (subsection 3.1.1); leveraging the full spectrum of the data openness continuum to balance risks and benefits (subsection 3.1.2); enhancing users' agency and control over data (subsection 3.1.3); supporting use of technological and organisational measures (subsection 3.1.4); and optimising technical interoperability of data across organisations and sectors (subsection 3.1.5).

3.1.1. Foster a culture of risk management and transparency across the data ecosystem

Stakeholders must be able to trust that risks related to data access, sharing and re-use are well managed. This will allow societies to fully reap the benefits of data-driven innovation. To this end, data governance policies should foster a culture of risk management across the data ecosystem.

OECD (2021[6]) defines the data ecosystem as "the integration of and interaction between different relevant stakeholders including data holders, data producers, data intermediaries and data subjects, that are involved in, or affected by, related data access and sharing arrangements, according to their different roles, responsibilities and rights, technologies, and business models." All stakeholders, including individuals, should be aware of risks as much as possible through appropriate transparency practices (OECD, 2019[3]).

Increasingly, privacy and data protection frameworks require businesses to evaluate risks before making decisions relating to data. For example, data controllers might need to conduct "risk assessments" to determine how best to protect organisational data. Depending on results, they may also have to report data breaches.

However, risk-based approaches remain challenging to implement for organisations. This is especially true where the rights of third parties are concerned. For example, it can be difficult to comply with third party rights with respect to personal data and the IPRs of organisations and individuals (OECD, 2019[3]).

In this context, stakeholders need different levels of transparency to trust the data ecosystem and make informed decisions. This transparency relates to what, how and by whom data are collected, accessed and used. Transparency is also crucial with respect to how data are governed. This includes information on processing, including with whom the data are shared, for what purpose and under what conditions access may be granted to third parties. It also addresses the rights, responsibilities and respective liabilities in case of violations.

Privacy and data protection frameworks have a variety of mechanisms to promote transparency in support of risk management. Data controllers, for example, may have obligations of transparency towards data subjects (individuals) and privacy enforcement authorities. At the same time, individuals could have the right to be informed, and to access and correct data. Such mechanisms, for example, can help identify risks related to discrimination. This measure empowers individuals to learn why certain decisions about them are made.[5]

However, mechanisms to promote transparency can fail. For example, the amount or complexity of the provided information may be too much for individuals to process (information overload). In addition, the presentation of information and related choices may make it difficult for individuals to exercise their rights effectively. This might be due to "dark patterns" in privacy consent mechanisms (subsection 3.1.3) or behavioural and cognitive biases (OECD, 2022[7]; 2022[8]).

New initiatives aim to rebalance asymmetries between stakeholders, better enabling individuals to reap the benefit and value of their data and enhancing transparency. These include, for example, a more user-friendly presentation of privacy policies, informed by behavioural insights and, potentially, facilitated through information intermediaries or the standardisation of information. They also include better defaults that make the choice for privacy protection settings at least as easy as privacy-intrusive ones (OECD, 2022[8]; 2022[7]).

How OECD legal instruments encourage risk management

The OECD (2021[6]) *Recommendation of the Council on Enhancing Access to and Sharing of Data* (hereafter the *EASD Recommendation*) calls on Adherents to

> *[t]ake necessary and proportionate steps to protect … legitimate public and private interests as a condition for data access and sharing"* (V.b); *[and] ensure that stakeholders are held accountable in taking responsibility, according to their roles … for the systematic implementation of risk management measures…, including [data security] measures*. (V.c)

The OECD (2021[9]) *Recommendation of the Council on Access to Research Data from Public Funding* (hereafter the *Research Data Recommendation*) recommends that Adherents

> *[t]ake steps to transparently manage risks posed by enhancing access to sensitive categories of research data and other research-relevant digital objects from public funding, including personal data, by applying specific risk mitigation measures, as well as providing for a "right to know" in cases of digital security incidents affecting the rights and interests of stakeholders.* (III.3)

The OECD (2016[10]) *Recommendation of the Council on Health Data Governance* (hereafter the *Health Data Governance Recommendation*) states that national health data governance frameworks should provide for control and safeguard mechanisms which

> *include formal risk management processes, updated periodically that assess and treat risks, including unwanted data erasure, re-identification, breaches or other misuses, in particular when establishing new programmes or introducing novel practices.* (III.11.iv)

The OECD (2015[11]) *Recommendation of the Council on Digital Security Risk Management for Economic and Social Prosperity* provides a set of operational principles that apply specifically to digital security risk management in data governance. These include the principles on "risk assessment and treatment cycle" that emphasises that "risk assessment should be carried out as an ongoing systematic and cyclical process", as well as the principle on "security measures", according to which

> *[digital security] risk assessment should guide the selection, operation and improvement of security measures to reduce the digital security risk to the acceptable level determined in the risk assessment and treatment. Security measures should be appropriate to and commensurate with the risk and their selection should consider account their potential negative and positive impact on the economic and social activities they aim to protect, on human rights and fundamental values, and on the legitimate interests of others.*

The OECD (2014[12]) *Recommendation of the Council on Digital Government Strategies* (hereafter the *Digital Government Recommendation*) suggests that governments

> *[d]evelop and implement digital government strategies which create a data-driven culture in the public sector, by balancing the need to provide timely official data with the need to deliver trustworthy data, managing risks of data misuse related to the increased availability of data in open formats.* (II.3)

It further specifies that these strategies should "reflect a risk management approach to addressing digital security and privacy issues and include … effective and appropriate security measures …" (II.4).

The OECD (2013[13]) *Recommendation of the Council concerning Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (hereafter the *Privacy Guidelines*) in Part Three on Implementing Accountability recommends that "a data controller should … have in place a privacy management programme that … provides for appropriate safeguards based on privacy risk assessment."

How OECD legal instruments encourage transparency across the data ecosystem

The OECD (2021[6]) *EASD Recommendation* calls on Adherents to

> *[e]nhance transparency of data access and sharing arrangements to encourage the adoption of responsible data governance practices … that meet applicable, recognised, and widely accepted … standards and obligations, including codes of conduct, ethical principles and privacy and data protection regulation. Where personal data is involved, Adherents should ensure transparency in line with privacy and data protection frameworks with respect to what personal data is accessed and shared, including with whom it is shared, for what purpose, and under what conditions access may be granted to third parties.* (III.c)

The OECD (2021[9]) *Research Data Recommendation* recommends that Adherents should

> *[f]oster searchable access to metadata that describes those datasets while respecting legal rights, ethical, principles, and/or, legitimate interests* (III.2.b) *[and] promote, and require where appropriate, the inclusion of information about rights and licensing in the metadata of all research data and other research-relevant digital objects from public funding ….* (V.4)

The OECD (2016[14]) *Recommendation of the Council on Consumer Protection in E-commerce* contains principles on transparency in consumer data practices, such as the principles on Fair Business, Advertising and Marketing Practices. It states that "businesses should not make any representation or omission or engage in any practice that is likely to be deceptive, misleading, fraudulent or unfair" (paragraph 4). Furthermore:

- "Businesses should not engage in deceptive practices related to the collection and use of consumers' personal data." (paragraph 8)
- "Businesses should comply with any express or implied representations they make about their adherence to industry self-regulatory codes or programmes, privacy notices or any other policies or practices relating to their transactions with consumers" (paragraph 11)
- "Businesses should protect consumer privacy by ensuring that their practices relating to the collection and use of consumer data are lawful, transparent and fair, enable consumer participation and choice, and provide reasonable security safeguards" (paragraph 48).

The OECD (2016[10]) *Health Data Governance Recommendation* provides guidance on "transparency, through public information mechanisms which do not compromise health data privacy and security protections or organisations' commercial or other legitimate interests" (III.7). These mechanisms should clarify:

- "the purposes for the processing of personal health data, and the health-related public interest purposes that it serves, as well as its legal basis"
- "the procedure and criteria used to approve the processing of personal health data, and a summary of the approval decisions taken, including a list of the categories of approved data recipients"
- "information about the implementation of the health data governance framework and how effective it has been." (III.7.iii)

The OECD (2013[13]) *Privacy Guidelines* in their Openness Principle recommend "a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller."

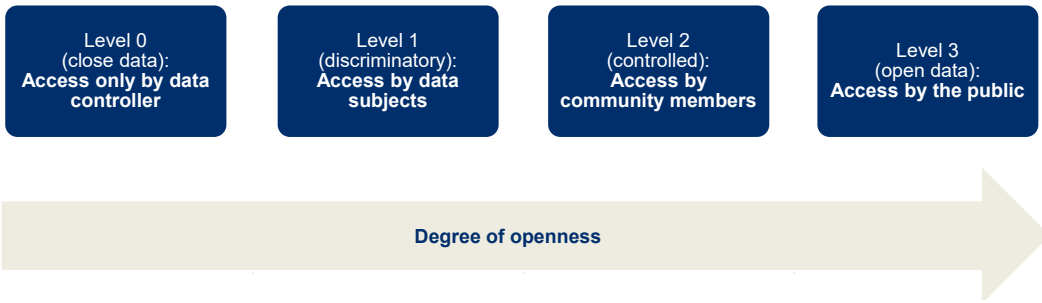
3.1.2. Balance risks and benefits via the full spectrum of the data openness continuum

Data governance approaches can be characterised by the different levels of openness that they establish. At one end is full and unrestricted openness (e.g. open access to data; no regulation of cross-border data flows). Along the continuum are arrangements that condition or limit the access, sharing, transfer and/or use of data to specific users or countries, or for specific use cases. Governing approaches often seek a balance so that data are "as open as possible, as closed as necessary" (OECD, 2021[6]).

In the case of data access and sharing, policy makers can leverage a wide range of arrangements along the data openness continuum (that moves from Level 0 to 3 as indicated in Figure 3.1). At one extreme, arrangements include those that limit access and re-use of data only to the data holder (Level 0) (OECD, 2019[3]). At the other extreme are open data arrangements (Level 3). These are "non-discriminatory data access and sharing arrangements, where data is machine readable and can be accessed and shared, free of charge, and used by anyone for any purpose subject, at most, to requirements that preserve integrity, provenance, attribution, and openness" (OECD, 2021[6]). Open arrangements are the most prominent approach to enhance access to data in particular in the public sector[6] (OECD, 2018[15]) and in science (OECD, 2020[16]).

In between the two extremes, data can be made available to a specific external stakeholder (Level 1). In data portability arrangements, for example, a natural or legal person can request that a data holder transfer to the person, or to a specific third party, data concerning that person in a structured, commonly used and machine-readable format on an ad hoc or continuous basis (OECD, 2021[17]). Data portability has become an essential tool for enhancing access to and sharing of data across digital services and platforms (subsection 3.1.3). It can empower users to play a more active role in the re-use of their data and increase interoperability. In this way, it can enhance competition and innovation by reducing switching costs and lock-in effects (OECD, 2021[17]; 2021[2]).

Figure 3.1. The degrees of data openness in approaches to data access and sharing



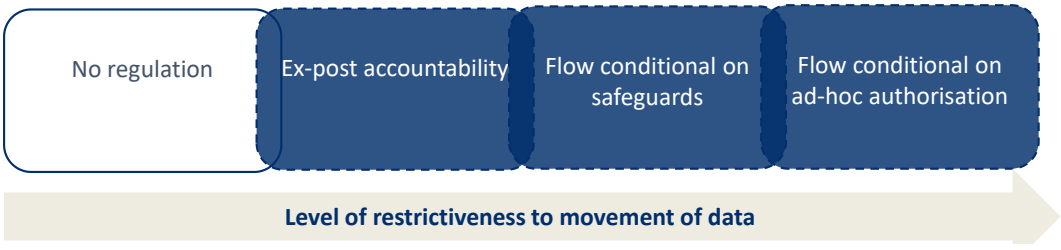
Source: OECD (2019[3]), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-Use Across Societies*, http://doi.org/10.1787/276aaca8-en.

Data can also be shared non-discriminatorily among members of a community[7] (Level 2), for instance, through conditioned data access and sharing arrangements. These arrangements "permit data access and sharing subject to terms that may include limitations on the users authorised to access the data (discriminatory arrangements), conditions for data use including the purposes for which the data can be used, and requirements on data access control mechanisms through which data access is granted" (OECD, 2021[6]). Conditioned data access and sharing arrangements are typically used in two scenarios. In the first, data are considered too confidential to be shared openly with the public (as open data arrangements). In the second, there are legitimate (commercial and non-commercial) interests opposing such sharing. In the latter cases, however, there can still be a strong economic and/or social rationale for sharing data between data users within a restricted set of trusted members of a community, under voluntary and mutually agreed terms.

Arrangements governing cross-border data flows is another example. Here, the continuum spans four categories of approaches to cross-border data transfers (Casalini, López González and Nemoto, 2021[18]) from the most restrictive to the most open (Figure 3.2):

- **Data flows are conditional on ad hoc authorisation**. These typically "relate to systems that only allow data to be transferred on a case-by-case basis subject to review and approval by relevant authorities" (Casalini, López González and Nemoto, 2021[18]).
- **Data flows are conditional on safeguards**. This includes approaches relying on the determination of adequacy or equivalence as ex-ante conditions for data transfer.
- **Ex-post accountability**. This approach does not prohibit cross-border transfer of data or require fulfilment of conditions. However, it provides for ex-post accountability for the data holder sending data abroad in case of misuse.
- **No regulation** of cross-border data flows. This is the most common open approach, having no privacy and data protection legislation at all.

Figure 3.2. The degrees of openness in approaches to cross-border data flow regulation



Source: OECD based on Casalini, López González and Nemoto (2021[18]), "Mapping commonalities in regulatory approaches to cross-border data transfers", https://dx.doi.org/10.1787/ca9f974e-en.

The optimal level of data openness along possible continuums will vary on a case-by-case basis. The level will depend on the risks associated with data openness, and therefore largely on the use cases or the type of data concerned. For example, the transfer of health data will tend to be less open than data related to product maintenance, which tends to pose lower risks (OECD, 2022[19]). In other words, high risks entail a greater need for control measures to protect the rights of stakeholders, and thus a tendency for less data openness. The presence of risks will in many cases incentivise and justify more controlled and restrictive arrangements, such as "conditioned" data access and sharing arrangements. Conversely, fully open arrangements such as open data will tend to be more appropriate where risks, including digital security and privacy risks, are insignificant.

How OECD legal instruments encourage leveraging of the data openness continuum to better balance risks and benefits

The OECD (2021[6]) *EASD Recommendation* states that adherents should "encourage data access and sharing arrangements that ensure that data are as open as possible to maximise their benefits and as closed as necessary to protect legitimate public and private interests." (V.a)

The OECD (2021[9]) *Research Data Recommendation* recommends to

> *[f]oster and support open access by default to research data and other research-relevant digital objects from public funding. … In cases where access needs to be partially or totally restricted to conform to legal rights, ethical principles and/or to protect legitimate private, public, or community interests, and with the ultimate objective of facilitating access which is as open as possible: … foster more limited forms of access such as access to aggregated or de-identified data, restricted access within safe and secure environments to certified users with clearance adapted to the sensitivity of data, or access via analyses that share only de-identified results.*

The OECD (2008[20]) *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information* (hereafter the *PSI Recommendation*) recognises, under its "Openness" principle, the importance of access but also the need for refusal and limitations to access by

> *[m]aximising the availability of public sector information for use and re-use based upon presumption of openness as the default rule to facilitate access and re-use. … [d]efining grounds of refusal or limitations, such as for protection of national security interests, personal privacy, preservation of private interests for example where protected by copyright, or the application of national access legislation and rules.*

3.1.3. Enhance users' agency and control over data through legal means

Different types of consent models

Protection of privacy and personal data is a common concern in data governance. It raises questions for policy makers on how best to balance data openness and control of personal data, while maximising trust. In this context, consent has been recognised as a mechanism to allow individuals to control the collection and (re-use) of personal data about them. Consent generally requires organisations to provide clear information to individuals about what personal data about them are being collected and used, and for what purpose. These rules are specified in the data protection and privacy laws of most countries and in the OECD (2013[13]) Privacy Guidelines (see subsection 3.1.1 on transparency). Individuals are also to be given "reasonable means to extend or withdraw their consent over time" (OECD, 2015[21]). As such, consent can be seen as an effective tool to empower individuals with respect to openness and control decisions, thereby helping to maximise trust.

At the same time, individuals may often not read or comprehend information regarding use of their data (including from mandatory disclosures) due, for example, to information overload (OECD, 2022[7]). Furthermore, some businesses have used "dark commercial patterns"[8] in privacy consent mechanisms such as cookie consent notices. These make it easier for individuals, for example, to opt into privacy-intrusive settings than to opt out. Often, this approach exploits behavioural and cognitive biases, effectively obstructing a deliberate consent decision (OECD, 2022[8]).

Additionally, data can be used and re-used, often in ways unforeseen at the time of collection. In this context, to retain flexibility in their compliance with privacy legislation, some organisations rely on one-time general or broad types of consent. These types of consent ostensibly provide some degree of control to individuals about the collection and use of data about them, without requiring consent for each specific instance of use. On the one hand, these types of consent seem to respect individuals' rights without pre-empting the social benefits that may accrue when those data are used for additional, previously unspecified aims. On the other, data subjects may not realise the full implications of broad consent, particularly in the context of artificial intelligence (AI) and big data analytics.

Increasingly, as highlighted in subsection 3.2.1, individuals cannot be fully aware of how observed, derived, inferred and personal data could reveal information about them. Nor can they know how these data could be used and shared between data controllers and third parties. They also cannot anticipate to what extent their data will be used for purposes that may transgress their moral values.

Consent remains a useful tool, especially in some contexts and circumstances. However, uncertainty about its use and effectiveness has led to the proposition of new models in the scientific literature, including "adaptive" or "dynamic" forms of consent (OECD, 2022[22]; 2015[21]). In time-restricted consent models, for example, individuals consent to use of personal data about them only for a limited period. This enables them to consent to new projects or to alter their consent choices in real time as circumstances change, while having confidence these changing choices will be respected.

However, behavioural and cognitive limitations or the presence of dark commercial patterns can call into question when and to what extent individuals can meaningfully consent to the collection and use of their personal data. Therefore, measures based on consent and information disclosure are likely insufficient in isolation. They may thus need to be accompanied by other regulatory measures. These could include requirements for "privacy-by-design" or prohibition of certain harmful data practices, such as privacy-intrusive dark patterns or potentially exploitative data-driven personalisation practices (OECD, 2022[8]; forthcoming[23]).

Data portability

Data portability has become an essential tool for enhancing access to and sharing of data across digital services and platforms. It can help increase interoperability and data flows and thus enhance competition and innovation by reducing switching costs and lock-in effects (OECD, 2021[2]; 2021[17]). Most importantly, an increasing number of countries is considering data portability as a means to grant users better agency and control over "their" data. This, in turn, empowers users to play a more active role in the re-use of these data.

Data portability initiatives and arrangements differ significantly along five key dimensions (OECD, 2021[17]):

- sectoral scope, including whether they are sector-specific or horizontal and thus directed potentially at all data holders regardless of the sector
- beneficiaries, including whether only natural persons (individuals) or also legal persons (i.e. businesses) have a right to data portability
- the type of data subject to data portability arrangements, including whether data portability is limited to personal data and includes volunteered, observed or derived data
- legal obligations, especially the extent to which data portability is voluntary or mandatory and if the latter, how it is enforced

- modalities of data transfer, meaning the extent to which data transfers are limited to or include ad hoc (one-time) downloads of data in machine-readable formats (regarded as "data portability 1.0"), ad hoc direct transfers of data to another data holder ("data portability 2.0"), or real-time (continuous) data transfers between data holders that enables interoperability between their digital services ("data portability 3.0").[9]

How OECD legal instruments address the need to enhance users' agency and control over data

The OECD (2021[6]) *EASD Recommendation* recommends that Adherents should

> *[e]mpower individuals, social groups, and organisations through appropriate mechanisms and institutions such as trusted third parties that increase their agency and control over data they have contributed or that relate to them, and enable them to recognise and generate value from data responsibly and effectively (III.d)… and [f]oster competitive markets for data through … appropriate measures, including enforcement and redress mechanisms that increase stakeholders' agency and control over data.* (VI.a)

The OECD (2021[9]) *Research Data Recommendation* recommends that Adherents should

> *[r]equire that consent or comparable legal basis be sought consistently for all collections of sensitive human subject data and metadata, including personal data, and that any use be in conformity with the consent granted, applicable privacy regulations, and ethical principles. Where it is proposed that personal data be used in ways not initially foreseen in the consent granted and seeking consent for such new use is impractical, specific case-by-case arbitration implemented by ethics review boards or similar authorities may be appropriate. Such case-by-case arbitration should also be accompanied by a review taking into account legal aspects of the planned change of purpose.* (III.5)

Data portability is addressed in the section on "Sustainable Infrastructures", where the Recommendation calls on Adherents to "[e]ncourage private investment in research data infrastructures … while taking measures to facilitate their openness, reliability and integrity, and to protect the public interest over the long term by avoiding vendor lock in and ensuring data portability." (VII.2)

The OECD (2016[10]) *Health Data Governance Recommendation* recommends that

> *[c]onsent mechanisms should provide: Clarity on whether individual consent to the processing of their personal health data is required, and, if so, the criteria used to make this determination; what constitutes valid consent and how consent can be withdrawn; and lawful alternatives and exemptions to requiring consent, including in circumstances where obtaining consent is impossible, impracticable or incompatible with the achievement of the health-related public interest purpose, and the processing is subject to safeguards consistent with this Recommendation.* (III.5.a.i)

Where the processing of personal health data is based on consent, "such consent should only be valid if it is informed and freely given, and if individuals are provided with clear, conspicuous and easy to use mechanisms to provide or withdraw consent for the future use of the data." (III.5.a.ii)

The OECD (2013[13]) *Privacy Guidelines* and their Individual Participation Principle recommends that

> *[a]n individual should have the right: a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him; b) to have communicated to him, data relating to him within a reasonable time; … and d) to challenge data relating to him …*

The Individual Participation Principle is complemented by two other principles which refer to the consent of the data subject, in both cases, however, with an alternative legal basis. According to the Collection

Limitation Principle "[t]here should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject." According to the Use Limitation Principle, "[p]ersonal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9 except: a) with the consent of the data subject; or b) by the authority of law."

3.1.4. Support the adoption of technological and organisational measures to enhance control

The sphere of control of a data holder[10] or a data producer[11], as well as governments, over data is defined by technological boundaries (e.g. a personal device of a consumer), organisational boundaries (e.g. an organisation's information system) or national boundaries (e.g. a government's legislation without extraterritorial effects). All these boundaries affect the costs (including opportunity costs) of excluding others from accessing and using those data. Beyond this sphere of control, data holders and producers, as well as governments, progressively lose their ability to control how data are used and to oppose any such use. Therefore, once data leave their sphere of control, stakeholders rely mainly on available enforcement and redress mechanisms for their rights to be respected.

A range of technical and/or organisational measures helps preserve (or enhance) the sphere of control of data holders and producers even as data are accessed, used and re-used across society. This enables data openness, while minimising associated risks (OECD, 2022[19]). Policy makers therefore may consider the range of technological and organisational measures available for different contexts and different data types and promote their use through policies for data governance. This includes, for example, providing guidance on their implementation or pioneering their research and development (R&D).

Technical and organisational measures do not guarantee that the rights of stakeholders are always protected. Their use therefore requires an appropriate risk assessment. This, in turn, may need to be combined with other data governance mechanisms, such as enforceable legal commitments to e.g. not re-identify de-identified information (OECD, forthcoming[24]).

PETs as key technological measure

Privacy-enhancing technologies (PETs) are a prominent example of technological measures that can expand the sphere of control over data (OECD, forthcoming[24]). According to OECD (2002[25]), PETs "commonly refer to a wide range of technologies that help protect personal privacy [ranging] from tools that provide anonymity to those that allow a user to choose if, when and under what circumstances personal information is disclosed." Typically, they are not stand-alone tools but can be viewed as new functionalities to manage data. OECD (forthcoming[24]) differentiates between the following four classes of PETs:

- Data accountability tools provide enhanced control to the sources over how data can be gathered, access and used, or else greater transparency and immutability into tracking data access and use. These include, for example, accountable software systems that manage the use and sharing of data by controlling and tracking how data are collected and processed, and when they are used. Such a design aims, in part, to grant access with limitations attached to the data. In other words, data remain within the sphere of control wherever they flow.
- Data obfuscation tools, like encrypted data processing and distributed analytics tools (see bullets below), reduce the need for sensitive information to leave a data source's sphere of control, where the underlying data are kept and processed. Unlike other types of tools, however, obfuscation alters the data by adding "noise" or removing identifying details. Data obfuscation tools commonly include anonymisation and pseudonymisation techniques.[12]

- Federated and distributed analytics allows executing analytical tasks (e.g. training models) upon data that are not visible or accessible to those executing the tasks. In federated learning, a technique gaining increased attention, raw data are pre-processed at the data source. In this way, only the summary statistics/results are transferred to those executing the tasks. These results can then be combined with similar data from other sources for all of the data to be indirectly processed centrally. In distributed analytics, sensitive data never leave the custody of a data source but may be analysed by third parties who are members of a common network. Distributed analytics is increasingly applied in the health sector and was particularly important for accelerating the pace of global COVID-19 research (OECD, 2022[22]).
- Encrypted data processing tools allow running computations over encrypted data that are never disclosed. Unlike data obfuscation tools, the underlying data remain intact but hidden by encryption. Examples include homomorphic encryption through which a data processor can perform increasingly complex calculations over the encrypted data. It extracts an encrypted result that can only be unlocked with the original data source's cryptographic key.

Developments in data analytics and AI have combined with the increasing volume and variety of available data sets and the capacity to link these different data sets. This has made it easier to infer and relate seemingly non-personal or anonymised data to an identified or identifiable entity (OECD, 2019[3]; forthcoming[24]).[13] This, in turn, limits the use of some PETs as a single means of protection. Therefore, PETs-related initiatives need to be complemented by legally binding and enforceable obligations to protect the rights and interests of data subjects and other stakeholders as articulated in the OECD (2021[6]) *EASD Recommendation*.

Organisational measures

Organisational measures refer to institutional arrangements that may involve contractual arrangements (OECD, 2022[19]) (subsection 3.2.3). They are often used in combination with PETs to help manage and enhance control over data, enabling a more balanced approach to data openness. Organisational measures can vary greatly in their model and implementation.

Organisational measures play a key role where data are considered a collective resource of common interest and thus for the implementation of "data commons".[14] Data commons are formal or informal governance institutions to enable a viable and shared production and/or use of data that reflects the interests of stakeholders. They facilitate joint production or co-operation with suppliers, customers (consumers) or even potential competitors (co-opetition). Data commons have often led to open data arrangements. However, other arrangements that foresee restrictions on data access and sharing (conditioned data access and sharing arrangements) are increasingly in use as well.

Indeed, a wide range of approaches can be used to enable data commons to enhance individual and organisational control over data while unlocking the benefits of greater data use and sharing:

- **Independent ethics review bodies** (ERBs) are committees that review proposed research methods to ensure they are ethical. ERBs have been highlighted in some cases as critical trusted third parties, especially around access to and sharing and re-use of personal data. For example, in the scientific community, the evaluation of applications for access to publicly funded personal data for research can depend on ERBs. The OECD Global Science Forum concluded that ERBs can increase trust between parties with an interest in the use of personal data for research purposes, particularly in situations where consent for research use is impractical or impossible (OECD, 2016[5]).[15]
- **Data sandboxes** are isolated environments through which data are accessed and analysed. Analytic results are only exported, if at all, when they are non-sensitive. Data sandboxes typically require execution of the analytical code at the same physical location as the data. These sandboxes can be realised through technical means of various complexity. One example is isolated virtual machines that cannot be connected to an external network or federated learning solutions. It is also possible to maintain a physical presence within the facilities where data are located (OECD, 2019[3]).

- **Data intermediaries**: The OECD (2021[6]) EASD Recommendation defines data intermediaries as "service providers that facilitate data access and sharing under commercial or non-commercial agreements between data holders, data producers and/or users." As the Recommendation notes, data intermediaries can be data holders. By defining common legal and technical standards, data intermediaries facilitate data access and sharing (e.g. data spaces). However, they can also be trusted third parties (e.g. data trusts) that essentially act on behalf of other stakeholders, including data holders, data producers and/or users.
  - Data spaces have been proposed as a decentralised type of data intermediary. At a technical level, data spaces rely on common standards for accessing, linking, processing, using and sharing data between different endpoints where the data are located. Data spaces have been proposed as a solution, especially with respect to "industrial" or "non-personal" data. For example, Gaia-X is a European initiative that aims to develop a software framework for data governance. It would be implemented by cloud service providers based on common rules to enhance transparency, controllability, portability and interoperability across data and services.[16] This government-backed standard aims to establish an ecosystem in which data are made available, collated and shared in a trustworthy environment. In such an environment, generators of data maintain full control and visibility on the context and purpose for which other actors access data.
  - Data trusts are a more centralised type of data intermediary. Data trusts are institutional arrangements whereby a trusted third party (an informed person or organisation) takes on a fiduciary duty to steward/govern data use or sharing on behalf of its members in relation to third parties. This is intended as a means to increase access and sharing of the data, while safeguarding the rights and interests of the data holders (Hardinges, 10 July 2018[26]; Ruhaak, 2021[27]). Examples of data trusts include personal data stores or Personal Information Management Systems (PIMS). PIMS are service providers that enhance an individual's control over their personal data by choosing where and how they want their data stored, accessed or processed at a granular level. Examples also include trusted data-sharing platforms, where major data holders either designate an existing trusted organisation or create a new trusted organisation and platform to share data with third parties. In the United States, for example, health care and health insurance companies (e.g. Aetna, Humana, Kaiser Permanente and United Healthcare) designated the non-profit Health Care Cost Institute (HCCI) as a trusted organisation. After removing information about which company has provided the data, HCCI shares information about health care use and costs in the United States with selected research institutions (OECD, 2019[3]).

Governments can act as, or create, a trusted third party. In certain countries, national statistical offices have acted as in this capacity. In Australia, under the *Data Integration Partnership for Australia*, the government initiated a three-year AUD 131 million (around USD 90 million) investment to maximise the use and value of the government's data assets between 2017 and 2020 (Australian government, 2017[28]). Agencies in social services, health, education, finance and other government entities would provide data for linking and integration. Meanwhile, "sectoral hubs of expertise, independent entities that are funded by the Commonwealth" and denominated Accredited Integrating Authorities (AIAs), would enable integration of longitudinal data assets. These data would be "housed in a secure environment, using privacy preserving linking methods and best practice statistics to link social policy and business data" (Productivity Commission, 2017[29]). The Australian Bureau of Statistics – the national statistical office – was the country's first institution to be recognised as an AIA.[17]

All these organisational measures have strengths and weaknesses to address concerns around data access and sharing. On their own, they often cannot provide a solution. A combination of technological and organisational measures will be necessary to address stakeholders' concerns and risks while enabling co-operation. Policy makers could provide incentives or directly experiment with some organisational measures to ensure optimal levels of data sharing occur based on technological and organisational possibilities.

How OECD legal instruments promote adoption of technological and organisational measures that help reduce the risk of loss of control

The OECD (2021[6]) *EASD Recommendation* calls on Adherents to

> *[e]ncourage competition-neutral data-sharing partnerships, including Public-Private Partnerships (PPPs), where data sharing across and between public and private sectors can create additional value for society. In so doing, Adherents should take all necessary steps to avoid conflicts of interest or undermining government open data arrangements or public interests.* (III.b)
>
> *[e]mpower individuals, social groups and organisations through appropriate mechanisms and institutions such as trusted third parties that increase their agency and control over data they have contributed or that relate to them, and enable them to recognise and generate value from data responsibly and effectively.* (III.d)
>
> *[f]oster the adoption of conditioned data access and sharing arrangements with the use of technological and organisational environments and methods, including data access control mechanisms18 and privacy enhancing technologies, through which data can be accessed and shared in a safe and secure way between approved users, combined with legally binding and enforceable obligations to protect the rights and interests of data subjects and other stakeholders.* (V. d)

The OECD (2021[9]) *Research Data Recommendation* recommends that

> *[i]n cases where access needs to be partially or totally restricted to conform to legal rights, ethical principles and/or to protect legitimate private, public, or community interests, and with the ultimate objective of facilitating access which is as open as possible, [Adherents should] foster more limited forms of access, such as access to aggregated or de-identified data, restricted access within safe and secure environments to certified users with clearance adapted to the sensitivity of data, or access via analyses that share only de-identified results.* (II.2.a)

The OECD (2016[10]) *Health Data Governance Recommendation* recommends that governments establish and implement national health data governance frameworks that provide for "control and safeguard mechanisms". Such mechanisms should "include technological, physical and organisational measures designed to protect privacy and security while maintaining, as far as practicable, the utility of personal health data for health-related public interest purposes" (III.11.v). This should be done with

> *[m]echanisms that limit the identification of individuals, including through the de-identification of their personal health data, and take into account the proposed use of the data, while also allowing re-identification where approved. Re-identification may be approved to conduct future data analysis for health system management, research, statistics, or for other health-related public interest purposes; or to inform an individual of a specific condition or research outcome, where appropriate.* (III.11.v.a.) *… Where practicable and appropriate, … alternatives to data transfer to third parties, such as secure data access centres and remote data access facilities.* (III.11.v.c.)

The OECD (2013[13]) *Privacy Guidelines* in Part Five on National Implementation recommend that "[i]n implementing these Guidelines, Member countries should: … consider the adoption of complementary measures, including … the promotion of technical measures which help to protect privacy."

3.1.5. Enhance technical interoperability for data openness

Technical interoperability refers to the ability of two or more systems or applications to exchange information and to mutually use that information through seamless exchanges, updates or transfers (ISO, 2017[30]). If technical specifications are incompatible, however, it can lead to fragmentation across organisations, sectors and countries. Indeed, in some cases this result is inevitable. With respect to data, technical interoperability is usually associated with the need for machine readability, common standards and other interoperable technical specifications, and metadata. These enable what is commonly called the FAIR (findability, accessibility, interoperability and re-use) and the correct interpretation and analysis of data across systems.

Syntactic and semantic interoperability

Commonly used machine-readable formats are not enough to guarantee technical interoperability (OECD, 2019[3]; 2020[16]). These formats may enable *data syntactic interoperability*, i.e. the transfer of "data from a source system to a target system using data formats that can be decoded on the target system." However, they do not guarantee *data semantic interoperability*, defined as "transferring data to a target such that the meaning of the data model is understood within the context of a subject area by the target." Both syntactic and semantic interoperability are needed for data interoperability and to enable data openness.

The private sector, including data intermediaries (subsection 3.1.4), has played an important role to develop and promote data-related standards and other technical specifications. Data intermediaries often provide their data in common data formats that assure both syntactic and semantic interoperability – driving their adoption across industries. In so doing, they can define *de facto* standards for their industry. One example is Google's General Transit Feed Specification (GTFS), a common format for public transportation schedules and associated geographic information (OECD, 2019[3]).

Governments nonetheless play a role to promote development and adoption of interoperable specifications for effective data access, sharing and use, and to ensure these specifications maximise trust. These specifications include common standards for data formats and models, as well as open-source implementations (OECD, 2021[6]). Promotion of their adoption can enhance technical interoperability and be an effective strategy for fostering competition between data-driven services and platforms. This is especially the case where limited interoperability – through, for example, closed application programming interfaces – may serve as an entry barrier (OECD, 2021[2]).

Data quality, which can be understood as "fitness for use" in terms of the ability of the data to fulfil user needs, is an important factor for technical interoperability. In this respect, "government initiatives that focus on the production of good-quality data can also contribute to greater interoperability, sharing and openness in later stages" of the data value cycle (OECD, 2019[31]).

Impacts of interoperability on investment

Although achieving technical interoperability across societies and economies represents a great opportunity, it can also have unintended consequences. Interoperability requirements, for example, may have a negative effect on firms' willingness to invest in data. They may perceive that lower switching costs enabled by interoperability creates a risk of losing their user base. In addition, interoperability standards that are too prescriptive and have a high cost of compliance can harm competition and innovation. This is especially the case for small and medium-sized enterprises (SMEs) as they may lack capacity to develop and maintain interoperability specifications (Competition Bureau Canada, 2022[32]). Therefore, interoperability needs to be considered in connection with the cross-cutting tension *Incentivising investments in data and their effective re-use* (section 3.3).

How OECD legal instruments promote technical interoperability of data

The OECD (2021[6]) *EASD Recommendation* recommends that Adherents "[f]oster where appropriate the findability, accessibility, interoperability and reusability of data across organisations, including within and across the public and private sectors" (VIII.b). In particular, Adherents should "[a]ssess and, whenever possible, promote development and adoption of interoperable specifications for effective data access, sharing and use, including common standards for data formats and models as well as open source implementations."

The OECD (2021[9]) *Research Data Recommendation* recommends that Adherents

> *[p]romote interoperability by requiring the use of semantic (including ontologies and scientific terminology), legal (rights of use), and technical (such as machine readability) standards as appropriate"* (IV. 3) *[and] take necessary measures to support development and maintenance of sustainable infrastructures to support the findability, accessibility, interoperability, and reusability of research data and other research-relevant digital objects from public funding free of charge at the point of use* (VII). *[To this end, Adherents should] support efforts to improve interoperability among global research infrastructures to leverage national investments and innovation, and to encourage interdisciplinarity.* (VII.b)

The OECD (2016[10]) *Health Data Governance Recommendation* recommends that national health data governance frameworks should provide for "cooperation among organisations processing personal health data, whether in the public or private sectors" (III.2.i). To achieve this, it indicates they should "encourage common data elements and formats; quality assurance; and data interoperability standards." Further,

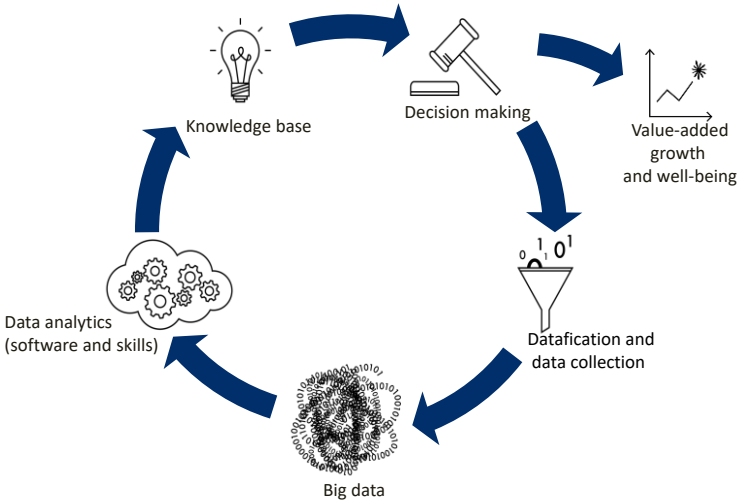
> *[It] recommends that governments engage with relevant experts and organisations to develop mechanisms consistent with the principles of this Recommendation that enable the efficient exchange and interoperability of health data whilst protecting privacy, including, where appropriate, codes, standards and the standardisation of health data terminology.*

The OECD (2008[20]) *PSI Recommendation* in its principle on "[n]ew technologies and long-term preservation" recommends that Member countries aim for "improving interoperable archiving, search and retrieval technologies and related research including research on improving access and availability of public sector information in multiple languages, and ensuring development of the necessary related skills." They should also aim for "avoiding fragmentation and promote greater interoperability and facilitate sharing and comparisons of national and international datasets and striving for interoperability and compatible and widely used common formats."

3.2. Managing overlapping and conflicting interests and regulations related to data governance

The second fundamental policy tension and objective in data governance relates to balancing the interests of different stakeholders and policy communities, while ensuring consistency across different policy and regulatory frameworks. This challenge reflects that multiple parties with potentially conflicting interests are involved at different phases of the data value cycle (data collection, analysis, use, deletion) (Figure 3.3). It is also the result of different policy communities being involved in the governance of data. This engagement gives rise to multiple, and sometimes overlapping, policy and regulatory frameworks (sectoral or cross-sectoral, as well as national and international) that focus on concerns mainly relevant to these policy communities.

Figure 3.3. The data value cycle



Source: OECD (2015[33]), *Data-Driven Innovation: Big Data for Growth and Well-Being*, http://dx.doi.org/10.1787/9789264229358-en.

Data can pertain to different domains and be subject to different policy and regulatory frameworks. Privacy and data protection frameworks are the most prominent legal and regulatory framework that define control and use rights over (personal) data. These include an individual's right to restriction of processing and, according to some frameworks, a right to data portability (Banterle, 2018[34]; Drexl, 2018[35]; Purtova, 2017[36]). On the other hand, data (including personal data) collected by an organisation will typically also be governed by IPRs. These are essentially copyrights in respect to the collection of data, sui generis database rights in some jurisdictions, and trade secrets for confidential business and technical data.[19]

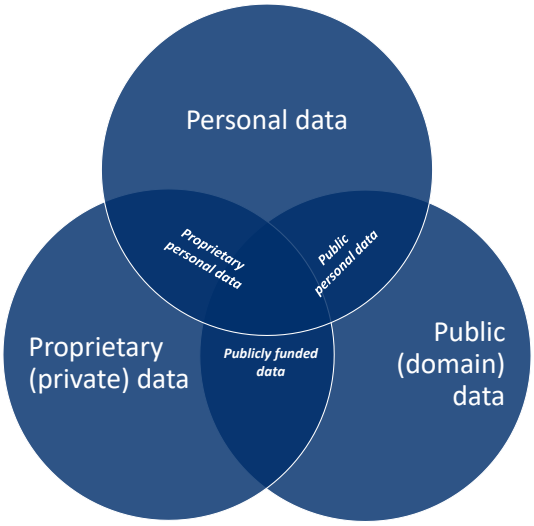
However, in the case of personal data, control or "ownership" rights of the organisation will hardly be comparable to other (intellectual) property rights. This is because most privacy regulatory frameworks give data subjects control rights over their personal data, which may interfere with "the full right to dispose of a thing at will" (Determann, 2018[37]). Therefore, in the case of personal data, no single stakeholder will have exclusive access and use rights. Different stakeholders will typically have different rights and obligations depending on their roles, which may be affected by different legal and regulatory frameworks.

Figure 3.4 illustrates the overlapping domains of data and the relevant policy and regulatory frameworks that can affect each domain differently, yet with overlaps:

- **the personal domain**, which covers data "relating to an identified or identifiable individual" (personal data) for which privacy and personal data protection frameworks apply
- **the private domain**, which covers proprietary data protected by IPRs or by other access and control rights (e.g. provided by contract, cyber-criminal or competition law), and for which there is usually an economic business interest to the exclusion of others
- **the public domain**, which covers data not protected by IPRs or rights with similar effects, but that most importantly lie in the "public domain" (understood more broadly than to be free from copyright protection), thus in principle available to access and re-use, as well as data for which there are a public interest.

Each of these domain-specific frameworks are important to foster trust in the data ecosystem. However, the "intricate net of existing legal frameworks" (Determann, 2018[37]) creates legal uncertainties. These concerns have been identified as a possible reason for low adoption of data analytics and related digital technologies such as AI. At the time, they may explain scarce use of new data governance approaches such as data portability (OECD, 2019[38]; 2019[3]; Business at OECD, 2020[39]).

Figure 3.4. The personal, private and public domains of data



Source: OECD (2019[3]), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-Use Across Societies*, http://doi.org/10.1787/276aaca8-en.

Data governance has proven to be challenging where public interest in data is concerned, especially when such interest "overlaps" with the interests in proprietary and personal data (the centre of the Venn diagram in Figure 3.4). This is the case with data of public or general interest, a relatively new class of data. Some countries[20] have started to recognise and regulate these data to fulfil societal objectives that otherwise would be impossible or too costly to fulfil. These objectives can include development of national statistics, development and monitoring of public policies, tackling of health care and scientific challenges and, in some cases, provision of public services (OECD, 2022[19]).

In this context, the challenge is how to disentangle and reconcile these multiple interests, and policy and regulatory frameworks, through clear policies that best reflect societal values and the public interest. Approaches to this trade-off aim to strike an effective balance between stakeholders' interests and ensure coherence across different policies and regulations. These include identifying and considering the contribution of different stakeholders throughout the data value chain, including by promoting multi-stakeholder engagement and participation (subsection 3.2.1); adopting a whole-of-government approach through national data strategies and cross-agency co-operation to help reconcile different domestic frameworks affecting data governance (subsection 3.2.2); leveraging contracts to address legal uncertainties in respect to data governance (subsection 3.2.3); and promoting international policy and regulatory co-operation to reconcile data governance differences across countries and enable cross-border data flows (subsection 3.2.4).

3.2.1. Promote multi-stakeholder engagement

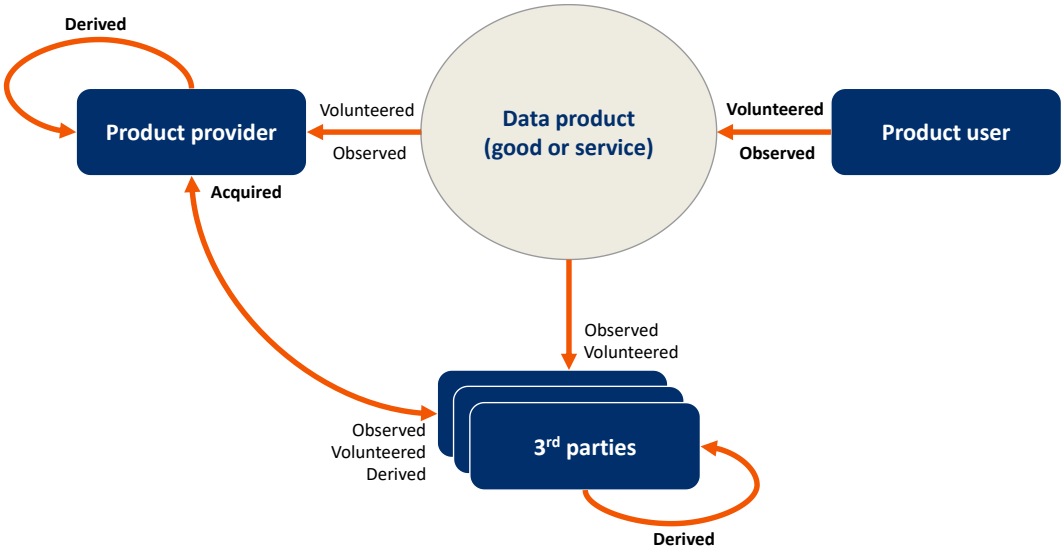
Data and the value derived from their use are often (co-)created through interactions of various parties in the global data ecosystem,[21] sometimes even without their awareness. Multiple stakeholders are involved in the collection, control and use of data at different stages of the data value chain. This engagement provides the rationale for each of them to be claiming rights and interests with respect to those data. However, some stakeholder interests may overlap or even conflict. In the case of personal data, individuals and businesses may have opposing interests regarding their sharing and re-use. As another example, businesses may also object to mandatory sharing of their data with public authorities for regulatory purposes.

The contributions of different stakeholders throughout the data-driven value creation process

Determining fair distribution of risks, benefits and liabilities across stakeholders for data governance is a multi-pronged process. One important first step is to identify and consider the respective contributions of different stakeholders throughout the data-driven value creation process (from data collection to processing through to use and re-use) and the different types of data that may be involved thereby.

Figure 3.5 presents a stylised process of how a product user (e.g. a consumer) typically interacts with a data product (e.g. an online service or portable smart device) provided by a product provider (e.g. a business). The data product (i) observes the activities of its user, creating *observed data*;[22] and/or (ii) collects input data volunteered by its user, i.e. *volunteered data.*[23] These data can then be accessed for further processing (iii) by the product provider and/or by a third party granted access, leading to *derived data.*[24] Derived data can also be enriched when a set of observed and/or volunteered data is combined with a set of *acquired data*[25] from (other) third parties (OECD, 2019[3]).

Figure 3.5. Data products and the different ways in which data originate



Note: Arrows represent potential data flows between the different actors and a data product (good or service). The type of data is highlighted in bold to indicate the moment at which the data are created.
Source: OECD (2019[3]), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-Use Across Societies*, http://doi.org/10.1787/276aaca8-en.

Observed, volunteered, derived and acquired data

This differentiation between observed, volunteered, derived and acquired data highlights the contribution of various stakeholders to the data value cycle. As such, it indicates the potential interests and claims of stakeholders with respect to those different types of data. Product users are most aware of the data they have voluntarily provided. Consequently, they typically tend to have stronger interests in volunteered data. For their part, product providers will have a particularly strong interest in derived and acquired data given their respective creation and acquisition directly involved financial and non-financial resources.

The differentiation between observed, volunteered, derived and acquired data has been reflected in recent policy considerations on the allocation of access and use rights on data. As countries increasingly adopt data portability regimes, to what extent should data portability rights apply to all types of personal data? Some data held by organisations may have been inferred or derived by the organisations. In other cases, users have been the main contributors (OECD, 2021[17]).

Recent policy decisions treat this question differently. The "Right to Data Portability" (Art. 20) of the European Union (2016[40]) General Data Protection Regulation (GDPR), for instance, only applies to personal data "provided by" the data subject with consent or under contract that is electronically processed. The (former) Article 29 Working Party indicates in its guidance that the definition of provided personal data should include data volunteered by the individual, as well as *observed* by virtue of their use of the service or device. However, it should not include personal data that are inferred or derived (OECD, 2015[33]).

The differentiation between observed, volunteered, derived and acquired data can also shed light on other aspects of interest to policy design. For example, it can help determine what product users (such as data subjects) may reasonably know about the scale and impact of activities by service providers with personal data about them. This is relevant, for example, when designing policies that seek to protect privacy by strengthening the control capacity of the product users such as through individual consent requirements (subsection 3.1.3).

Multi-stakeholder engagement and participation

The differentiation between observed, volunteered, derived and acquired data is crucial to disentangle the contribution of stakeholders at different phases of the data value cycle conceptually. However, multi-stakeholder engagement and participation is another fundamental instrument to disentangle the multiple interests entrenched in data governance. Available evidence shows that engagement about data governance with stakeholder communities can help better identify and make explicit their various interests, and better define responsibilities and acceptable risk levels across the data ecosystem (Frischmann, Madison and Strandburg, 2014[41]; OECD, 2021[4]; 2019[3]; 2016[5]). In this sense, it also helps build trust in data governance and the data ecosystem, strengthening subsequent buy-in and compliance from all stakeholders.

To be effective, multi-stakeholder engagement and dialogues require establishing whole-of-society, open and inclusive processes where:

- all relevant stakeholders are represented, their roles, responsibilities and modalities of participation are identified, and their various interests recognised and made transparent
- appropriate mechanisms ensure that all stakeholders' interests are considered in the design, implementation and monitoring of data governance frameworks, that stakeholder feedback is given appropriate attention and consideration and that a rationale is given where feedback is not taken on board (OECD, 2019[31]).

How OECD legal instruments help disentangle the various interests of stakeholders

The OECD (2021[6]) *EASD Recommendation* recommends that Adherents should

> *[p]romote inclusive representation of and engage relevant stakeholders in the data ecosystem – including vulnerable, underrepresented, or marginalised groups – in open and inclusive consultation processes during the design, implementation, and monitoring of data governance frameworks* (III.a) *[and] work together with key stakeholders to clearly define the purpose of [data access and sharing] arrangements and identify data relevant to these purposes, taking into account their benefits, costs, and possible risks.* (IV.a)

The OECD (2021[9]) *Research Data Recommendation* calls on Adherent to

> *[c]onsult with communities of stakeholders about open access to, sharing of, and re-use of research data and other research digital objects from public funding for reinforcing trust. This should include establishing open and inclusive processes that ensure equitable representation of stakeholder groups and consideration of their respective needs* (III.4) *[and] to prioritise, in consultation with stakeholders at the national and international level, research data and other research-relevant digital objects from public funding for short-, medium-, or long-term preservation.* (VII.1.a)

The OECD (2016[10]) *Health Data Governance Recommendation* calls on governments to establish and implement a national health data governance framework that provides for

> *[e]ngagement and participation notably through public consultation, of a wide range of stakeholders with a view to ensuring that the processing of personal health data under the framework serves the public interest and is consistent with societal values and the reasonable expectations of individuals for both the protection of their data and the use of their data for health system management, research, statistics or other health-related purposes that serve the public interest.* (III.1)

The OECD (2014[12]) *Digital Government Recommendation* recommends that governments develop and implement digital government strategies that "encourage engagement and participation of public, private and civil society stakeholders in policy making and public service design and delivery" (II.2). This is to be done by

> *[a]ddressing issues of citizens' rights, organisation and resource allocation, adoption of new rules and standards, use of communication tools and development of institutional capacities to help facilitate engagement of all age groups and population segments, in particular through the clarification of the formal responsibilities and procedures* (II.2.i) *[and] identifying and engaging non-governmental organisations, businesses or citizens to form a digital government ecosystem for the provision and use of digital services. This includes the use of business models to motivate the relevant actors' involvement to adjust supply and demand; and the establishment of a framework of collaboration, both within the public sector and with external actors.* (II.2.ii)

3.2.2. Support a whole-of-government approach

As highlighted above, data are typically governed by multiple, and sometimes overlapping, policy and regulatory frameworks. These frameworks include, most prominently, privacy and personal data protection frameworks and IPR frameworks, which can affect conditions for access, sharing and processing of data differently. Other relevant policy and regulatory frameworks include sector- and domain-specific regulations such as those targeting the financial sector (open banking), the public sector (open government data) or science and research (open science). Conditions for access, sharing and processing of data may also be defined by ex-post enforcement measures. In some cases, for example, data portability may be mandated as a competition enforcement mechanism (OECD, 2021[2]).

Against this background, policy makers must identify and address potential overlaps and incoherencies across their domestic policy and regulatory frameworks that directly or indirectly govern data (e.g. between personal data protection and sectoral data frameworks). This should ensure a proper interplay across the frameworks and provide legal certainty for stakeholders. For this purpose, policy makers need to acknowledge the strengths and weaknesses of different data governance approaches. Unlike general approaches that apply across sectors, for example, approaches specific to data types or sectors tend to better address the legal, organisational and technical needs of individual sectors and better protect specific categories of data. At the same time, varying standards and requirements across sectors or by data types can create challenges for data sharing within and across sectors. For example, due to the complexities of sharing, accessing and processing, databases may contain both personal and non-personal data that are subject to different rules.

Policy makers need strategic whole-of-government approaches to foster data access and sharing within and across society, while ensuring policy coherence across data governance frameworks. Such approaches should take advantage of both the general and sector/data-type specific approaches to data governance. National data strategies, in combination with sectoral data strategies, which integrate cross-cutting economic, social, cultural, technical and legal governance issues, could be instrumental. This combination of strategies should create the conditions for effective data governance frameworks to better protect the rights and interests of individuals and organisations. At the time, it should provide the flexibility for all to benefit from data openness (OECD, 2022[19]). Such national and sectoral data strategies would need to integrate, or be reinforced by, efforts to foster cross-agency co-operation.

National and sectoral data strategies

National and sectoral data strategies could help address many of the policy issues highlighted above in a comprehensive manner by incorporating a whole-of-government perspective. Many countries have adopted national digital economy and national AI strategies. However, equivalent data strategies are emerging spurred by the recognition of their potential to support the data economy strategically and the responsible use of data across society (OECD, 2022[19]).

National data strategies are cross-sectoral by nature. In many instances, they are designed to help reach higher level objectives. These could include gross domestic product growth, productivity, well-being and/or combatting climate change and fostering sustainable development. To achieve these strategic objectives, many national data strategies build on specific strengths of the country. For example, they often integrate pre-existing national digital economy strategies (Gierten and Lesher, 2022[42]), national broadband strategies, digital government strategies and national AI strategies, which national data strategies often complement. In turn, national data strategies are often complemented by sector-specific data strategies. The most prominent sectoral data strategies relate to health but also science and research, as well as smart cities, smart transportation and smart energy systems.

Approaches to governing national and sectoral data strategies vary across countries, but commonalities can be found. In the case of sectoral data strategies, development, co-ordination and monitoring fall essentially under the ministry in charge of the sector. For national data strategies, a few countries (still the exceptions so far) have appointed a high-level government official with leading development, co-ordination and monitoring. Several have tasked a ministry, ministerial position or other body dedicated to digital affairs. Not surprisingly therefore, several ministries, bodies or institutions may be involved when implementing national data strategies. In some cases, multiple stakeholders are also involved. Indeed, in almost all countries, multiple private and public stakeholders and bodies contribute input to the development of national and sectoral data strategies (OECD, 2022[19]).

Cross-agency co-operation

Interdisciplinary and cross-agency co-operation in both policy making and enforcement are needed in cases where data governance relies on overlapping policy and regulatory frameworks. Data portability initiatives, for example, involve personal data that are also important for businesses and can alter competition conditions. Therefore, governance intersects with competition, privacy and consumer protection. Even if only one of these three areas is the primary motivation for a data portability initiative, the policy will likely implicate the others. Multiple regulatory domains may also be implicated when data governance is implemented at a sectoral level. Open Banking initiatives, for example, involve financial market authorities, as well as privacy enforcement authorities.

Privacy enforcement authorities (PEAs) have long raised the need for closer dialogue between regulators and experts across jurisdictions but also across policy communities. Such a dialogue could achieve more effective competition and consumer protection enforcement. At the same time, it might stimulate the market for privacy-enhancing services (EDPS, 2016[43]; 2014[44]). Multiple OECD reports have discussed enhanced co-operation and co-ordination across agencies (OECD, 2020[45]; 2020[46]; 2019[47]; 2018[48]; 2015[33]; forthcoming[49]). The review of the OECD (2007[50]) *Recommendation on Cross-Border Co-operation in the Enforcement of Laws Protecting Privacy* found that PEAs are collaborating "generally with regulators responsible for consumer protection, competition/antitrust, cybersecurity, telecommunication, financial regulation, public health and transportation/infrastructure" (OECD, forthcoming[49]). The review notes that

> *[c]o-operation with other regulators is often necessary when a case involves broad policy areas, when data protection obligations are provided by sectoral laws and enforced by sectoral regulators, when developing legislative or administrative measures in relatively new policy areas (e.g. data portability), and/or where PEAs are consulted for regulatory opinions which may concern privacy and data protection.*

Good practices offer a starting point for other countries

Among good practices, the Office of Australian Information Commissioner (OAIC) has been working with the Australian Competition and Consumer Commission (ACCC) on a range of projects. One example is the development of the "Consumer Data Right" (Australia's data portability initiative) (OAIC, 2020[51]). Another example is the OAIC's engagement in the ACCC's Digital Platforms Inquiry. This inquiry examined the impact of digital platforms on competition in the media and advertising services markets, recommending reform of Australia's privacy framework. In support of this recommendation, the OAIC provided input from the viewpoint of a PEA acting internationally. This aimed to ensure the interoperability of Australia's data protection laws globally and optimal regulatory outcomes in the public interest (OAIC, 2019[52]).

Another example is the collaboration between the French data protection authority (CNIL) and the Directorate-General for Competition, Consumer Affairs and Fraud Control (DGCCRF) (the French authority responsible for consumer protection). Together, they have processed personal data by social networks, deceptive marketing practices related to compliance with the GDPR and the use of personal data in electronic commerce (CNIL, 2019[53]).

Despite the increasing number of collaborations, policy makers and regulators may still face challenges. OECD (2021[54]) notes that co-ordination for competition and consumer policy issues and enforcement is generally more straightforward than for those involving privacy enforcement. This is because a common agency has responsibility for both competition and consumer issues in more than 30 jurisdictions. In Germany, for example, legislation provides the basis for co-operation between these authorities (Stauber, 2019[55]).

Few agencies assume responsibility for privacy in addition to competition and consumer protection. In this regard, the United States' Federal Trade Commission is one of the exceptions. In addition, few countries have legislation to allow cross-disciplinary collaborations between privacy enforcement and other authorities similar to those between competition and consumer protection regulators.

However, there are promising developments in certain countries. The Information Commissioner's Office (ICO) of the United Kingdom, for instance, has established Memoranda of Understanding with the Financial Conduct Authority (FCA) and the Competition and Markets Authority (CMA) on matters related to data protection and other areas of mutual interest. The agreements set out a range of ways to co-operate, notably through regular communications and consultations. This may include sharing information about investigations, relevant action and relevant information regarding data breaches, fraud and criminal activity (ICO/CMA, 2021[56]; ICO/FCA, 2019[57]). The CMA and the ICO recently released a joint statement on their interactions and next steps, highlighting synergies and potential tensions between different policy areas (CMA/ICO, 2021[58]).

How OECD legal instruments support whole-of-government and cross-agency co-operation for data governance

The OECD (2021[6]) *EASD Recommendation* recognises

> *[t]he need for greater coherence across policy approaches to enhancing access to and sharing of all forms of data, including personal data, research data, public sector data, and that the development of general principles and policy guidance will support such coherence [and] that data access and sharing arrangements, including government access to proprietary and personal data held by the private sector, may involve activities governed by specific national and international legal frameworks that need to be taken into account in such arrangements.*

More specifically, the *EASD Recommendation* invites to

> *[a]dopt a strategic whole-of-government approach to data access and sharing to ensure that data access and sharing arrangements help effectively and efficiently meet specific societal, policy, and legal objectives that are in the public interest* (IV). *In particular, Adherents should:… [a]dopt and regularly review coherent, flexible, and scalable data governance frameworks – including national data strategies, which integrate cross-cutting economic, social, cultural, technical, and legal governance issues – in order to foster data access and sharing within and across society, public and private sectors, and jurisdictions.* (IV.b)

The OECD (2021[9]) *Research Data Recommendation* in its section on "Responsibility, Ownership, and Stewardship" recommends to

> *[t]ake measures to ensure a clear delineation and allocation of responsibility, ownership, and stewardship for access to publicly funded research data and other research-relevant digital objects from public funding across the research data ecosystem, while also tailoring and implementing licensing and other management of intellectual property rights to optimise scientific discovery and innovation and protect research data and digital object producers' rights.* (V)

3.2.3. Promote diverse tools for data governance

As a response to the complexity resulting from the "intricate net" of legal frameworks applying to data, many stakeholders recognise contract law as a tool for determining rights and obligations related to data control, access and re-use. This is especially the case in a business-to-business (B2B) context.

Contracts can allow adaptation to a specific context. However, parties may have unequal positions in the negotiation of terms and conditions. This could relate, for example, to market power or information asymmetries. The weaker parties are typically individuals (consumers) and SMEs.[26] These dynamics, paired with the difficulty of pricing data, may present high transaction costs when negotiating terms and conditions for data access and use. This can create a barrier to use of contracts for data governance overall.

Contract models from government and industry

Some governments and private sector actors have recognised both the importance of contracts and their related issues. In response, they provide guidance and/or voluntary model contracts or contractual clauses for data sharing agreements. These clauses are the default position for parties to consider when negotiating their data sharing agreements for complete, clear and fair contracts (European Commission, 2018[59]). Parties are free to deviate from the proposed clauses at their mutual agreement. Nevertheless, by providing a starting point that seeks to balance power or information asymmetries, contract guidelines and model contracts are seen as a promising means to assure fairer terms and conditions for data access, sharing and re-use.

In Japan, for example, the Ministry of Economy, Trade and Industry has formulated the *Contract Guidelines on Utilisation of AI and Data.* It summarises issues for private businesses to consider when drafting a contract related to data re-use or development and use of AI-based software (METI, 2019[60]).

Similarly, industry-developed voluntary codes of conduct and public sector data ethics frameworks guide the conduct of a specific sector, firm or public contractor to promote an environment of trust. Such codes generally focus on transparency, disclosure and consent. In the agricultural sector, codes for use of agricultural data complement legislation and encourage best practice in farm data management. They are seen as a means of improving transparency and fairness in agricultural data contracts. As such, they can be a viable option to generate trust in the relationship between farmers and technology providers. These codes are not without challenges, chiefly due to their voluntary nature. However, when they are developed by and for farmers, potentially with government seed funding, they can have broad buy-in (Jouanjean et al., 2020[61]).

As another example from the private sector, Microsoft published three data sharing agreements in July 2019 as templates: the *Open Use of Data Agreement* (O-UDA), the *Computational Use of Data Agreement* (C-UDA) and the *Data Use Agreement for Open AI Model Development* (DUA-OAI). These were complemented by a fourth in November 2019: the *Data Use Agreement for Data Commons* (DUA-DC) (Microsoft, 2019[62]).

In terms of data ethics frameworks, public procurement processes could enforce alignment with value-based data governance practices (OECD, 2021[4]). For instance, the UK Data Ethics Framework is used for assessing the ethical standards of projects, partners and suppliers in the Crown Commercial Service Dynamic Purchasing System. Suppliers applying to join the system must commit only to bidding where they are capable and willing to deliver the ethical and technical dimensions of a tender (where a buyer has stated there is an ethical dimension to their procurement). Following the bid, suppliers need to provide evidence on how they have been following the Data Ethics Framework during implementation of the product or a service (OECD, 2021[4]).

How OECD legal instruments encourage to leverage contracts to clarify and strengthen data governance

The (2021[6]) *OECD EASD Recommendation* recognises the need to "promote, where appropriate, self- or co-regulation mechanisms – including voluntary guidance, codes of conduct and templates for data access and sharing agreements – that provide legal flexibility while ensuring that all relevant stakeholders have certainty as to applicable laws and regulations." (VI.b)

The OECD (2013[13]) *Privacy Guidelines* under Part Five on National Implementation calls on Member countries to "encourage and support self-regulation, whether in the form of codes of conduct or otherwise." (19.d)

3.2.4. Reconcile data governance frameworks across countries

Countries develop different approaches to data governance, reflecting legitimate cultural and policy preferences. This means that as data move across borders to support the global digital economy, differing policy and regulatory frameworks might apply. Indeed, countries consider "uncertainty regarding legal privacy regimes", followed by "incompatibility of legal regimes" as the top challenges to transborder data flows (OECD, 2021[63]).

In this context, policy makers in different countries need to develop policies that facilitate cross-border data sharing with the desired oversight and/or protection (OECD, 2022[1]). In the context of privacy and data protection, the concept of interoperability has gained prominence to reconcile differences in personal data governance approaches across countries ("privacy interoperability"). This is reflected in the revised OECD (2013[13]) *Privacy Guidelines*, under Part Six on "International Co-Operation and Interoperability". These recommend that "Member countries should encourage and support the development of international arrangements that promote interoperability among privacy frameworks that give practical effect to these Guidelines."

Robinson, Kizawa and Ronchi (2021[64]) propose a working definition of privacy interoperability with the aim of moving towards a shared definition. They call it "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." In principle, there is broad agreement on the importance of privacy interoperability. Achieving it in practice, however, is less well understood.

Casalini, López González and Nemoto (2021[65]) identify several ways in which regulatory frameworks seek to be more interoperable with others, especially in the field of privacy. They point, for example, to contractual clauses, adequacy decisions, and inter-governmental arrangements and agreements of different nature, depth and scope. These different ways to govern cross-border data flows can leverage elements of convergence between domestic approaches (partly as a result of those inter-governmental arrangements as well). Governments have developed several approaches to enable the flow of data, especially personal data, across borders while protecting other public policy objectives (OECD, 2022[1]):

- **Unilateral mechanisms** embedded in domestic policy and regulatory frameworks enable the transfer of data to other countries provided certain safeguards are in place (e.g. through contracts). This approach is developed in several privacy frameworks, such as Argentina's Personal Data Protection Act 25.326 (PDPA) (Ley de Protección de los Datos Personales); Chapter V of the European Union (EU) General Data Protection Regulation (GDPR); or New Zealand's Privacy Act 2020.
- **Plurilateral arrangements** in the field of privacy and personal data protection aim to generate consensus around the transfer of specific types of data. They could also build networks of enforcement that promote 'trust' in cross-border data flows among enforcers irrespective of the data's geographical location (OECD, forthcoming[49]). This is the case of arrangements such as the OECD *Privacy Guidelines* or the Asia-Pacific Economic Cooperation Cross-Border Privacy Rules System.
- **Trade and digital economy agreements** are increasingly addressing issues around data flows (in the context of both personal and non-personal data). They combine provisions that enshrine the free flow of data between parties to the agreement, subject to certain exceptions and provided those countries maintain frameworks for data protection. Notable examples include the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP) FTA and the United States-Mexico-Canada Agreement. Digital Economy Agreements are also increasingly of interest. These include the Digital Economy Partnership Agreement between Chile, New Zealand and Singapore or the United Kingdom-Singapore Digital Economy Agreement.
- **Standards and technology-driven initiatives** are non-regulatory tools developed by non-governmental and private sector organisations to better handle issues around cross-border data transfers, often in the context of privacy and security protection. Data intermediaries that leverage PETs are a typical example of such initiatives (subsection 3.1.8).

Each approach addresses governance of cross-border data flows differently. In fact, these approaches are not mutually exclusive but potentially complementary. Countries can simultaneously use different approaches for different purposes, partners, types of data and situations (OECD, 2022[1]). For example, Casalini, López González and Nemoto (2021[18]) show that trade agreements since 2008 increasingly feature binding commitments on data flows. These are combined with provisions on the need for domestic privacy and data protection legislation (including references to inter-governmental arrangements outlined below). In this sense, they converge and complement inter-governmental arrangements in the field of privacy and personal data protection.

How OECD legal instruments address cross-border flow of data

The OECD (2021[6]) *EASD Recommendation* recommends to "further improve conditions for cross-border data access and sharing with trust" (VII). To this effect, Adherents should

> *[a]ssess, and to the extent possible minimise, restrictions to cross-border data access and sharing … taking into account the need to ensure respect for fundamental rights and vital interests …* (VII.a) *[and] ensure that measures that condition cross-border data access and sharing are non-discriminatory, transparent, necessary, and proportionate to the level of risk, taking into account, among others, the sensitivity of the data, the purpose and context of data access, sharing, and use, and the extent to which measures are in place to enforce accountability irrespective of the jurisdiction in which the data is stored*. (VII.b)

The OECD (2021[9]) *Research Data Recommendation* "[w]ith specific regard to sensitive data …, including personal data", recommends that Adherents "share good practices and experiences in enhancing access to them across borders, recognising that the data may themselves have to reside in the host originating country" (IX.4.). In this respect, Adherents should "explore interoperability of legal and ethical frameworks to enhance data access across borders while protecting legitimate private, public, or community interests" (IX.4.a); and "work towards developing internationally compatible procedures" to this effect. (IX.4.b)

The OECD (2016[10]) *Health Data Governance Recommendation* calls on governments to "facilitate the compatibility or interoperability of health data governance frameworks." (IV.2)

The OECD (2013[13]) *Privacy Guidelines* in Part Four recommend that

> *[a] data controller remains accountable for personal data under its control without regard to the location of the data. A Member country should refrain from restricting transborder flows of personal data between itself and another country where (a) the other country substantially observes these Guidelines or (b) sufficient safeguards exist, including effective enforcement mechanisms and appropriate measures put in place by the data controller, to ensure a continuing level of protection consistent with these Guidelines. Any restrictions to transborder flows of personal data should be proportionate to the risks presented, taking into account the sensitivity of the data, and the purpose and context of the processing.*

Part Six on "International Co-Operation and Interoperability" also indicates that "Member countries should encourage and support the development of international arrangements that promote interoperability among privacy frameworks that give practical effect to these Guidelines."

Building on these approaches, data governance policy makers should continue to promote pathways to foster interoperability across regulatory frameworks and enforcement within domestic policies and regulations, and through mechanisms for multilateral engagement and regulatory co-operation. They should ensure the effective protection of rights across borders. Finally, they should create a favourable environment for the digital economy to thrive.

How OECD legal instruments foster multilateral engagement and regulatory co-operation

The OECD (2021[6]) *EASD Recommendation* recommends Adherents "promote continued dialogue and international co-operation on ways to foster data access and sharing across jurisdictions … as well as the interoperability and mutual recognition of data access and sharing arrangements…" (VII.c) to "further improve conditions for cross-border data access and sharing with trust." (VII)

The OECD (2021[9]) *Research Data Recommendation* recommends that Adherents "collaborate at the international level … in order to enable free exchange of ideas and enhance scientific discovery..." (IX). To this respect, Adherents should "work together in international fora … to develop common definitions, data and security standards, and certification processes relating to access to research data … from public funding and design frameworks to enhance access … across different jurisdictions and national borders." (IX.1)

The OECD (2016[10]) *Health Data Governance Recommendation* calls on governments to "support … transborder co-operation in the processing of personal health data for health system management, research, statistics and other health-related purposes that serve the public interest." To that effect, governments should "identify and remove barriers to effective cross-border co-operation in the processing of personal health data … in a manner consistent with protecting privacy and data security, in light of all the circumstances." (IV.1)

The OECD (2013[13]) *Privacy Guidelines* in Part Six on "International Co-Operation and Interoperability" indicate that "Member countries should encourage and support the development of international arrangements that promote interoperability among privacy frameworks that give practical effect to these Guidelines." They should also "take appropriate measures to facilitate cross-border privacy law enforcement co-operation, in particular by enhancing information sharing among privacy enforcement authorities."

The Principle on "International access and use" of the OECD (2008[20]) *PSI Recommendation* recommends "supporting international co-operation and co-ordination …".

The OECD (2007[50]) *Recommendation of the Council on Cross-Border Co-operation in the Enforcement of Laws Protecting Privacy* encourages engagement with "civil society and business on their respective roles in facilitating cross-border enforcement of Laws Protecting Privacy, and in particular in helping raise awareness among individuals on how to submit complaints and obtain remedies, with special attention to the cross-border context." (IV.C.22.c)

3.3. Incentivising investments in data and their effective re-use

The third fundamental policy tension and objective in data governance relates to incentivising investments in data and their effective re-use. While the marginal costs of transmitting, copying and processing data can be close to zero, substantial investments are often required to generate and collect data and to enable data sharing and re-use. Investments may also be needed for data cleaning and data curation,[27] often beyond the scope and timeframe of the activities for which the data were initially collected (OECD, 2020[16]).

In many cases, complementary investments are also needed in data-related skills and competencies, as well as in information communication technologies (ICTs). This includes algorithms and software needed along the data value cycle (from data generation and collection to processing and re-use). Indeed, evidence shows that firms are increasingly buying start-ups to secure access to data and other complementary assets that may be critical for the development of their data-driven business (OECD, 2019[3]).

Policy makers should consider that data holders' incentives to invest in the collection, sharing and use of data will depend on expected returns. In some cases, policy measures to foster data openness may have unintentionally disincentivised investments in data and their re-use by lowering data collectors' expectations of returns on their investments.

Studies discussed in OECD (2019[3]) show that data openness can increase the value of data to data holders (direct impact). Indeed, they can help create 10-20 times more value to data users (indirect impact) and 20-50 times more value for the wider economy (induced impact). However, in some cases – when implemented in the form of open data, for example – data openness may reduce the producer surplus of data holders significantly.

Given significant investment requirements and potential disincentives to make them, data collection, sharing and use may not occur at sufficient levels across society. To address this issue, policy makers can rely on measures commonly used to address constraints to investments in innovation. As highlighted in the revised OECD (2015[66]) Innovation Strategy, such policies need to address the following:

- Low economic and social returns, due to fundamental – systemic – barriers to change and innovation. These could include factors linked to barriers to competition, lack of co-operation within an innovation system, prevailing norms and habits, and technology lock-in. They could also be linked to capacity constraints, or "low social returns", which are often linked, in turn, to lack of skills, infrastructure or inadequate institutions.
- Low appropriability of returns, due to market and government failures to prevent firms or other innovation actors from capturing the full value of their investments in innovation, thus leading to

underinvestment. Examples include the externalities associated with investment in R&D, where a firm can never capture all the returns to its investments due to spillover effects associated with knowledge production. Another example concerns the negative externalities related to environmental damages. Such damages are often not priced by the market, which adds to difficulties faced by investors to fully appropriate the returns from innovation in more environmentally friendly products.

While considering established innovation policy measures, policy makers also need to acknowledge the specific characteristics of data and how they intersect with better-known reasons of low investment. For example, the low appropriability of returns from investments in data is rooted in their economic properties. In fact, given the non-rivalrous nature of data, the original data holder may not be able to privatise all the benefits of their re-use due to their imperfect excludability (OECD, 2015[33]).[28]

Data openness, then, may benefit others more than the original data holder. Granting private property rights, and IPRs in the case of intangibles goods such as data, is often suggested as a solution to this incentive problems related to free riding.[29] However, as highlighted in section 3.2, IPR frameworks only apply to data under certain conditions. This limitation, combined with difficulties in quantifying the overall spillover benefits of data use and the detriments of their misuse, makes addressing the disincentive to invest in data and their effective re-use particularly challenging.

This section presents policy approaches to incentivise investments in data and their effective re-use by addressing the risks of low economic and social returns and of low appropriability of returns. These approaches complement those discussed in sections 3.1 and 3.2: balancing data openness and control, while maximising trust and managing overlapping and potentially conflicting interests and regulations related to data governance, are also necessary to incentivise investments in data and their re-use.

Policy approaches in this discussion thus include promoting a culture of responsible data sharing and use, and investments in related skills (subsection 3.3.1); encouraging investments in and adoption of financially viable ICT infrastructures for data (subsection 3.3.2); fostering competition in data-driven markets and addressing barriers to entry for new firms (subsection 3.3.3); and promoting standardised approaches for evaluating the social and economic value of data (subsection 3.3.4).

3.3.1. Promote appropriate knowledge and skills for responsible data sharing and use

Investment requirements for effective data access, sharing and use are not limited to data. In many cases, complementary investments are needed. Besides investments in secure ICT infrastructures, which are discussed in subsection 3.3.2, investments in complementary skills, competences and organisational capital are essential.

“[A] key barrier to data use in firms is a lack of knowledge on the potential offered by data for enhancing business performance, and a lack of training in the digital technologies required to process and analyse these data.” (OECD, 2022[67]) This also results in insufficient levels of investment in organisational capital, including for changes in business models, organisational structures and business processes that would be needed to make use of data. Consequently, “policies that encourage the adoption of these enablers, including efforts to increase firm-level skills utilisation, could be an asset in helping firms of all sizes thrive in the data-driven future” (OECD, 2022[68]).

Today, many governments recognise that availability of data-related skills and competences can be a bottleneck to effective use of data both in the private and public sector. Some governments have established dedicated initiatives to support development of data-related skills. For example, the United Kingdom’s Digital Skills Partnership brings together public, private and charity sector organisations “to boost skills for a world-leading, inclusive digital economy”. Similarly, the United Kingdom’s Data Ethics Framework aims to ensure that public servants from across disciplines in the country understand insights from data and emerging technologies and use data-informed insight responsibly (Department for Digital, Culture, Media & Sport, 2018[69]).[30]

The above-mentioned initiatives are complemented by the Centre for Data Ethics and Innovation. As part of the United Kingdom's Industrial Strategy and the AI Sector Deal, the Centre aims to drive the uptake and adoption of AI (Department for Digital, Culture, Media & Sport, 2018[70]).[31] It helps identify measures needed to strengthen and improve the way data and AI are used and regulated. This includes articulating best practice and advising on how to address potential gaps in regulation. The Centre helps ensure that those who govern and regulate the use of data across sectors do so effectively. By ensuring data and AI are used ethically, the Centre also aims to promote trust in AI and data analytics technologies.

How OECD legal instruments promote investments in data-related skills and a culture of responsible data use

The OECD (2022[71]) *Recommendation of the Council on SME and Entrepreneurship Policy* calls on Adherents to support "the adoption of digital technologies, services and data by all SMEs and entrepreneurs in line with their needs, digital maturity and aspirations by enhancing access to digital infrastructure; strengthening digital skills, data literacy and management of digital security risk; and ensuring open and well-functioning markets for digital goods and services."

The OECD (2021[6]) *EASD Recommendation* recognises the "need to further the promotion of a consistent culture of responsible data access and sharing and the legal and technical skills and capabilities necessary for responsible data access and sharing across society, including in the private and public sectors." It recommends that Adherents "adopt measures to enhance the capacity of all stakeholders to effectively use data responsibly along the data value cycle" including by promoting the development of data-related skills and competencies (IX.b). It also recommends to "promote appropriate incentive mechanisms that enable the fair distribution of the benefits of data access and sharing arrangements and ensure that stakeholders are enabled, encouraged, recognised, and rewarded for engaging in data access and sharing arrangements." (VI.d)

The OECD (2021[9]) *Research Data Recommendation* recommends that Adherents should "support the development of the human capital necessary to realise the full potential benefits of enhancing access to research data and other research-relevant digital objects from public funding" (VIII). To this effect, Adherents should

> *[i]dentify gaps and formulate strategies to develop and maintain the diverse skills necessary for data-driven research and innovation* (VIII.1); *develop appropriate learning and training programs and resources* (VIII.2); *[and] attract and retain data scientists and research software engineers across the breadth of scientific disciplines.* (VIII.3)

The section on "Incentives and Rewards" recommends that Adherents

> *[f]oster and support the development and implementation of effective models of reward and recognition that provide incentives and remove disincentives for researchers and research support staff to provide access to research data and other research-relevant digital objects from public funding.* (VI)
>
> *[f]oster, and require where appropriate, the adoption of measures to recognise and reward the provision of access to, and maintenance of, research data and other research-relevant digital objects from public funding as a recognised research output.* (VI.1)
>
> *[r]ecognize that researchers and institutions may require a reasonable limited period of exclusive use of the research data and other research-relevant digital objects they produce, for example to provide time for data analysis and preparation of final results and/or intellectual property claims*. (VI.2)

> *[p]romote, and require as appropriate, the use of unique digital identifiers for individual researchers, and research-relevant digital objects to facilitate and improve citation and provision of due credit to authors and contributors.* (VI.3)

The OECD (2016[10]) *Health Data Governance Recommendation* briefly addresses the issue of skills. It recommends that health data governance framework should provide for "establishment of appropriate training and skills development in privacy and security measures for those processing personal health data, that are in line with prevailing standards and data processing techniques." (III.10)

The OECD (2013[13]) *Privacy Guidelines* in Part Five on national implementation refer to skills development where they recommend that "in implementing these Guidelines, Member countries should: … consider the adoption of complementary measures, including education and awareness raising, skills development, and the promotion of technical measures which help to protect privacy."

3.3.2. Fostering investments in and adoption of financially viable ICT infrastructures for data

Investments in ICT infrastructures for data

In many cases, the effective use and re-use of data require complementary investments in metadata, data models and algorithms for data storage and processing, and in secure ICT infrastructures for (shared) data storage, processing and access. For example, the use and correct interpretation of data depend on the possibility to identify their source, the methodology of collection, as well as the various steps of curation. For this reason, investments in metadata and documentation are as critical as those in data collection itself. Complex datasets can also often be interpreted and used correctly only if suitable data processing algorithms and software are available to use.

Investments in and access to data storage processing and analytic infrastructures are thus a key condition for the effective re-use of data across society. This is especially true for SMEs and individuals such as scientists who may find it disproportionately costly to operate, maintain and scale their own data infrastructures. Cloud computing has been a major catalyst for the use and re-use of data and big data in particular. However, its diffusion remains much below expectations in particular among SMEs (OECD, 2021[72]; 2019[73]).

In this context, initiatives such as those presented above that aim to develop data-related skills are often complemented by those that establish ICT infrastructures for data, including data analytic support centres. For example, Australia's initiative to provide a big data infrastructure and technology foundation covers both support of data-related know-how and of the necessary infrastructure. This includes support to develop and provide statistical and analytical methods and tools with the help of data scientists for different target groups (supporting outcome monitoring and developing indicators based on heterogeneous data sources).

In the public sector, digital investments should be coherent and align with relevant procurement, technology, digital, data and services standards. For this reason, governments across OECD countries are investing to develop shared data infrastructures and standards. This aims to improve the efficiency of digital investments, while promoting data access and sharing within the public sector by reducing proliferation of fragmented digital solutions. Examples in this area include Italy´s National Digital Data Platform and the United Kingdom's Digital Marketplace (OECD, 2019[31]).

Business models for sustaining the function of ICT infrastructures for data openness

Lack of sustainable funding for ICT infrastructures for data openness, especially in open data infrastructures (such as open data portals and data repositories) remains a concern. This situation is exacerbated by the lack of viable business models for these infrastructures, especially since access to (open) data is often

assumed to be free rather than at the cost to the data holder (OECD, 2017[74]). In the area of science and research, many institutions (including funding agencies) are struggling to keep up with demand for help. There is a great demand to fund the infrastructure for data stewardship and data sharing, as well as the necessary training to support these activities.

Against this background, market-based approaches have been recognised as essential for encouraging the use, access and sharing of data. Already in 2017, the G7 ICT and Industry Ministers in Turin recognised that "market-based approaches to access and sharing of data are important to foster innovation in production and services, entrepreneurship and development of SMEs" (G7, 2017[75]). In this context, a growing range of data intermediaries that provide added-value services such as a payment and data exchange infrastructure can facilitate data sharing, including their commercialisation. They thus play a role in sustaining the function of ICT infrastructures for data openness.

However, policy makers need to acknowledge and, where possible, address challenges to enhance functioning of market-based approaches. In particular, the risk of data erosion has increased due to the combination of several factors. First, there is a lack of dedicated funding for data sharing infrastructures and limited pathways for their sustainment, even in critical areas such as science and health care research. This gap combined with misalignment of incentives to invest in, curate and share data has increased the risk of data erosion over time.

To address this issue, policy makers need to promote adoption of viable business models that ensure financing matches the desired duration of data sharing (OECD, 2019[3]). To this end, different revenue models can be used in combination. Promising revenue models include freemium,[32] advertisement, subscription, usage fees, selling of value-added goods and/or services, licensing and commission fees.

For research data repositories more specifically, policy makers, research funders and other stakeholders need to consider the ways in which data repositories are and can be funded. They should also reflect on the advantages and disadvantages of those business models in different circumstances (OECD, 2017[74]):

- **Structural funding** typically involves a trade-off between funding for data repositories and funding for other research infrastructure or for research itself. That allocation will best be made by informed actors making choices, such as through a funding allocation process involving widespread research stakeholder participation, expert consultation and "road-mapping".
- **Host or institutional funding** may divorce informed actors from the funding decisions or require additional processes to ensure greater stakeholder understanding of the value of the repository services.
- **Data deposit fees** bring the trade-off closer to researchers, but their success in optimising allocation will depend on the extent to which the actors are informed and on their freedom of choice. The latter may be constrained by open data mandates (regulation).

How OECD legal instruments highlight the importance of investments in complementary ICT infrastructures

The OECD (2021[6]) *EASD Recommendation* recognises

> *[t]hat data management, … and managing the associated risks, can require substantial investments over time and may involve a wide range of complementary digital resources, including algorithms, software, hardware, and other foundational infrastructures from multiple parties.*
>
> *[It then recommends] that Adherents adopt measures to enhance the capacity of all stakeholders to effectively use data responsibly along the data value cycle, including by facilitating access to and the adoption of the sustainable, open, scalable, safe, and secure foundational infrastructures.* (IX.c)

The OECD (2021[9]) *Research Data Recommendation* recommends that

> *Adherents take necessary measures to support development and maintenance of sustainable infrastructures to support the findability, accessibility, interoperability, and reusability of research data and other research-relevant digital objects from public funding free of charge at the point of use.* (VII)
>
> *To this end, Adherents should develop] strategies, including road-maps, funding plans, and business models, to ensure sustainable infrastructures for research data and other research-relevant digital objects from public funding, including data and software repositories and services.* (VII.1) *[They should also] encourage private investment in research data infrastructures with investment in the skills needed to manage and use them …* (VII.2)

3.3.3. Foster competition in data-driven markets and address barriers to entry for new firms

The dynamics of markets are increasingly transformed by firms' investments in data and their adoption of data-driven business models. As a result, some have expressed concern that the control of data may become a source of market power. As such, it could disincentivise non-incumbents from investing in data and data use. Therefore, some have suggested a need for competition authorities to better address and correct these developments. For competition authorities,

> *[t]his can include considering multi-sidedness and business models that involve the provision of products at a price of zero, often in exchange for consumer data. Similarly, the role of data in reinforcing demand-side characteristics of a market, including in relation to search and switching costs, and choice and information overload, could further contribute to entrenching market power for a firm in a dominant position. (OECD, 2022[67])*

Recent OECD (2022[76]) work on "The Evolving Concept of Market Power in the Digital Economy" (Box 3.1) also identified key considerations for assessing whether data are an input that gives rise to market power.

Box 3.1. Considerations for assessing whether data contribute to market power

As data exhibit some characteristics that set them apart from some more traditional important inputs, understanding their contribution to market power may require considering the following attributes:

- **The precise scope of data being considered**: The data relevant in a given market may be just a subset of a larger database, or it may be contained in multiple different databases. Moreover, the structure of data may differ significantly across firms. Further, the types of data used by different market participants may vary. Thus, relevant data can be amorphous and difficult to define relative to more traditional inputs.
- **The importance of a specific dataset for competing in a market**: The objective importance of a given dataset for a firm's ability to compete, and thus its contribution to market power if held exclusively, is a fundamental question. It can be divided as follows:
  - Are data an indispensable input in the market? In other words, can the product be offered to consumers at a competitive level of quality without this type of data? This assessment can consider whether alternative types of data or approaches can be used to provide the product.
  - Are there substitutes for the specific dataset held by the firm being assessed? In particular, the identification of potential substitutes for competitors, or alternative approaches not requiring the data in question, can be challenging. This is especially true in rapidly developing markets where firms collecting data may not themselves know exactly how they will apply them in the future. One consideration is whether comparable datasets can be purchased from third parties, including data aggregators, and at what cost (Autorité de la concurrence and Bundeskartellamt, 2016[77]).

- **The date of collection and the nature of data flows**: In some instances, a fixed static dataset may be sufficient. In other cases, the continual flow of data is most valuable. Data may have an expiration date in terms of usefulness (Oxera, 2018, p. 8[78]).
- **The quality and accuracy of the data**: Datasets based on inputs from users, or inferences, may be subject to accuracy limitations. Further, there may be measurement errors or data corruptions. Thus, quality and accuracy are crucial to the comparison of different datasets and the determination of whether a given dataset contributes to market power (for example, ACCC (2019[79])).
- **Whether scale is needed for the data to be useful**: Some datasets are useful on an individual consumer level, whereas others only generate value once they attain a given scale. For example, the marginal value of an individual search engine enquiry is likely to be minimal, since the algorithm will require large volumes of data to improve predictions[33] (OECD, 2016[80]; Autorité de la concurrence and Bundeskartellamt, 2016[77]). In contrast, individual-level data may be valuable for products like social media platforms, where access to that data is an important part of the platform's value (OECD, 2021[2]). When scale is crucial, or when individual-level data are not portable, data may contribute to incumbent market power.
- **Whether other data, or specific resources, are needed for the data to be useful**: Some datasets may need to be combined with others to be useful. For instance, a given dataset may need to be matched to individual user profiles in order to be valuable. Further, not all firms in a market may have access to the requisite skills and resources to be able to handle, process and analyse a dataset.

Source: OECD (2022[76]), *The Evolving Concept of Market Power in the Digital Economy*, https://www.oecd.org/daf/competition/the-evolving-concept-of-market-power-in-the-digital-economy-2022.pdf.

Policy makers and competition authorities need to consider how firms' control and use of data shape competitive dynamics. At the same time, they should consider whether asymmetric approaches are warranted in some areas either through competition enforcement or regulation. Any such interventions should ensure that measures to foster competition are focused on large incumbents and do not create barriers to entry for new firms. "Small and medium-sized enterprises, in particular, may need targeted assistance in restructuring and adapting to the new environment" (OECD, 2022[67]).

How OECD legal instruments help foster competition

The (2021[6]) *OECD EASD Recommendation* recognises

> *[t]hat market-based approaches, including commercialisation of data and freedom of contract, are essential for incentivising data access and sharing and related investments, but that there may be costs, risks, and limitations to these approaches' ability to fully meet demand for data.*

Its Section 2 on "Stimulating Investment in Data and Incentivising Data Access and Sharing" recommends that Adherents

> *[f]oster competitive markets for data through sound competition policy and regulation that addresses possible exploitation of market dominance and other appropriate measures, including enforcement and redress mechanisms that increase stakeholders' agency and control over data and ensure an adequate level of consumer, intellectual property, and privacy and personal data protection* (VI.a) *[and]*
>
> *[p]romote appropriate incentive mechanisms that enable the fair distribution of the benefits of data access and sharing arrangements and ensure that stakeholders are enabled, encouraged, recognised, and rewarded for engaging in data access and sharing arrangements.* (VI.d)

The (2008[20]) *OECD PSI Recommendation* recognises, under its "Openness" principle, the importance of access but also the need for refusal and limitations to access:

> *Maximising the availability of public sector information for use and re-use based upon presumption of openness as the default rule to facilitate access and re-use. … Defining grounds of refusal or limitations, such as for protection of national security interests, personal privacy, preservation of private interests for example where protected by copyright, or the application of national access legislation and rules.*

3.3.4. Promote standardised approaches for evaluating the social and economic value of data

Data have become a strategic social and economic resource, including for research, innovation, production and trade. In this context, one of the most important and complex challenges faced by policy makers is how to measure the social and economic value of data (Mitchell, Ker and Lesher, 2021[81]). Measuring this value is critical to provide a solid evidence base for policy making. Moreover, it also helps market participants more fully internalise the benefits and risks of data in their decision making. A range of approaches has tried to find ways of valuing data, including in the context of the System of National Accounts (SNA). OECD (2022[82]) presents three different approaches for assessing the value of data, databases, and data flows:

- **The market-based approach** determines value based on the price of comparable products on the market. However, most data are not traded. As a consequence, only a small portion of their value can be measured based on market statistics. These mostly include firms' revenues, international trade and market valuation.[34]
- **The net present value approach** determines value by estimating future cash flows that can be derived from the data. This approach may theoretically provide an accurate measure of the value of data. However, the non-rivalry nature of data and their unlimited use potential make it challenging for national statistical organisations (NSOs) to estimate the stock based on potential future earnings.
- **The sum-of-cost approach** determines value by the cost of producing the information/know-how derived from data. This appears to be the most promising approach to estimate the value of own-account data, i.e. produced by a firm not for sale but for its own use.[35]

Several experimental estimates of investment in data assets are available at this point, both from NSOs and academia, all of which rely on the sum-of-cost approach. In the most comprehensive effort to date, Statistics Canada (2019[83]) used wage and employment information for preliminary estimates of investment in and stock of total data assets.[36] These efforts were followed by initiatives at Statistics Netherlands (de Bondt and Mushkudiani, 2021[84]), the Australian Bureau of Statistics (Smedes, Nguyen and Tenburren, 2022[85]), and the US Bureau of Economic Analysis (Calderón and Rassier, 2022[86]). Goodridge and Haskel (2015[87]), Goodridge, Haskel and Edquist (2021[88]), and more recently Corrado et al. (2022[89]) also created estimates for several economies. International guidance on the implementation of the sum-of-cost approach in the SNA is being developed.

However, data have value for consumers, firms, governments, research activities and society at large that goes beyond the scope of macroeconomic statistics. (OECD, 2022[82]) This social value, which may also include negative value when data use is detrimental, is not captured by any of the approaches listed above. Therefore "developing conceptual schemes to think about these channels of value creation and statistical frameworks to measure them should be part of any measurement agenda for data" (OECD, 2022[82]).

How OECD legal instruments address the measurement of data

The (2013[13]) *Privacy Guidelines* under Part Six on "International co-operation and interoperability" calls on Member countries to "encourage the development of internationally comparable metrics to inform the policy making process related to privacy and transborder flows of personal data."

References

ACCC (2019), *Digital Platforms Inquiry: Final Report*, Australian Competition and Consumer Commission, Canberra, https://www.accc.gov.au/publications/digital-platforms-inquiry-final-report. [79]

AG Data Transparent (2016), *Ag Data's Core Principles: The Privacy and Security Principles for Farm Data*, webpage, http://www.agdatatransparent.com/principles/ (accessed on 10 May 2020). [102]

Australian government (2017), "Information about the Data Integration Partnership for Australia", http://www.pmc.gov.au/sites/default/files/publications/DIPA-information.pdf. [28]

Autorité de la concurrence and Bundeskartellamt (2016), *Competition Law and Data*, 10 May, Autorité de la concurrence and Bundeskartellamt, Bonn, https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf?__blob=publicationFile&v=2. [77]

Banterle, F. (2018), "The interface between data protection and IP law: The case of trade secrets and the database sui generis right in marketing operations, and the ownership of raw data in big data analysis", in Bakhoum, M. et al. (eds.), *Personal Data in Competition, Consumer Protection and Intellectual Property Law. MPI Studies on Intellectual Property and Competition Law*, Springer, Berlin, http://dx.doi.org/10.1007/978-3-662-57646-5_16. [34]

Barker, A. (2021), "Consumer data and competition: A new balancing act for online markets?", *OECD Going Digital Toolkit Notes*, No. 5, OECD Publishing, Paris, http://dx.doi.org/10.1787/e22e3a47-en. [54]

Bird & Bird (2019), "Big data & issues & opportunities: Intellectual property rights", 4 March, Bird & Bird, http://www.twobirds.com/en/insights/2019/global/big-data-and-issues-and-opportunities-ip-rights. [99]

Business at OECD (2020), *Regulatory Sandboxes for Privacy: Analytical Report*, OECD, Paris. [39]

Calderón, J. and D. Rassier (2022), "Valuing stocks and flows of data for the U.S. business sector", prepared for the BEA Advisory Committee, 13 May. [86]

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, http://dx.doi.org/10.1787/ca9f974e-en. [18]

Casalini, F., J. López-González and T. Nemoto (2021), *Mapping commonalities in regulatory approaches to cross-border data transfers*, OECD Publishing, Paris, http://dx.doi.org/10.1787/ca9f974e-en (accessed on 3 March 2022). [65]

Chisholm, M. (2011), "What is Data Ownership?", webpage, http://www.b-eye-network.com/view/15697 (accessed on 11 February 2019). [93]

CMA/ICO (2021), "Competition and data protection in digital markets: A joint statement between the CMA and the ICO", Competition & Markets Authority and Information Commissioner's Office, London, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/987358/Joint_CMA_ICO_Public_statement_-_final_V2_180521.pdf. [58]

CNIL (2019), "La CNIL et la DGCCRF font évoluer leur protocole de coopération pour renforcer la protection des consommateurs et de leurs données personnelles", 31 January, La Commission Nationale de l'Informatique et Libertés, Paris, https://www.cnil.fr/fr/la-cnil-et-la-dgccrf-font-evoluer-leur-protocole-de-cooperation-pour-renforcer-la-protection-des. [53]

Competition Bureau Canada (2022), "Unlocking the Power of Health Data: Digital Health Care Market Study – Part One", webpage, http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/04669.html (accessed on 3 November 2022). [32]

COPA-COGECA et al. (2018), *EU Code of Conduct on Agricultural Data Sharing by Contractual Agreement*, COPA-COGECA et al., https://copa-cogeca.eu/img/user/files/EU%20CODE/EU_Code_2018_web_version.pdf. [94]

Corrado, C. et al. (2022), "Data, digitization, and productivity", in *Technology, Productivty and Economic Growth*, University of Chicago Press, https://www.nber.org/conferences/criw-conference-technology-productivity-and-economic-growth-spring-2022. [89]

Council of Europe (2010), "Recommendation CM/Rec(2010)13 of the Committee of Ministers to member states on the protection of individuals with regard to automatic processing of personal data in the context of profiling", webpage, https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805cdd00 (accessed on 15 March 2022). [95]

de Bondt, H. and N. Mushkudiani (2021), "Estimating the value of data in the Netherlands", prepared for the IARIW-ESCoE Conference, 12 November, https://iariw.org/wp-content/uploads/2021/10/bondt_paper.pdf. [84]

Department for Digital, Culture, Media & Sport (2018), *Centre for Data Ethics and Innovation Consultation*, Department for Digital, Culture, Media & Sport, United Kingdom, http://www.gov.uk/government/consultations/consultation-on-the-centre-for-data-ethics-and-innovation/centre-for-data-ethics-and-innovation-consultation (accessed on 1 October 2022). [70]

Department for Digital, Culture, Media & Sport (2018), *Guidance Data Ethics Framework*, webpage, http://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework (accessed on 1 October 2022). [69]

Determann, L. (2018), "No one owns data", *UC Hastings Research Paper*, No. 265, SSRN, http://dx.doi.org/10.2139/ssrn.3123957. [37]

Drexl, J. (2018), "Legal challenges of the changing role of personal and non-personal data in the data economy", in Franceschi, A. and R. Schulze (eds.), *Digital Revolution: Data Protection, Smart Products, Blockchain Technology and Bitcoins Challenges for Law in Practice, Max Planck Institute for Innovation & Competition Research Paper No. 18-23*, Beck, Munich, https://ssrn.com/abstract=3274519. [35]

EDPS (2016), "EDPS opinion on coherent enforcement of fundamental rights in the age of big data", Opinion 8/26, European Data Protection Supervisor, Brussels, https://edps.europa.eu/sites/edp/files/publication/16-09-23_bigdata_opinion_en.pdf. [43]

EDPS (2014), "Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy", 26 March, Preliminary Opinion, European Data Protection Supervisor, Brussels, https://edps.europa.eu/data-protection/our-work/publications/opinions/privacy-and-competitiveness-age-big-data_en. [44]

European Commission (2018), "Guidance on sharing private sector data in the European", *Commission Staff Working Document*, No. SWD(2018) 125 final, European Commission. [59]

European Union (2016), *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC*, European Union, Brussels, http://data.europa.eu/eli/reg/2016/679/oj. [40]

Frischmann, B. (2012), *Infrastructure: The Social Value of Shared Resources*, Oxford University Press, Oxford. [91]

Frischmann, B., M. Madison and K. Strandburg (eds.) (2014), *Governing Knowledge Commons*, Oxford University Press. [41]

G7 (2017), *G7 ICT and Industry Ministers' Declaration: Making the Next Production Revolution Inclusive, Open and Secure*, http://www.g8.utoronto.ca/ict/2017-ict-declaration.html. [75]

Gierten, D. and M. Lesher (2022), "Assessing national digital strategies and their governance", *OECD Digital Economy Papers*, No. 324, OECD Publishing, Paris, http://dx.doi.org/10.1787/baffceca-en. [42]

Goodridge, P. and J. Haskel (2015), "How much is UK business investing in big data?", *Discussion Paper*, No. 2015/05, Imperial College London, Business School, https://spiral.imperial.ac.uk/bitstream/10044/1/25159/2/Goodridge%202015-05.pdf. [87]

Goodridge, P., J. Haskel and H. Edquist (2021), "We see data everywhere except in the productivity statistics", *Review of Income and Wealth*, http://dx.doi.org/10.1111/roiw.12542. [88]

Hardinges, J. (10 July 2018), "A data trust provides independent, fiduciary stewardship of data", ODI blog, https://theodi.org/article/what-is-a-data-trust/. [26]

Henze, M. et al. (2013), "Maintaining user control while storing and processing sensor data in the cloud", *International Journal of Grid and High Performance Computing*, Vol. 5/4, pp. 97-112, http://dx.doi.org/10.4018/ijghpc.2013100107. [96]

Hess, C. and E. Ostrom (eds.) (2007), *Understanding Knowledge as a Commons: From Theory to Practice*, MIT Press, Cambridge, MA. [100]

ICO/CMA (2021), "Memorandum of Understanding between the Information Commissioner and the Competition and Markets Authority", Information Commissioner's Office and the Competition & Markets Authority, London, https://ico.org.uk/media/about-the-ico/mou/2619798/ico-cma-mou-20210430.pdf. [56]

ICO/FCA (2019), "Memorandum of Understanding between the Information Commissioner and the Financial Conduct Authority", Information Commissioner's Office and the Financial Conduct Authority, London, https://ico.org.uk/media/about-the-ico/documents/2614342/financial-conduct-authority-ico-mou.pdf. [57]

ISO (2017), "ISO/IEC 19941:2017(en) Information Technology – Cloud Computing – Interoperability and Portability", webpage, https://www.iso.org/obp/ui/#iso:std:iso-iec:19941:ed-1:v1:en (accessed on 9 March 2021). [30]

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, http://dx.doi.org/10.1787/53ecf2ab-en. [61]

Madison, M. (2014), "Commons at the intersection of peer production, citizen science, and big data: Galaxy zoo", in *Governing Knowledge Commons*, Oxford University Press. [97]

METI (2019), "Contract guidance on utilization of AI and data 1.1 formulated", 9 December, News Release, Ministry of Economy Trade and Industry, Tokyo, http://www.meti.go.jp/english/press/2019/1209_005.html (accessed on 4 November 2022). [60]

Microsoft (2019), "Removing barriers to data innovation: Empowering people and organizations to share and use data more effectively", (fact sheet), Microsoft, Seattle, WA, https://3er1viui9wo30pkxh1v2nh4w-wpengine.netdna-ssl.com/wp-content/uploads/prod/sites/560/2019/12/Backgrounder-FAQ-Sheet_FINAL.pdf. [62]

Mitchell, J., D. Ker and M. Lesher (2021), "Measuring the economic value of data", *OECD Going Digital Toolkit Notes*, No. 20, OECD Publishing, Paris, http://dx.doi.org/10.1787/f46b3691-en. [81]

OAIC (2020), "ACCC/OAIC Compliance and Enforcement Policy for the Consumer Data Right", webpage, https://www.oaic.gov.au/consumer-data-right/compliance-and-enforcement-policy (accessed on 4 November 2022). [51]

OAIC (2019), *Digital Platforms Inquiry Final Report — Submission to the Australian Government*, Office of the Australian Information Commissioner, Sydney, https://www.oaic.gov.au/engage-with-us/submissions/digital-platforms-inquiry-final-report-submission-to-the-australian-government. [52]

OECD (2022), *Cross-border Data Flows: Taking Stock of Key Policies and Initiatives*, OECD Publishing, Paris, http://dx.doi.org/10.1787/5031dd97-en. [103]

OECD (2022), "Dark commercial patterns", *OECD Digital Economy Papers*, No. 336, OECD Publishing, Paris, http://dx.doi.org/10.1787/44f5e846-en. [8]

OECD (2022), "Data shaping firms and markets", *OECD Digital Economy Papers*, No. 344, OECD Publishing, Paris, http://dx.doi.org/10.1787/7b1a2d70-en. [67]

OECD (2022), "Enhancing online disclosure effectiveness", *OECD Digital Economy Papers*, No. 335, OECD Publishing, Paris, http://dx.doi.org/10.1787/6d7ea79c-en. [7]

OECD (2022), "Fostering cross-border data flows with trust", *OECD Digital Economy Papers*, No. 343, OECD Publishing, Paris, http://dx.doi.org/10.1787/139b32ad-en. [1]

OECD (2022), *Health Data Governance for the Digital Age: Implementing the OECD Recommendation on Health Data Governance*, OECD Publishing, Paris, http://dx.doi.org/10.1787/68b60796-en. [22]

OECD (2022), "Measuring the value of data and data flows", *OECD Digital Economy Papers*, No. 345, OECD Publishing, Paris, http://dx.doi.org/10.1787/923230a6-en. [82]

OECD (2022), *Recommendation of the Council on SME and Entrepreneurship Policy*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0473. [71]

OECD (2022), "Responding to societal challenges with data: Access, sharing, stewardship and control", *OECD Digital Economy Papers*, No. 342, OECD Publishing, Paris, http://dx.doi.org/10.1787/2182ce9f-en. [19]

OECD (2022), "The evolving concept of market power in the digital economy", *Background Note by the Secretariat*, OECD, Paris, https://one.oecd.org/document/DAF/COMP(2022)5/en/pdf. [76]

OECD (2022), "Turning data into business", in *Financing Growth and Turning Data into Business: Helping SMEs Scale Up*, OECD Publishing, Paris, http://dx.doi.org/10.1787/f5fcdf71-en. [68]

OECD (2021), "Data portability, interoperability and digital platform competition", *OECD Competition Committee Discussion Paper*, OECD, Paris, http://www.oecd.org/daf/competition/data-portability-interoperability-and-digital-platform-competition-2021.pdf. [2]

OECD (2021), *Good Practice Principles for Data Ethics in the Public Sector – OECD*, OECD Publishing, Paris, https://www.oecd.org/gov/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.htm (accessed on 14 April 2021). [4]

OECD (2021), "Mapping data portability initiatives, opportunities and challenges", *OECD Digital Economy Papers*, No. 321, OECD Publishing, Paris, http://dx.doi.org/10.1787/a6edfab2-en. [17]

OECD (2021), *Recommendation of the Council concerning Access to Research Data from Public Funding*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347. [9]

OECD (2021), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463. [6]

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/official-document/C(2021)42/en. [63]

OECD (2021), *The Digital Transformation of SMEs*, OECD Studies on SMEs and Entrepreneurship, OECD Publishing, Paris, http://dx.doi.org/10.1787/bdb9256a-en. [72]

OECD (2020), "Consumer Data Rights and Competition", webpage, http://www.oecd.org/daf/competition/consumer-data-rights-and-competition.htm (accessed on 4 November 2021). [45]

OECD (2020), *Enhanced Access to Publicly Funded Data for Science, Technology and Innovation*, OECD Publishing, Paris, http://dx.doi.org/10.1787/947717bc-en. [16]

OECD (2020), "Going Digital integrated policy framework", *OECD Digital Economy Papers*, No. 292, OECD Publishing, Paris, http://dx.doi.org/10.1787/dc930adc-en. [46]

OECD (2019), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris, http://dx.doi.org/10.1787/276aaca8-en. [3]

OECD (2019), *Going Digital: Shaping Policies, Improving Lives*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264312012-en. [47]

OECD (2019), *Measuring the Digital Transformation: A Roadmap for the Future*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264311992-en. [73]

OECD (2019), *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, http://dx.doi.org/10.1787/059814a7-en. [31]

OECD (2019), *The Policy Environment for Blockchain Innovation and Adoption: 2019 OECD Global Blockchain Policy Forum Summary Report, OECD Blockchain Policy Series*, http://www.oecd.org/finance/2019-OECD-Global-Blockchain-Policy-Forum-Summary-Report.pdf. [38]

OECD (2018), *Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact*, OECD Digital Government Studies, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264305847-en. [15]

OECD (2018), *Toolkit for Protecting Digital Consumers: A Resource for G20 Policy Makers*, Directorate for Science, Technology and Innovation, OECD, Paris, https://www.oecd.org/going-digital/topics/digital-consumers/toolkit-for-protecting-digital-consumers.pdf. [48]

OECD (2017), "Business models for sustainable research data repositories", *OECD Science, Technology and Industry Policy Papers*, No. 47, OECD Publishing, Paris, http://dx.doi.org/10.1787/302b12bb-en. [74]

OECD (2016), "Big data: Bringing competition policy to the digital era", *Background Note by the Secretariat*, OECD, Paris, https://one.oecd.org/document/DAF/COMP(2016)14/en/pdf. [80]

OECD (2016), *Recommendation of the Council on Consumer Protection in E-commerce*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0422. [14]

OECD (2016), *Recommendation of the Council on Health Data Governance*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433. [10]

OECD (2016), "Research Ethics and New Forms of Data for Social and Economic Research", *OECD Science, Technology and Industry Policy Papers*, No. 34, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jln7vnpxs32-en. [5]

OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264229358-en. [33]

OECD (2015), *Enquiries into Intellectual Property's Economic Impact*, OECD, Paris, https://www.oecd.org/sti/ieconomy/KBC2-IP.Final.pdf. [101]

OECD (2015), *Health Data Governance: Privacy, Monitoring and Research*, OECD Health Policy Studies, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264244566-en. [21]

OECD (2015), *Recommendation of the Council on Digital Security Risk Management for Economic and Social Prosperity*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0415. [11]

OECD (2015), *The Innovation Imperative: Contributing to Productivity, Growth and Well-Being*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264239814-en. [66]

OECD (2014), *Recommendation of the Council on Digital Government Strategies*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0406. [12]

OECD (2013), *Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0188. [13]

OECD (2008), *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0362. [20]

OECD (2007), *Recommendation of the Council on Cross-border Co-operation in the Enforcement of Laws Protecting Privacy*, OECD, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0352. [50]

OECD (2002), "Inventory of Privacy-Enhancing Technologies (PETs)", OECD, Paris, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=dsti/iccp/reg%282001%291/final. [25]

OECD (forthcoming), "Consumer Vulnerability in the Digital Age", *OECD Digital Economy Papers*, OECD Publishing, Paris. [23]

OECD (forthcoming), "Emerging privacy enhancing technologies: Maturity, opportunities and challenges", *OECD Digital Economy Papers*, OECD Publishing, Paris. [24]

OECD (forthcoming), "Review of the 2007 OECD Recommendation on Cross-Border Co-operation in the Enforcement of Laws Protecting Privacy", OECD, Paris. [49]

Ostrom, E. (1990), *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge University Press, http://dx.doi.org/10.1017/CBO9780511807763. [92]

Oxera (2018), "Market power in digital platforms", report prepared for the European Commission, Oxera, 30 September, https://ec.europa.eu/competition/information/digitisation_2018/contributions/oxera/oxera_market_power_in_digital_markets.pdf. [78]

Productivity Commission (2017), "Data availability and use", *Productivity Commission Inquiry Report*, No. 82, Productivity Commission, Canberra, https://www.pc.gov.au/inquiries/completed/data-access/report/data-access.pdf. [29]

Purtova, N. (2017), "Do property rights in personal data make sense after the big data turn?: Individual control and transparency", *Journal of Law and Economic Regulation* 10(2), https://ssrn.com/abstract=3070228. [36]

Robinson, L., K. Kizawa and E. Ronchi (2021), "Interoperability of privacy and data protection frameworks", *OECD Going Digital Toolkit Notes*, No. 21, OECD Publishing, Paris, http://dx.doi.org/10.1787/64923d53-en. [64]

Ruhaak, A. (2021), "How data trusts can protect privacy", *MIT Technology Review* 24 February, http://www.technologyreview.com/2021/02/24/1017801/data-trust-cybersecurity-big-tech-privacy/. [27]

Scassa, T. (2018), "Data ownership", *CIGI Papers*, No. 187, Centre for International Governance Innovation, https://www.cigionline.org/sites/default/files/documents/Paper%20no.187_2.pdf. [98]

Smedes, M., T. Nguyen and B. Tenburren (2022), "Valuing data as an asset, implications for economic measurement", prepared for Economic Implications of the Digital Economy Conference, 9-10 March. [85]

Statistics Canada (2019), "The value of data in Canada: Experimental estimates", *Latest Developments in the Canadian Economic Accounts*, 19 July, Statistics Canada, Ottawa, https://www150.statcan.gc.ca/n1/en/pub/13-605-x/2019001/article/00009-eng.pdf?st=ifEOEPUK (accessed on 3 February 2022). [83]

Stauber, P. (2019), "Facebook's abuse investigation in Germany and some thoughts on cooperation between antitrust and data protection authorities", *Antitrust Chronicle*, Vol. 2/2, pp. 36-43, https://www.competitionpolicyinternational.com/wp-content/uploads/2019/02/AC_February_2.pdf. [55]

Sundareswaran, S., A. Squicciarini and D. Lin (2012), "Ensuring distributed accountability for data sharing in the cloud", *IEEE Transactions on Dependable and Secure Computing*, Vol. 9/4, pp. 556-568, http://dx.doi.org/10.1109/TDSC.2012.26. [90]

Endnotes

[1] "Data access control mechanisms" thereby play an important role. The OECD (2021[6]) *EASD Recommendation* defines these mechanisms as "technical and organisational measures that enable safe and secure access to data by approved users including data subjects, within and across organisational borders, protect the rights and interests of stakeholders, and comply with applicable legal and regulatory frameworks".

[2] Studies discussed in OECD (2019[3]) show that while data openness can increase the value of data to holders (direct impact), it can help create 10-20 times more value to data users (indirect impact), and 20-50 times more value for the wider economy (induced impact). In some cases, however, data openness, and in particular open data, may also reduce the producer surplus of data holders, which is the cause of the incentive problem discussed in subsection 3.3. Overall, these studies suggest that data openness can help generate social and economic benefits worth between 0.1% and 1.5% of gross domestic product (GDP) in the case of public sector data, and between 1% and 2.5% of GDP when also including private sector data.

[3] It is broadly considered that IPR frameworks are applicable to data, but the extent and conditions of such applicability remains debated (Determann, 2018[37]; Scassa, 2018[99]; Bird & Bird, 2019[100]). In particular: **Copyright** typically "protects and rewards literary, artistic and scientific works, whatever may be the mode or form of their expression, including those in the form of computer programs" (OECD, 2015[102]). The protection afforded to databases (as collections of data or other elements) is established – or confirmed – by both Art. 10(2) of the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) and the almost identical Art. 5 of the WIPO Copyright Treaty. The arrangement or selection thus provides a separate layer of protection without prejudice to any rights to the content of the database itself. **Sui generis database right**: in some jurisdictions, such as the European Union, Japan and Korea, databases are also protected by a so-called sui generis database right (SGDR). This provides an additional layer of protection for databases regardless of the intellectual creation (i.e. "selection or arrangement") that may or may not be present. In other words, protection under the SGDR is granted without the requirement for human creativity or originality unlike IPRs such as copyright. What is protected more specifically is the investment in generating the database, i.e. in the obtaining, verification or presentation of the data. **Trade secrets** encompass "confidential business and technical information and know-how that a firm makes reasonable efforts to keep secret and that has economic value as a result" (OECD, 2015[102]). Trade secrets may protect the information conveyed by data, but only under some conditions, the most important one being that the information must be kept "secret". Not all data can thus be protected as trade secret. However, even where data can be protected, the dissemination of the data will only be possible to authorised persons (subject to confidentiality agreements) to a limited extent. That said, "by offering a measure of protection for valuable information and relieving businesses of the need to invest in more costly security measures, some trade secret laws may encourage businesses to invest in the development of such information" (OECD, 2015[102]).

[4] The cost of excluding others from using data will typically be low enough for the original data holder if the data are kept within a controlled environment. This is an environment where technical and organisational measures only allow approved users (including data subjects), within and across organisational borders, to access and use the data. However, once the data are transferred outside of the controlled environment, unless specific data stewardship and processing provisions are in place, the original data holder loses control of them (Sundareswaran, Squicciarini and Lin, 2012[91]; Henze et al., 2013[97]). For example, the ubiquitous nature of digital technologies, coupled with technological advances in data analytics and AI, make it increasingly easy to generate inferences about individuals from data collected in commercial or social contexts, even if these individuals never directly shared this information with anyone. Consequently, data holders and individuals, respectively, lose the ability to control how their data are re-used or object to or (technically) oppose such uses. As a result, they must rely solely on law enforcement and redress.

[5] The Council of Europe (2010[96]) recommends that in some circumstances the transparency extend to include the logic underpinning the processing in the context of profiling.

[6] Examples include open data initiatives such as data.gov (United States), data.gov.uk (United Kingdom), data.gov.fr (France) or data.go.jp (Japan).

[7] The term "community" as used in this report does not imply that access to data is free or that it is unregulated. The term rather describes social groups (of individuals or organisations) that have something in common, such as norms, values or interests.

[8] While there is no universally accepted definition of dark patterns, according to the OECD's Committee on Consumer Policy current working definition, "[d]ark commercial patterns are business practices employing elements of digital choice architecture, in particular in online user interfaces, that subvert or impair consumer autonomy, decision-making or choice. They often deceive, coerce or manipulate consumers and are likely

to cause direct or indirect consumer detriment in various ways, though it may be difficult or impossible to measure such detriment in many instances" (OECD, 2022[8]).

[9] Another dimension sometimes considered is whether third-party data recipients need to be accredited to participate in data portability arrangements.

[10] "Data holders" refers to organisations or individuals who, according to applicable laws or regulations, are competent to decide on granting access to or sharing data under their control, regardless of whether or not such data are managed by that organisation or individual or by an agent on their behalf (OECD, 2021[6]).

[11] "Data producers" refers to organisations or individuals that create, co-create, generate or co-generate data, including as a by-product of their social and economic activities, and can therefore be considered a primary data source (OECD, 2021[6]).

[12] Anonymisation is the process of removing identifying elements from data that would allow them to be linked to an individual or organisation. Anonymised data, in theory, should not be linkable back to an individual even when combined with outside/additional data sets.

[13] Once linked with sufficient other information, the likelihood that an individual will possess certain characteristics can be predicted to build a profile. The inferences may not be accurate, but even where correct, there remains a risk that they could be used against an individual's best interests, wishes or expectations.

[14] The concept of "commons" has been used to describe natural resources that are shared. The consumption of these commons is governed by mechanisms (including informal norms and values) that reflect the collective interest (Hess and Ostrom, 2007[101]; Madison, 2014[98]). Commons are therefore defined as collective goods, in which stakeholders have common interests, and which are characterised by the governance mechanisms surrounding their production and consumption. The establishment of data commons can be quite complex as it involves a number of considerations. For example, it touches on the scope of the community, institutional design, the relevant regulatory framework, boundaries and exclusion of non-members, pricing and congestion management to assure the sustainability of the commons, and in some cases even on exceptions from the non-discrimination rule.[14] The understanding, establishment and support of commons therefore require systematic analysis of all relevant contextual factors (Ostrom, 1990[93]; Hess and Ostrom, 2007[101]).

[15] The OECD Global Science Forum recommended that "[e]thics review bodies should, where consent for research use of personal data is not deemed possible or would impact severely upon potential research findings, evaluate the potential risks and benefits of the proposed research. If the proposed project is deemed ethically and legally justified without obtaining consent, ethics review bodies should ensure that information is made publically [sic] available about the research and the reasons why consent is not deemed practicable and should impose conditions that minimise the risk of disclosure of identities" (OECD, 2016[5]).

[16] A first example of application of Gaia-X is Catena-X. Catena-X plans to organise itself as a registered association in Germany. "Catena-X sees itself as an extensible ecosystem in which automotive manufacturers and suppliers, dealer associations and equipment suppliers, including the providers of applications, platforms and infrastructure, can all participate equally. The purpose of the association is to create a uniform standard for information and data sharing throughout the entire automotive value chain" (OECD, 2022[104]).

[17] It was followed by the Australian Institute of Health and Welfare, the Australian Institute of Family Studies, the Department of Social Services, the Queensland Government Statistician's Office, the Centre for Victorian Data Linkage, and South Australia Northern Territory DataLink. See https://toolkit.data.gov.au/Data_Integration_-_Accredited_Integrating_Authorities (accessed 24 August 2022).

[18] "Data access control mechanisms" are defined as "technical and organisational measures that enable safe and secure access to data by approved users including data subjects, within and across organisational borders, protect the rights and interests of stakeholders, and comply with applicable legal and regulatory frameworks."

[19] For example, in McInerney v. McDonald (1992[39] cited in Scassa, 2018[40]), "one of the theories considered, and ultimately rejected, by the court was that a patient owned their personal medical information" (Scassa, 2018[99]). Instead, the court found that the "physician, institution or clinic compiling the medical records owns the physical records".

[20] See Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique 2016 (Law for a Digital Republic). It defines "data of general interest" (données d'intérêt général) as including: (i) private sector data from delegated public services such as utility or transportation services, (ii) private sector data that are essential for granting subsidies and (iii) private sector data needed for national statistics.

[21] The *OECD* (2021[6]) *EASD Recommendation* defines the data ecosystem as "the integration of and interaction between different relevant stakeholders including data holders, data producers, data intermediaries and data subjects, that are involved in, or affected by, related data access and sharing arrangements, according to their different roles, responsibilities and rights, technologies and business models".

[22] Observed data are created where activities are captured and recorded. In contrast to volunteered data where the data subject is actively and purposefully sharing its data, the role of the data subject in case of observed data is rather passive and it is the data controller that plays the active role. Examples of observed data include location data of cellular mobile phones and data on web usage behaviour.

[23] Volunteered (or surrendered or contributed or provided) data are data provided by individuals when they explicitly share information about themselves or others. Examples include creating a social network profile and entering credit card information for online purchases.

[24] Derived (or inferred or imputed) data are created based on data analytics, including data "created in a fairly 'mechanical' fashion using simple reasoning and basic mathematics to detect patterns". In this case, it is (only) the data controller or processor that plays the active role in the creation of data. The data subject typically has little awareness over what is inferred about her or him, especially since that personal information can be derived from several pieces of seemingly anonymous or non-personal data. Examples of derived data include credit scores calculated based on an individual's financial history.

[25] Acquired (purchased or licensed) data are obtained from third parties based on commercial (licensing) contracts (e.g. when data are acquired from data brokers) or other non-commercial means (e.g. when data are acquired via open government initiatives). As a result, contractual and other legal obligations may affect the re-use and sharing of these data.

[26] See the conflict between farmers and agriculture technology providers that led in the United States to AG Data Transparent (AG Data Transparent, 2016[103]) and to the "EU Code of Conduct on Agricultural Data Sharing by Contractual Agreement" (COPA-COGECA et al., 2018[95]) in the European Union.

[27] Data curation embodies data management activities necessary to assure long-term data quality across the data value cycle.

[28] The argument that follows is that if data are shared, free-riding users can "consume the resources without paying an adequate contribution to investors, who in turn are unable to recoup their investments" (Frischmann, 2012[92]). However, the assumption that positive externalities and free riding always diminish incentives to invest cannot be generalised and needs careful case-by-case scrutiny. In this regard, Frischmann

(2012[92]) notes: "There is a mistaken tendency to believe that any gain or loss in profits corresponds to an equal or proportional gain or loss in investment incentives, but this belief greatly oversimplifies the decision-making process and underlying economics and ignores the relevance of alternative opportunities for investment."

[29] The often raised question about who "owns the data" is therefore essentially motivated by the recognition that ownership rights provide a "powerful basis for control" (Scassa, 2018[99]) as "to have legal title and full property rights to something" (Chisholm, 2011[94]) implies "the right to exclusive use of an asset" and the "full right to dispose of a thing at will" (Determann, 2018[37]).

[30] This Framework sets out principles and practical advice for using data, including building and procuring advanced analytics software, for designing and implementing policies and services. The framework is aimed broadly at anyone working directly or indirectly with data in the public sector, including data practitioners (statisticians, analysts and data scientists), policy makers, operational staff and those helping produce data-informed insights. The framework includes a Data Ethics Workbook with questions to probe ethical, information assurance and methodological considerations when building or buying new technology.

[31] The Centre will identify the measures needed to strengthen and improve the way data and AI are used and regulated. This will include articulating best practice and advising on how we address potential gaps in regulation. The Centre's role will be to help ensure that those who govern and regulate the use of data across sectors do so effectively. By ensuring data and AI are used ethically, the Centre will promote trust in these technologies, which will in turn help drive the growth of responsible innovation and strengthen the United Kingdom's position as one of the most trusted places in the world for data-driven businesses to invest in.

[32] The term "freemium" is a portmanteau of "free" and "premium". This revenue model is one of the most frequently used: products are provided free of charge, but money is charged for additional, often proprietary features (i.e. premium). This model is often combined with the advertising-based revenue model, where the free product is offered with advertisement while the premium is advertisement-free.

[33] While scale may not always improve a search algorithm's results, it may be particularly important for dealing with rare inquiries that have a higher chance of appearing with large volumes.

[34] "In the United States, the revenues from direct sales of data were estimated at USD 33.3 billion in 2019. In the same year, exports of data services from the European Union (EU27) and the United States were equal to USD 18.6 and 6.7 billion, respectively. Venture capital investments in "big data" firms, which reflect the investors' evaluation of the future revenues of these firms, reached USD 35.6 billion in 2021" (OECD, 2022[82]).

[35] As noted in (OECD, 2022[82]), "[t]his is consistent with the valuation of other own-account intellectual property products, e.g., software, and research and development."

[36] Total data assets consist of data, database and data science.

Going Digital Guide to Data Governance Policy Making

The ubiquitous collection, use, and sharing of data that power today's economies challenge existing governance frameworks and policy approaches. Drawing on the extensive research and analysis conducted at the OECD on data governance, on countries' policies and practices, and the OECD legal instruments in this area, the *Going Digital Guide to Data Governance Policy Making* supports policy makers in navigating three fundamental policy tensions that characterise efforts to develop, revise, and implement policies for data governance across policy domains in the digital age: balancing data openness and control while maximising trust; managing overlapping and potentially conflicting interests and regulations related to data; incentivising investments in data and their effective re-use. The operative part of the guide consists of a checklist of questions to orient policy makers as they develop and revise effective policies for data governance, based on possible policy approaches and real-life examples.

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 978-92-64-84416-2
PDF ISBN 978-92-64-84995-2