

Data in the Digital Age

The growing interactions between data, algorithms and big data analytics, connected things and people are opening huge new opportunities. But they are also giving rise to issues around “data governance” at the national and international levels. These include questions around the management of data availability, accessibility, usability, integrity and security, as well as concerns about ownership, impacts on trade and competition, implications for personal privacy, and more.

Policy makers across government are grappling with these issues. The use of digital technologies and data underpins digital transformation across all sectors of economies and societies, meaning that any policy decisions on data can have wide effects.

Instilling trust in the use of data is a pre-condition for fully realising the gains of digital transformation. This in turn calls for policy makers to better understand and account for the heterogeneity of data, to take a strategic approach to its governance and ensure all policy objectives are considered, and to upgrade their own capabilities to harness data for better policy making.

Key policy directions

- **Recognise that there are different types of data, and measure the economic and social value of data and data flows.** Many different types of data are collected and used, and their purpose and value can differ widely. Without recognising these different kinds of data, governments will not be able to effectively address issues such as personal data protection or competition. Better measurement of data and data flows and their role in production and value chains will also help support policy making.
- **Develop national data strategies.** The central role of data requires a high-level and strategic policy perspective that can balance multiple policy objectives. National data strategies and internationally interoperable approaches can help unleash the economic and social potential of data while effectively protecting privacy, intellectual property and other policy goals.
- **Invest in data capabilities and complementary assets to support policy making.** All stakeholder groups, including governments, must build data analytic and data management capabilities. Investments in data-related innovation and research and development (R&D), as well as in data related standards, skills and infrastructures, can enable governments to develop better policy across the board.

Data is a critical resource...

Data is not new, but it is now growing at an unprecedented pace. Humans have been recording facts as symbols, numbers and letters for thousands of years. What is new is the volume of data produced every day as a result of the proliferation of devices, services, and sensors throughout the economy and society. The volume of data only increases with its collection and use, creating a deep well of possibility for scientific discovery and for improving existing or inventing new products and services.

Not all data is valuable, but its general properties mean it is an increasingly critical resource. Since digital data can be used, reused, copied, moved, and processed cheaply, without degradation, at very fast speeds, and globally, it can drive economies of scale and scope.

...and is characterised by a high degree of variety

Data comes in many types, and the best way to categorise it will likely depend on the application, policy issue or business model at stake. For example, in some instances, it might be important to distinguish between public sector and private sector data; in others, whether the data is personal or non-personal (see Box 1 for more examples of possible data categories).

We therefore must not refer to data as a uniform entity, as this may lead to misunderstandings, oversimplifications and less effective policy. For instance, presuming that all data that could be linked to a person is personal data may lead to applying stringent regulations to all data sharing and transfers, whereas the risks to privacy in some cases may be small or inexistent. Taking into account the heterogeneity of data can help policy makers to formulate approaches that allow exploiting the potential of data to stimulate innovation and productivity while adhering to policy objectives such as protecting privacy and intellectual property rights, and ensuring security.

Box 1. Disentangling different types of data

There is no one “right” way to classify data into different types. One approach that could be relevant to policy making distinguishes the parties involved in data flows and assigns a proximate assessment of the personal content of that data. The example below is for illustrative purposes, to show the variations that may exist across data types.

Type of data	Personal content
Business to Business (B2B)	0 1 2 3 4 5 6 7 8 9 10
GVC data	⇒
Engineering (M2M)	⇒⇒
IoT (M2M)	⇒⇒⇒
Financial / human resources	⇒⇒⇒⇒
Business to Consumer (B2C)	
Media	⇒⇒⇒⇒
Consumer	⇒⇒⇒⇒⇒
Services (e.g. health, financial)	⇒⇒⇒⇒⇒
Government to Citizen (G2C)	
Services (e.g. health, tax, identity, social-welfare protection)	⇒⇒⇒⇒⇒
IoT (e.g. metro, CCTV)	⇒⇒⇒⇒⇒
Citizen to Citizen (C2C)	
Social media	⇒⇒⇒⇒⇒
Communications (e.g. e-mail, messages, voice)	⇒⇒⇒⇒⇒

Notes: GVC = global value chain; M2M = machine-to-machine ; IoT = Internet of Things; CCTV = closed-circuit television. Scale of personal content from 1 to 10, where 1 is low personal content and 10 is high personal content.

Source: Based on OECD (forthcoming a), *Enhanced access to and sharing of data: Reconciling risks and benefits of data re-use across societies*.

Extracting information from data creates value

Data becomes valuable when it is used to improve social and economic processes, products, organisational methods, and markets. Such data-driven innovation underpins many new business models that transform markets and sectors, ranging from agriculture to transportation to finance, and drives productivity growth (OECD, 2015). It can equally underpin improvements in the design and enforcement of government policies (Box 2).

Certain characteristics of data may be more valuable for some users than for others. For instance, velocity or the speed at which data is available is crucial for an application providing traffic updates. Similarly, the veracity or accuracy of data is critical for systems providing real-time feedback to trigger safety and maintenance warnings.

The value of data emerges through analysis, which requires human input to be most effective. Techniques including software, artificial intelligence (AI) and visualisation tools can extract insights, convert data into information and ultimately generate knowledge to support decision-making, push scientific frontiers or transform production.

Box 2. Harnessing data for better policy making

The application of digital technologies holds the potential to reshape existing policies, enable innovative policy design and rigorous impact evaluation, and expand citizen engagement in local and national policy making.

- **Data can be used to increase the efficiency of enforcement and targeting of existing policies.** In agriculture, for example, remote sensing and digital land parcel identification systems allow countries to grant direct subsidies to farmers and to enforce other regulatory measures related to the sustainability of agriculture.
- **Digital technologies can broaden the suite of policy instruments available to governments** and can lower the cost of policy experimentation and evaluation. In cities, digital cameras that automatically register the license plate of vehicles entering a congestion zone have made it more feasible to design, implement and enforce congestion pricing schemes.
- **Digital transformation can reshape government-citizen interaction and expand stakeholder engagement.** Many countries are making more data freely available to enhance accountability in the public sector and allow for evaluating the effects of current policies. Making pollutant release and transfer registers publicly available online can facilitate civil society oversight on regulated entities, making compliance efforts more transparent and breaches more open to public scrutiny.

Source: OECD (2019a), “Using digital technologies to improve the design and enforcement of public policies”, <https://doi.org/10.1787/99b9ba70-en>.

Enhancing access to data is critical to unleash its potential

The benefits of data are realised when data is available and accessible. For example, opening access to scientific data, particularly when the research is publicly funded, can boost collaboration, dissemination, reproduction and application of the results of scientific endeavour (Box 3).

There is no single optimal level of data openness and access, and no single optimal way to price data. Data does not need to be in the public domain to serve the public good; in some instances, restricted or preferential access may be sufficient and more appropriate. Business-to-government (B2G) data sharing initiatives for the public good (e.g. to share automated vehicle crash data), including, for instance, access to micro-data for informing policy makers, can be valuable, and data can be made available at different prices (or for free).

“Open government data” initiatives have the potential to spur innovation and new business models. Recent efforts by London’s transport authority to release consistent, up-to-date information, has been estimated to yield savings up to GBP 130 million per year for customers, road users, London and the authority itself (Deloitte, 2017) and added GBP 12 million to GBP 15 million in gross value per year for firms (OECD, forthcoming a).

Box 3. Supporting science and innovation through improved data access and sharing

Research is increasingly data intensive, making access to data a key enabler for science and innovation. Big data analytics is spreading across all scientific domains, with expectations of new scientific breakthroughs, less duplication and better reproducibility of results, as well as improved trust and innovation.

Publicly funded data are one important category of data used for science and innovation. The findability, accessibility, interoperability and reusability (FAIR) data principles are now a mainstream reference for policy makers. Many countries also have policy initiatives to support open science and open access, frequently related to infrastructure (such as repositories) and strategies.

Policy concerns remain, however, especially regarding:

- **Data governance for trust:** Opening up data can help advance the scientific agenda, but need to be balanced against issues of cost, privacy, security and prevention of malevolent uses.
- **Technical standards and practices:** Technology development is outpacing standard-setting, creating regulatory gaps.
- **Defining responsibility and ownership:** Intellectual property right protection can be an important incentive for private sector investment in R&D but there is a risk of “privatising” and preventing access to data derived from publicly funded research.
- **Incentives:** Academic reward systems currently motivate publication and do not attach enough value to the sharing of data.
- **Business models and funding:** Data providers may bear the costs of providing open data while users accrue the benefits, implying a need for business models that can sustain long-term financial commitments to open research data.
- **Building human and institutional capabilities:** Skills are needed along the full data value chain, including data management skills of researchers, curation skills with data stewards, and data literacy among users.

Making headway on these issues is critical, especially given that the importance of AI in assisting scientific discovery is expected to grow. Access to well-managed data is a key enabler of this development.

Source: OECD (2018), “Enhanced access to publicly funded data for STI”, https://doi.org/10.1787/sti_in_outlook-2018-11-en.

Key questions in the data debate

As data is now a social and economic resource, policy makers are facing a number of cross-cutting policy issues. Many hinge around understanding and dealing with the value, ownership and flows of data, the protection of personal data (including across borders), and the potential for data concentration and new digital divides. Some progress has been made on policy directions, and national data strategies offer one path for greater coherence, but these issues remain high on the digital transformation research agenda.

How should we value data?

A key dimension of many data policy debates – ranging from measures of gross domestic product (GDP) to consumer protection to taxation – is around data’s value, but no ideal methodology has yet emerged to measure it. Attempts to identify the value of data remain imperfect, because data’s value depends on its usefulness in a particular context. In the case of social networks, users interact by contributing content to various platforms (e.g. Facebook, Twitter), which enables the companies to collect data. Because users volunteer the data, it is not priced (and thus escapes the tax net), yet the data clearly has a value to social network companies (evidenced in part through advertising revenues) (OECD, 2019b).

Who “owns” data?

The rights to control access, copy, use and delete data – the main rights associated with the concept of “data ownership” – are currently affected by different legal frameworks, e.g. copyright and related rights, or, where personal data is concerned, privacy protection law (OECD, 2019c; 2015). This intricate, multi-stakeholder and often cross-border legal backdrop motivates many stakeholders to rely on contract law as the primary legal means for apportioning proprietary rights related to data access and use. However, uncertainties and divergent expectations remain, in particularly for those that are in a weaker position to negotiate fair terms and conditions. Alternative ownership mechanisms that have emerged include data commons (governance institutions which facilitate sharing data) and data trusts (which aim to help users trust how their data will be used and reused). While not implying full ownership as such, data portability mechanisms allow users to access data they have given to an organisation and to transmit it to another organisation. More work is required to advance thinking on data ownership and its policy implications.

How can we effectively address privacy concerns in a data-driven economy?

Privacy is becoming a flashpoint in digital transformation. As data proliferates and analytical techniques advance the ability to link what once was non-personal data to an individual, concerns about potential privacy violations have risen. Personal data is increasingly collected without people's explicit awareness or being used in ways not anticipated at the time of collection. With the growth in use and value of data, personal data breaches have become more common (OECD, 2017).

Paradoxically, while perception surveys show that people are concerned that their privacy is at risk, many individuals do not use options to change their privacy settings. Even when informed about the intended uses of their data, consumers may feel they have no other choices than to use a service despite disagreeing with terms and conditions, or to fully opt out. Whatever the reasons are that can explain this paradox – and behavioural insights can help better understand them – privacy and personal data need to be protected and managed more effectively. Indeed, privacy risks implicate not only the individuals concerned, but the core values and principles which privacy and personal data protection seeks to promote, including individual autonomy, equality and free speech.

To better protect privacy while recognising the importance of collection, processing and use of massive amounts of all kinds of data for innovation, the implementation of personal data protection regulation would benefit from being considered in the wider economic and social context, and as part of the broader data governance agenda, which includes also access to, and transfer of, data. Fostering interoperability of privacy approaches is also a promising path forward especially in a context where countries apply different privacy frameworks. More dialogue is needed, to help bring this about.

How should we think about cross-border data flows?

In a globalised world, many interactions span national borders, making cross-border data flows a critical enabler of economic and social activity. Today's trade and production activities are heavily dependent on moving, storing and using digital information (data), increasingly across borders. Data enables the co-ordination of international production processes through global value chains, helps small firms reach global markets, can be an asset that can be traded, or a conduit for delivering services, and is a key component for automation in trade facilitation. By some estimates, cross-border data flows contribute around USD 2.8 trillion to global economic activity, or 3.5% of global GDP (MGI, 2016).

At the same time, cross-border data flows raise questions about how to achieve important public policy objectives, such as the protection of privacy, security and intellectual property rights, in the new digital landscape. Building on work since 1980, the OECD recommended in 2013: "Any restrictions to transborder data 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" (OECD, 2013). Currently, an array of measures is making some cross-border data flows conditional or banning them altogether, with economic consequences, not least for trade transactions (OECD 2019d).

Reaping the benefits of digital transformation requires multi-stakeholder dialogue on regulatory approaches that ensure the interoperability of differing regulatory regimes, particularly for transversal issues such as cross-border data flows. The challenge is to preserve the significant economic and social benefits flowing from data-enabled trade, research and other activities. This underscores the importance also of better understanding the heterogeneity of data flows.

Is data concentration a problem?

Like many resources, data is not equally distributed. Concentration of data is visible, for example, in countries with high numbers of data centres (often being countries with a large population and uniform policies). Concentration is also visible at the sector or firm level with some companies holding far more data than others do as well as greater ability to analyse the data. Data savvy firms may have a comparative advantage enabling and inducing them to broaden their scope and expand to additional sectors either as new entrants or through acquisitions of existing firms (OECD, 2019b).

This may potentially create new types of competition in markets, as well as new ways to achieve anticompetitive outcomes, such as the use of algorithms to collude. It may also have implications for the distribution of power, with shifts away: (i) from individuals to organisations (incl. consumer to business and citizen to governments); (ii) from traditional businesses to data-driven businesses; (iii) from governments to data-driven businesses; and (iv) from lagging economies to data-driven ones. If these shifts result in new divides, it raises issues for social cohesion and economic resilience (OECD, 2015). Ongoing research and debate at the international level is essential.

Towards national data strategies

The formulation of national data strategies can help realise the potential of data for economies and societies. Some countries are in the process of developing such strategies, and some data-related aspects are already addressed in open government data strategies as well as in national digital economy and/or security strategies (OECD, forthcoming b). Building on their experiences, governments could consider developing consolidated data strategies that set out a comprehensive and coherent approach to leverage the potential of data for value creation while addressing the many challenges arising with data, including across borders.

Further reading

Deloitte (2017), *Assessing the value of Tfl's open data and digital partnerships*, <http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf>.

MGI (McKinsey Global Institute) (2016), "Digital Globalization: The new era of global flows", McKinsey & Company, March, www.mckinsey.com/business-functions/mckinseydigital/our-insights/digital-globalization-the-new-era-of-global-flows.

OECD (2019a), "Using digital technologies to improve the design and enforcement of public policies", *OECD Digital Economy Papers*, No. 274, OECD Publishing, Paris, <https://doi.org/10.1787/99b9ba70-en>.

OECD (2019b), "Vectors of digital transformation", *OECD Digital Economy Papers*, No. 273, OECD Publishing, Paris, <https://doi.org/10.1787/5ade2bba-en>.

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

OECD (2019d), "Trade and Cross-border data flows", *OECD Trade Policy Papers*, No. 220, OECD Publishing, Paris, <http://dx.doi.org/10.1787/b2023a47-en>.

OECD (2019e), *Measuring the Digital Transformation: A Roadmap for the Future*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264311992-en>.

OECD (2018), "Enhanced access to publicly funded data for STI", in *OECD Science, Technology and Innovation Outlook 2018: Adapting to Technological and Societal Disruption*, OECD Publishing, Paris, https://doi.org/10.1787/sti_in_outlook-2018-11-en.

OECD (2017), *OECD Digital Economy Outlook 2017*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264276284-en>.

OECD (2016), *The OECD Ministerial Declaration on the Digital Economy: Innovation, Growth and Social Prosperity (the Cancun Declaration)*, OECD, Paris, <https://www.oecd.org/internet/Digital-Economy-Ministerial-Declaration-2016.pdf>.

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

OECD (2013), "Privacy Expert Group Report on the Review of the 1980 OECD Privacy Guidelines", *OECD Digital Economy Papers*, No. 229, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5k3xz5zmj2mx-en>.

OECD (forthcoming a), *Enhanced access to and sharing of data: Reconciling risks and benefits of data re-use across societies*, OECD Publishing, Paris.

OECD (forthcoming b), "Towards national privacy strategies", *OECD Digital Economy Papers*, OECD Publishing, Paris.

Please cite this note as:

OECD (2019), “Data in the digital age”, *OECD Going Digital Policy Note*, OECD, Paris, www.oecd.org/going-digital/data-in-the-digital-age.pdf.

This document, as well as any data and any map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

www.oecd.org/going-digital – goingdigital@oecd.org –  [@OECDInnovation](https://twitter.com/OECDInnovation) – [#GoingDigital](https://twitter.com/GoingDigital)