Bennett Institute for Public Policy Cambridge

ODI

Nuffield Foundation

The value of data

The Value of Data summary report / 2020

How is the value of data created, captured and distributed?

Introduction

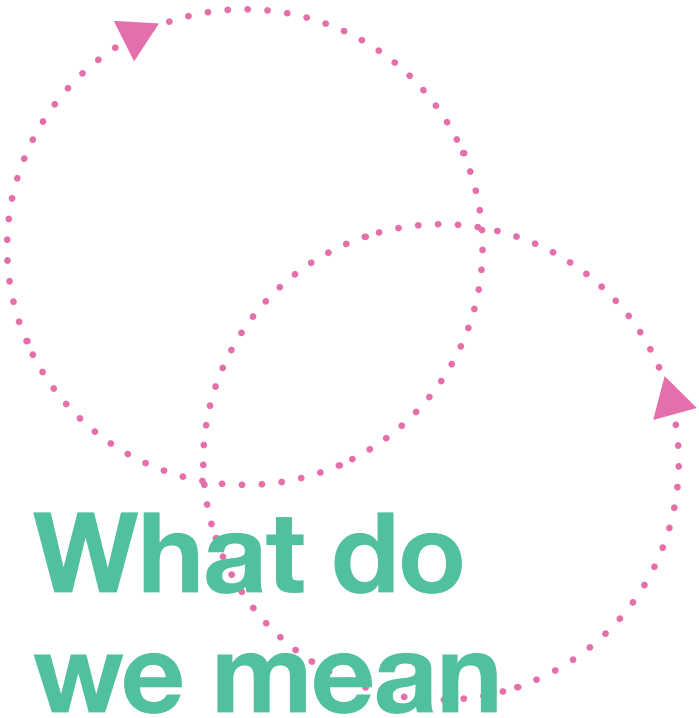
We are living in an age where data is becoming central to how our economies and societies function. Putting a precise figure to the value of data is extremely difficult, not least because there are lots of types of data and it is used in lots of different and unanticipated ways. But most analyses agree that the value of data is not being fully realised or evenly spread. Economies and societies that find ways to unlock this value and distribute it fairly will be at an advantage.

In this research, we looked at existing literature on the value of data, conducted interviews, and consulted with an expert advisory group. We looked at both the economic and informational characteristics of data and the data economy. We reviewed the different ways people are measuring the value of data. We examined how sharing data affects what value is unlocked and how that value is distributed.

We hope this research can help inform the answers to questions like:

- Which datasets should the government invest in maintaining?
- How should public sector data be made available?
- How should the financial benefits that arise from having access to data be distributed?
- What interventions would help unlock the value of data?

Our work shows that the economic characteristics of data and the data economy mean the market alone will not unlock data's full potential value. But it is possible to gain more from data with the right data policies, and an institutional framework that supports trustworthy access to data.



What do we mean by value?

People mean different things when they talk about the value of data. Some people think of it only in terms of monetary value. They might focus on how much organisations should be charged for data or what figure to include on a balance sheet.

In this report 'value' refers to the economic concept of social welfare: the wellbeing of all society. Value arises from data when businesses create jobs or become more productive; when governments deliver more effective public services; when our environment is clean and diverse; and when people live happier and healthier lives.

What characteristics does data have?

Economic characteristics

Data is unlike other goods and assets we are used to buying and selling. It has particular economic characteristics that affect how we should think about its value:

- **Data is non-rival.** Many people can use the same data at the same time without it being used up. It is not useful to think of data as being owned or exchanged, though it is often described as such.
- **Data varies in whether it is excludable.** In economic language, data is a *public good* when anyone can use it, or a *club good* when it is possible to exclude people from accessing it. It is hard to exclude people from data that is easy to collect. For example, *environmental data*, such as air quality or images scraped from websites, can be collected by anyone with the right equipment, whereas *administrative data* can only be collected by those delivering a service.
- **Data involves externalities.** Positive externalities happen when one dataset is enhanced through being combined with others. Negative externalities arise when harms arise through data being collected or used. Externalities make the market do too much with data (undermining privacy and personal autonomy) and too little (limiting potential insights and opportunities).
- **Data may have increasing or decreasing returns.** Sometimes, collecting more data can provide additional insights. At other times, more data adds little extra value but organisations may continue to accumulate and hoard it to prevent others entering the market.
- **Data has a large option value.** It is hard to predict how the value of data might change. New data, new technologies or algorithms, and new questions, may mean existing data has unpredictable future importance. Organisations may keep data for its potential value rather than its current value.
- **Data collection often has a high up-front cost and low marginal cost.** Collecting data can mean investment in hardware (such as sensors), digitisation, or quality improvement. It can then be cheap to collect, particularly when that is highly automated or a side-effect of other activities. High up-front costs can prevent firms collecting data or make them look for a financial return.
- **Data use requires complementary investments.** Organisations may need to invest in things like software, compute resources, and skilled people. These costs can be a barrier to getting value from data.

Data about people has some particular features.

First, data about any one person may have limited value, but combined with data about other people it can reveal trends or patterns. People create positive and negative impacts for others as they share data about themselves. This makes it hard to draw sensible conclusions about the value of data about any one person.

Second, organisations that collect and use personal data need to comply with data protection legislation. Costs arise from implementing technical and process compliance. Risks arise from the possibility of fines and reputational damage.

Because data is non-rival and has externalities, the market is unlikely to provide the best overall outcomes. Value comes from data being brought together, and that requires organisations to let others use the data they hold. But if they do, organisations will not get all the benefits from data they have collected, and perhaps not enough benefits to cover the cost of collecting and storing the data in the first place.

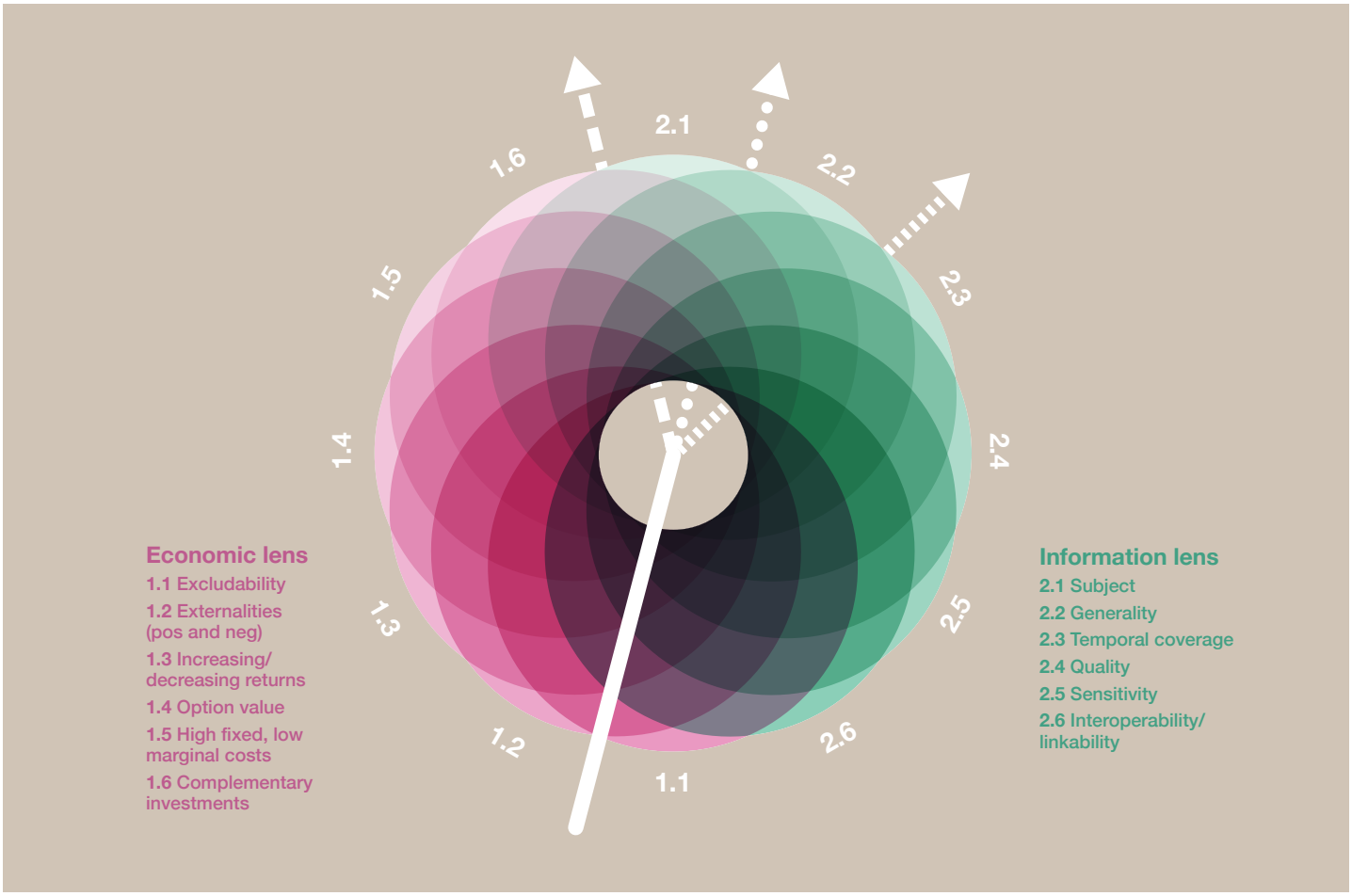
Informational characteristics

For a data scientist, the value of a dataset relates to its informational characteristics. These include characteristics like:

- **What is the data about?** The subject matter and geographic coverage of a dataset determines what kind of problems it can be used to tackle. For example, transport data in a city can help with transport problems in that city.
- **How general purpose is the data?** Some datasets, such as maps, might be useful across a range of analyses; others are only valuable in a few situations. More detailed or granular data can be filtered and aggregated for different uses.
- **What temporal coverage does the data have?** Datasets can be *forecasts*, *real-time*, *historic* or *backcasts*. These have different value for people *planning*, *operating* or performing *historic analyses*. For example, forecasts are useful during planning, while real-time data is valuable for those operating a service.
- **What is the quality of the data?** Quality incorporates several characteristics. For example, *completeness* measures the proportion of data in a dataset compared to what's required; *accuracy* assesses how well the data reflects a real-world situation; and *timeliness* expresses how up to date the data is. Higher quality data is more valuable because it reduces the uncertainty and risk in an analysis, but the required quality of data depends on what it is used for. Greater automation in the collection of data tends to mean higher quality data.
- **How sensitive is the data?** Data is sensitive when people could use it in damaging ways. It might be personal, commercially sensitive, about national security targets, endangered animals or protected environments. Keeping sensitive data can entail costs and risks for organisations, and sharing it must be done carefully.
- **How interoperable and linkable is the data?** Interoperability affects how easy it is to work with and aggregate datasets; linkability is the ease with which they can be joined up. Both make it easier to get value from data as combining datasets enables new insights. *Data standards* help to improve interoperability and *reference data* provides common reference points that improve linkability.

Quality and interoperability can sometimes be improved through investment to increase the potential value of a dataset.

The value of data should be viewed through both economic and information lenses



How can you estimate the value of data?

People have tried to estimate the value of data in a variety of ways. These fall into two main groups: market-based and non-market-based valuations.

Market-based valuations

Stock market valuations demonstrate the advantage gained by companies who invest in data and data capability. PwC found data-driven firms have higher stock-market valuations, and those with data analytics capabilities perform better than peers within the same industry.[1]

Income-based valuations aim to estimate current and future income that may be derived from data. Income-based valuations have several limitations:

- They cannot reflect wider benefits of data to an organisation or system, beyond the financial transactions involving that data. These benefits may be substantial. Indeed in most cases very little of the data created by an organisation will be sold by that organisation.
- Technological changes provide new opportunities for use and enable new competitor datasets to be created. Market prices at any moment in time are unlikely to factor in the full option value (potential future value) of the data.
- Most markets for data are not *thick markets*, which means they do not have enough buyers and sellers to converge on a market price that reflects the true economic value of data. This is amplified by the fact data markets are complicated, non-transparent, increasingly concentrated on a few firms, and in some cases unlawful.

Cost-based valuations reflect the cost of creating and maintaining data. The national accounts reflect the costs of preparing data in a useful format, which are mainly labour costs. There is active debate on whether and how national accounts should reflect the cost of purchasing or producing data. For organisations, cost-based valuation can look at their investment into data or the cost of replacing data.

Non-market-based valuations

The economic value of open data has been examined for a number of free and open datasets, such as Landsat data ($2bn/year)[2] or Transport for London (TfL) data (£130m/year from a £1m/year investment).[3] These estimates are often created using contingent valuation: asking how much organisations would be willing to pay for access to data. They typically look at how value arises from data for one or more of three groups:

- **Organisations that steward data** bear the costs of collecting, using and sharing data. They may benefit from that data internally, gain revenue by charging for access to that data, or reduce their own costs by sharing data. For example, by publishing open data, TfL saves £1m/year in customer support costs they would otherwise have to bear.
- **Intermediaries that reuse data to create products and services** bear the costs of gaining access to data, which may be financial and/or be investment in capability to use that data. They build businesses that sell the products and services they develop, generating jobs, innovation and economic growth. For example, TfL's open data contributes £14m/year to London's economy and has generated over 700 jobs. Companies House data supports £23m/year of revenue for intermediaries.[4]
- **People and organisations who use those products and services** benefit from data that informs decisions and action. For example, Landsat data contributes to monitoring and protecting the environment, enforcing regulations and increasing human safety. TfL data saves time worth £70–90m/year for those travelling around London, and increases the accessibility of public transport for those with accessibility needs.

The value of personal data can also be estimated through contingent valuation. The figures this generates depends on the question asked. For example, a survey of American consumers showed they would be willing to pay $5/month to use a privacy-preserving service, but would charge $80/month to allow access to personal data.[5] Consumers say they would need to be compensated by about $48 to give up Facebook for a month.[6]

Other values for personal data can be derived from the profits of companies that rely on it. In 2018, Facebook generated about $10/year income per active daily user.[7] On the dark web, login details can fetch between £1 and £280.[8] Fines issued by data protection authorities give figures that are skewed by legal limits on fines:

- The US Federal Trade Commission fined Equifax $575m – $4 per person – for its 2019 data breach; people affected could also claim up to $20k in compensation.[9]
- The UK Information Commissioner's Office fined Facebook £500k in 2018 (the maximum allowable fine at the time) for allowing access to data: 0.6p per person affected.[10]

Regardless of these figures, privacy is a human right and many argue it should not be negotiable.

Why is access to data important?

Access to data determines how much of the potential value of data can be unlocked and who can benefit from that value. Data is inherently *non-rival*, and does not get used up as more people use it. However uses can be excluded: access to data can be limited by technology, licences, governance, and price.

Restricting access to data has two kinds of impacts. First, it limits who can use data to run an analysis or develop a product or service. Efforts to share data more widely often focus on expanding this group, for example by enabling academic researchers and startups to access data. Second, restrictions limit who can use data to make decisions. Sharing data can help widen that set of people and organisations, for example to include excluded populations or small and medium-sized enterprises.

Technology can restrict the use of data in a variety of ways. This can be purposeful, such as when Digital Rights Management is used to enforce legal restrictions. It can also be unintended. For example, proprietary or non-standard data formats can increase barriers to the use of data.

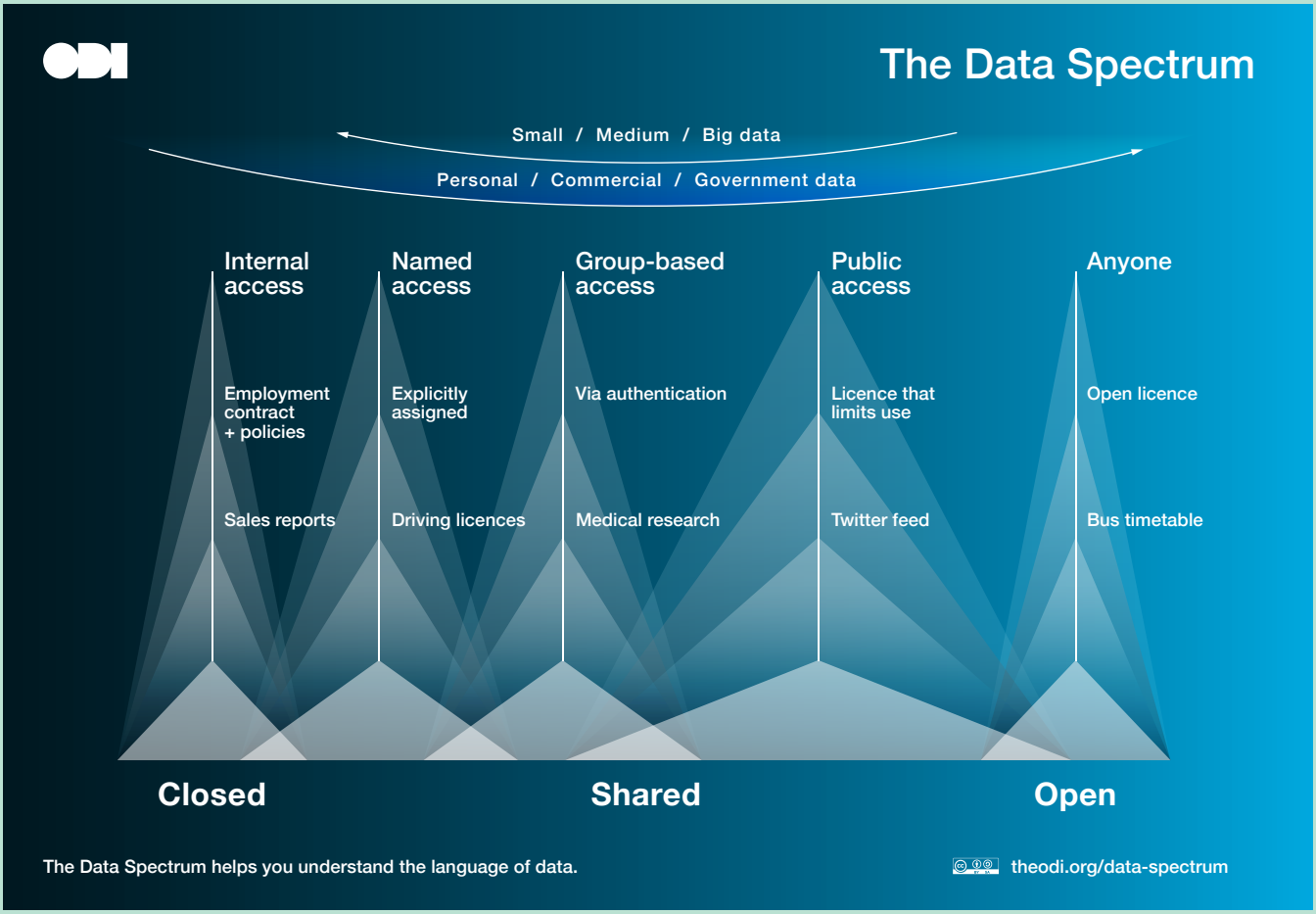
The Data Spectrum shows how **licensing** restricts access to data. Data can be closed and only benefit the organisation that holds it (a *private good*); shared in different ways with a select group of organisations (a *club good*); or open (a *public good*) for anyone to access, use and share. Beyond licensing terms themselves, the complexity of licensing can itself increase the transaction costs and risks of accessing data for reusers.

Legal and ethical governance constrains how widely data can or should be shared. Sharing sensitive data can cause harm to the people, organisations or environment the data describes. For instance, while data about people's habits is valuable to advertisers, data protection laws may restrict access to that data. That said, personal data can be found at both ends of the Data Spectrum.[11]

Cost can also restrict the use of data. But questions about licensing and regulation are distinct from those about charging. It is possible to restrict access to data – for example to accredited researchers – without charging for it. It is also possible to charge for reliable or custom access to open data.

The *public good* character of open data, and the network effects of the *positive externalities* that arise as data is combined with other data, means increasing access to data will increase its societal value. However, there are several trade-offs to consider:

1) The need to incentivise investment and innovation by retaining or granting exclusive access to data is similar to the well-known trade-off in intellectual property, where copyright protection incentivises investment in creativity and discovery. However when data is created as 'exhaust' – a by-product of the provision of a digital service for example – these incentives may be unnecessary.
2) The costs and risks borne by organisations that share data, in particular when it is personal or otherwise sensitive data, also need to be shared. Organisations may have little visibility or control over third-party uses but bear the burden of liability and reputational impacts if it is misused
3) The ability to assess the current and potential future value of data is limited which makes it hard to be certain about whether increasing the quality or accessibility of data will be worth it. Creating such an evidence base is costly in itself.



How can access to data be improved?

Data access approaches need to address several characteristics of the data economy:

- **Coordination problems:** those that collect data are unaware of who might be interested in using it, and those who might benefit are unaware of who is collecting it.
- **Information asymmetries:** some people or organisations know more than others about datasets, such as their quality or how they could be used.
- **Principal-agent misalignments:** the incentives of those that collect data, those that could use data, and those who are affected by it, are mismatched.
- **Incomplete contracts:** uncertainty about how other people will behave because of ambiguity in norms, regulations, licensing or terms and conditions, or because there are few ways of assessing the trustworthiness of other parties.

Opening data

Unlocking the economic and social value of public sector information is one motivation for making government data open, alongside driving democratic accountability.

Estimates for the value of open government data as a percentage of GDP have ranged from 0.08% to 7.19%.[12] A recent OECD report cites a range of 1% to 2.5% of GDP.[13]

Creating, maintaining and publishing data has a cost and governments often consider charging for public sector data. However, requiring payment can impede its use and the value derived from it. Research for the ODI found that a shift from cost recovery to open access would create 0.5% more GDP growth per year for the UK economy.[14]

There is ongoing research into which public sector datasets might be of most value as open data. The European Commission identified six data types that appear to have the most value: geospatial, earth observation and environmental, meteorological, statistics, company data, and transport data.[15]

Private sector organisations can also open the data they hold.[16] The UK's Digital Competition Expert Panel highlighted the significant potential for shared or open data to promote competition and innovation in the economy.[17] The UK's Bus Services Act 2017, which mandates bus operators to open data about bus timetables, demonstrates how the government can regulate to open data.[18]

Sharing data

There have long been arrangements and markets that enable data to be shared as a *club good*. Newer data sharing schemes aim to unlock value from data by minimising access restrictions that arise from technology, licensing or cost while retaining restrictions based on legal and ethical considerations.

Organisations that open or share data are concerned about what reusers might do with it, and consequent regulatory or reputational costs. They worry about unintentionally revealing commercially confidential information. They find it hard to predict how data might be used in the future, and what value they might be giving away.

Despite this, there are a number of initiatives where data is being shared, for example:

Organisations providing restricted access to sensitive data. The Office for National Statistics's Secure Research Service enables approved researchers can gain access to government administrative data.[19] Private sector examples include Goldcorp sharing its proprietary geological database to invite third parties to locate gold deposits.

People having better control over the way personal data is used. A number of initiatives mediate between individuals and organisations wishing to use personal data. Databox, for example, helps individuals get access to data held about them and Internet of Things devices, and grant access to third parties.[20] Solid similarly centres on individuals controlling access and storage of personal data in a decentralized model.[21]

Institutions balancing the different needs and incentives of those who collect data and those who want to use it. For example, the Data Communications Company manages energy-data sharing between competing energy suppliers, energy network operators and switching services.[22] Data trusts give fiduciary responsibility for data-sharing decisions to named trustees.[23] Intermediary data institutions like these may provide a route to distribute costs and benefits, as well as reducing transaction costs and increasing efficiency in data sharing.

Who gets value from data and how?

DATA ABOUT ENVIRONMENT
DATA ABOUT PEOPLE
DATA ABOUT OBJECTS
DATA ABOUT ORGANISATIONS

ECONOMIC VALUE
EVIRONMENTAL VALUE
SOCIAL VALUE

Data is collected by sensors and systems; excessive data collection can be harmful

STEWARD

Some organisations get paid public money to collect and maintain data for everyone

Data about the world is collected by those who steward data; doing this is costly

Organisations that steward data can get value from using it internally

Organisations that collect data choose how much they share and with whom

Organisations that collect data can share it with others

Insights from data inform decisions and action that have social, economic and environmental impacts

Intermediaries create products and services using data that automate actions or inform decisions

Some organisations charge for access to data they steward

INFORM
ADVISE
ACT

Organisations that create value from data also create economic growth and jobs

CREATE

People and organisations pay to buy software or subscribe to services

The flow of data
The flow of value
The flow of money

What kinds of institutions do we need?

It is helpful to think of data as a form of virtual infrastructure on which we now rely. The regulatory and institutional structures we design need to consider:

- **Data assets** such as datasets, identifiers and registers
- **Standards and technologies** used to curate and provide access to those data assets
- **Guidance and policies** that inform the use and management of data assets and the data infrastructure itself
- **Organisations** that govern the data infrastructure
- **Communities** that contribute to or maintain data infrastructure, and those who are affected by decisions that are made using it

There has been a lot of economic work on institutional frameworks that support getting the best value from non- rival goods, like data, through norms, regulations, and laws. This work also examines the mix of market, collective and government decisions about what goods get produced and how the value created from them is allocated.

> **“ Intermediary organisations like data trusts may provide a route to distribute both costs and benefits, as well as reducing transaction costs and increasing efficiency in data sharing”**

This work is highly relevant to the data economy. The same analytical tools can be used to design data access regulation and institutions. To complement this traditional government regulation, there is already innovation and experimentation with institutional frameworks. This includes voluntary sharing arrangements, establishing data trusts and cooperatives, and the mandatory data access schemes gradually being adopted in some jurisdictions.

An institutional framework for the data economy has to do two things:

1) It needs to provide models for sharing data that increase its use, capture positive externalities and limit negative ones so we can maximise the value of data to society.
2) It needs to include institutions that together govern who can access what data. These institutions need to be trustworthy and act in accordance with the social and legal permissions they are given.

Elinor Ostrom’s framework for the management of shared resources offers particularly useful insights for data regulation and governance. Her work considers contexts where people need to reach agreement about rules of access to and management of a resource when some people will have to sacrifice private benefit for the greater common good.[24]

An example is an upstream farmer, who would benefit themselves from not sharing water for irrigation with those downstream, but would enable higher crop yields for their community as a whole if they do share. Similarly, the holder of data may need to sacrifice some private economic benefits by sharing data to unlock potentially much larger benefits for their sector or supply chain.

Ostrom's principles	Data economy parallel
There are clear boundaries and rules about who is entitled to what	Clarity on the rights of different entities to control, access, use and share data
Monitoring actions is feasible	Transparency and auditability of how data is being collected, used and shared
There are mechanisms for resolving conflicts	Regulators who can enforce both mandating and limiting access to data
Individual responsibilities and benefits broadly balance	Transparency and better understanding of both rights and how value from data returns to people and organisations
Users themselves are responsible for monitoring and enforcement	Transparency and contractual terms to enable monitoring and auditing of data use and sharing; in a data economy this may require agents who can act on behalf of data subjects
Sanctions for abuse are possible and graduated, getting progressively tougher	Enforcement of a range of consequences for the misuse of data, ranging from the withdrawal of access permissions to fines and other penalties
Decisions are legitimated by the participation of users	For individuals, consent and opt outs need to be informed and viable (which requires competitive alternative services) Organisations need to proactively engage with communities such as through representative data governance bodies and public participation exercises
Decisions are also legitimated by government recognition	A comprehensive data strategy and institutional/ regulatory framework

The table sets out Ostrom's design principles and their data economy parallels. Note that the commons – resources such as fish or grazing land – that Ostrom studies are rival, whereas data is non-rival.

What should policymakers and researchers do?

It is clear data could have much greater value to our economy, society and environment. It underpins developments in areas such as autonomous vehicles, drug discovery and energy efficiency. But the economic characteristics of data and the data economy mean the market will not naturally unlock this value. New technologies, such as deep learning or edge computing have and will continue to create new opportunities and challenges. Governments that adapt quickly, make the right investments and create the right incentives will have a significant advantage in a world where data is a new form of infrastructure.

There are substantial challenges to creating a thriving data economy:

- How to fund data as a public good, when it may need large up front investment
- How to incentivise investment in data when the benefits from it are then more widely shared
- How the benefits that do arise from using data should be fairly distributed
- How to compensate those who steward data for the costs and risks they take
- How to gain value from aggregated personal data while respecting people's privacy
- How to ensure data can be linked and combined to create positive externalities
- How to keep options open for potential future uses of data

Meeting these challenges calls for a strategic approach to data policymaking.

Incentivise investment without disincentivising sharing
There is a trade-off between wider access to data and incentivising investment in the creation of data and services. Exclusive access can give commercial advantages. Policymakers should re-examine existing legislation on intellectual property rights for data and consider other approaches such as time-limited exclusive rights, patent pools or compulsory licensing.

Limit exclusive access to public sector data
Selling exclusive access to public sector data provides a short-term financial gain but more open access usually provides greater long-term benefits. Policymakers should explore when exclusive access to public sector data is lawful and necessary. They should build confidence that deals involving public sector data will benefit the public, and be transparent about these deals.

Use competition policy to distribute value
Big incumbent companies currently capture a large proportion of the value of data as private profit. They can invest in collecting and using data and in specialist skills. However, preventing them from using data to provide valued services would be counterproductive. We recommend using competition policy to open data-driven markets to other providers.

Explore mandating access to private sector data
Policymakers should examine areas where mandating access to privately held data could enable innovation, competition and growth in priority policy areas. Public bodies should explore increasing access to data through procurement contracts and when issuing licences to operate services. Regulators should support and instigate initiatives that standardise access to data, as they have with open banking.

Provide a trustworthy institutional and regulatory environment
The value data has is dependent on the environment in which it exists. Institutions are needed to regulate who has access to data, monitor impact, and enforce compliance with regulation, technical standards and codes of conduct. The Information Commissioner's Office, sector-specific regulators, professional bodies and industry associations all have a role to play, including through the development of new data institutions that may be needed to create wider commercial and social value from data. And many of these organisations could benefit by working with international institutions, or national bodies in other countries, as the nature of data and the digital economy means that regulatory questions are best considered with a cross-border mindset that encourages international regulatory cooperation.

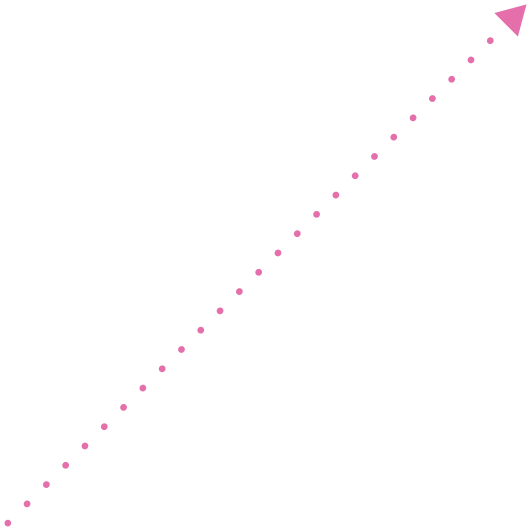
Simplify data regulation and licensing
Complex and overlapping regulation and intricate licensing schemes create uncertainties that hold back organisations from using and sharing data. Existing regulation should be simplified, new regulation should be coherent, and clear guidance should be provided.

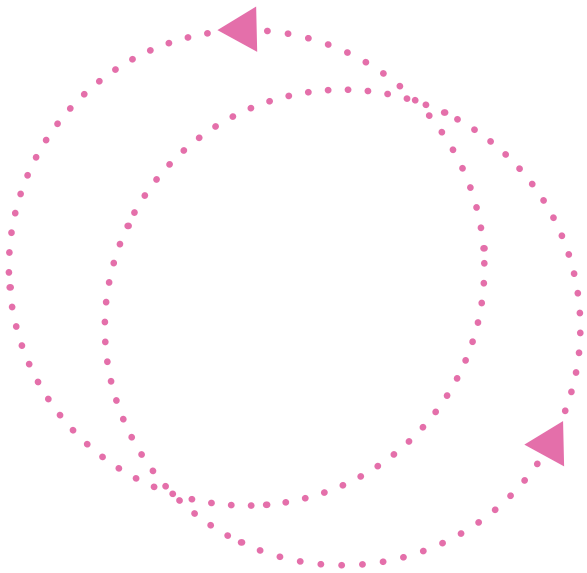
Monitor impacts and iterate
Changing the institutional and regulatory environment for data will also change the return on investment for collecting and cleaning data, and in the skills, software and other resources that help organisations make the most of data. These knock-on effects should be monitored. Experimentation in sectors or regions is useful for building evidence of what works.

Support further research
The full impact of data on our economies and societies is yet to play out. Further work is needed to:

- develop and apply methodologies and toolkits for understanding the value of data to the economy, the value of individual datasets, and how to value personal data and privacy
- explore how data is currently bought and sold in existing data markets and between the private and public sectors
- explore business models in the private sector and commercial models in the public sector that motivate providing access to data
- explore the distributional impacts, particularly on vulnerable populations, of the collection, use and sharing of data
- explore institutional models that can support trustworthy access to data, such as data trusts, pools, cooperatives and clubs
- bring existing work on institutional and regulatory economics to bear on data policy problems

The precautionary principle is often applied when there are unknown future risks. In data policy, where there are unknown future opportunities, we would argue an optionality principle should also apply. As the UK develops its National Data Strategy and makes investments in data, the data economy and AI, it should create the conditions for greater access, sharing and use of data, within a framework of regulation and trustworthy institutions.





Learn more

You can find out more about our work, and explore the findings summarised in this report at: bennettinstitute.cam.ac.uk/valuingdata

If you would like to contribute to our research, or discuss the value of data with us, please contact office@hermes.cam.ac.uk

References

1 PwC (2019), 'Putting a value on data', pwc.co.uk/issues/data-analytics/insights/putting-value-on-data.html

2 US Geological Survey (2013), 'Users, uses, and value of Landsat satellite imagery—Results from the 2012 survey of users', doi.org/10.3133/ofr20131269

3 Deloitte (2017), 'Assessing the value of TfL's open data and digital partnerships', content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf

4 UK Government (BEIS), 'Companies House data: valuing the user benefits.' (2019), gov.uk/government/publications/companies-house-data-valuing-the-user-benefits

5 Journal of Consumer Policy 42:425 (2019), 'How Much Is Data Privacy Worth? A Preliminary Investigation', doi.org/10.1007/s10603-019-09419-y

6 Proceedings of the National Academy of Sciences of the United States of America (2019), 'Using massive online choice experiments to measure changes in well-being', pnas.org/content/116/15/7250

7 Based on 2018 annual net income of $22bn and 2.2bn active monthly users. investor.fb.com/investor-news/press-release-details/2019/Facebook-Reports-Fourth-Quarter-and-Full-Year-2018-Results

8 Moneywise (2018), 'How much is your data worth to hackers?', moneywise.co.uk/news

9 The Verge (2019), 'Equifax agrees to settlement of up to $700 million over 2017 data breach', theverge.com/2019/7/22/20703497/equifax-ftc-fine-settlement-2017-data-breach-compensation-fund

10 Information Commissioner's Office (2018), 'ICO issues maximum £500,000 fine to Facebook for failing to protect users' personal information', ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2018/10/facebook-issued-with-maximum-500-000-fine

11 ODI, 'Anonymisation and open data: An introduction to managing the risk of re-identification' (2019). theodi.org/article/anonymisation-report

12 European Data Portal (2017), 'Analytical Report 9: The Economic Benefits of Open Data', europeandataportal.eu/en/highlights/economic-benefits-open-data

13 OECD (2019), 'Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies', doi.org/10.1787/276aaca8-en

14 ODI (2016), 'Research: The economic value of open versus paid data', theodi.org/article/research-the-economic-value-of-open-versus-paid-data

15 Open Knowledge Foundation (2019), 'What data counts in Europe? Towards a public debate on Europe's high value data and the PSI Directive', blog.okfn.org/tag/open-government-data-2

16 See for example blog.google/technology/research/open-source-and-open-data

17 HM Treasury (2019), 'Unlocking digital competition Report of the Digital Competition Expert Panel', gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel

18 UK Government (2017), 'Bus Services Act 2017: new powers and opportunities', gov.uk/government/publications/bus-services-act-2017-new-powers-and-opportunities

19 Office for National Statistics (2019), 'Accessing secure research data as an accredited researcher', ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme

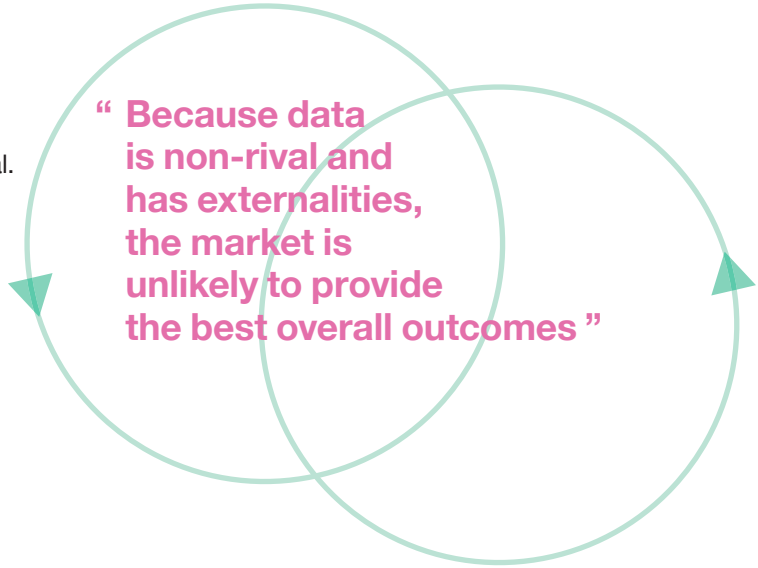
20 Databox Project (2019), databoxproject.uk

21 Solid, solid.mit.edu

22 Data Communications Company, smartdcc.co.uk

23 ODI (2019), 'Data trusts: Lessons from three pilots', theodi.org/article/odi-data-trusts-report

24 Cambridge University Press (1990), 'The Evolution of Institutions for Collective Action', doi.org/10.1017/CBO9780511807763



Bennett Institute for Public Policy Cambridge

ODI

Nuffield Foundation

About the Bennett Institute for Public Policy

The Bennett Institute for Public Policy, established in 2018, conducts high-level academic and policy research, as well as expanding the portfolio of public policy education and training offered at the University of Cambridge. The institute aims to become a world-leader in achieving successful and sustainable solutions to some of the most pressing problems of our time.

bennettinstitute.cam.ac.uk | **@BennettInst**

About the ODI

The Open Data Institute is an independent, non-profit, non-partisan company headquartered in London. The ODI was co-founded in 2012 by the inventor of the web Sir Tim Berners-Lee and Artificial Intelligence expert Sir Nigel Shadbolt to advocate for the innovative use of data to affect positive change across the globe. The ODI works with companies and governments to build an open, trustworthy data ecosystem, where people can make better decisions using data and manage any harmful impacts.

theodi.org | **@ODIHQ**

About the Nuffield Foundation

The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics and the Ada Lovelace Institute. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation.

nuffieldfoundation.org | **@NuffieldFound**

Credits

Authors: (Bennett Institute) Diane Coyle, Stephanie Diepeveen, Julia Wdowin. (ODI) Lawrence Kay, Jeni Tennison

Production and editing: Becky Ghani

Design and art direction: Adrian Philpott / @PhilpottDesign

Illustration: Ian Dutnall (page 5 and page 9)

Bennett Institute for Public Policy Cambridge

ODI

Nuffield Foundation

“Access to data determines how much of the potential value of data can be unlocked and who can benefit from that value”

bennettinstitute.cam.ac.uk/valuingdata