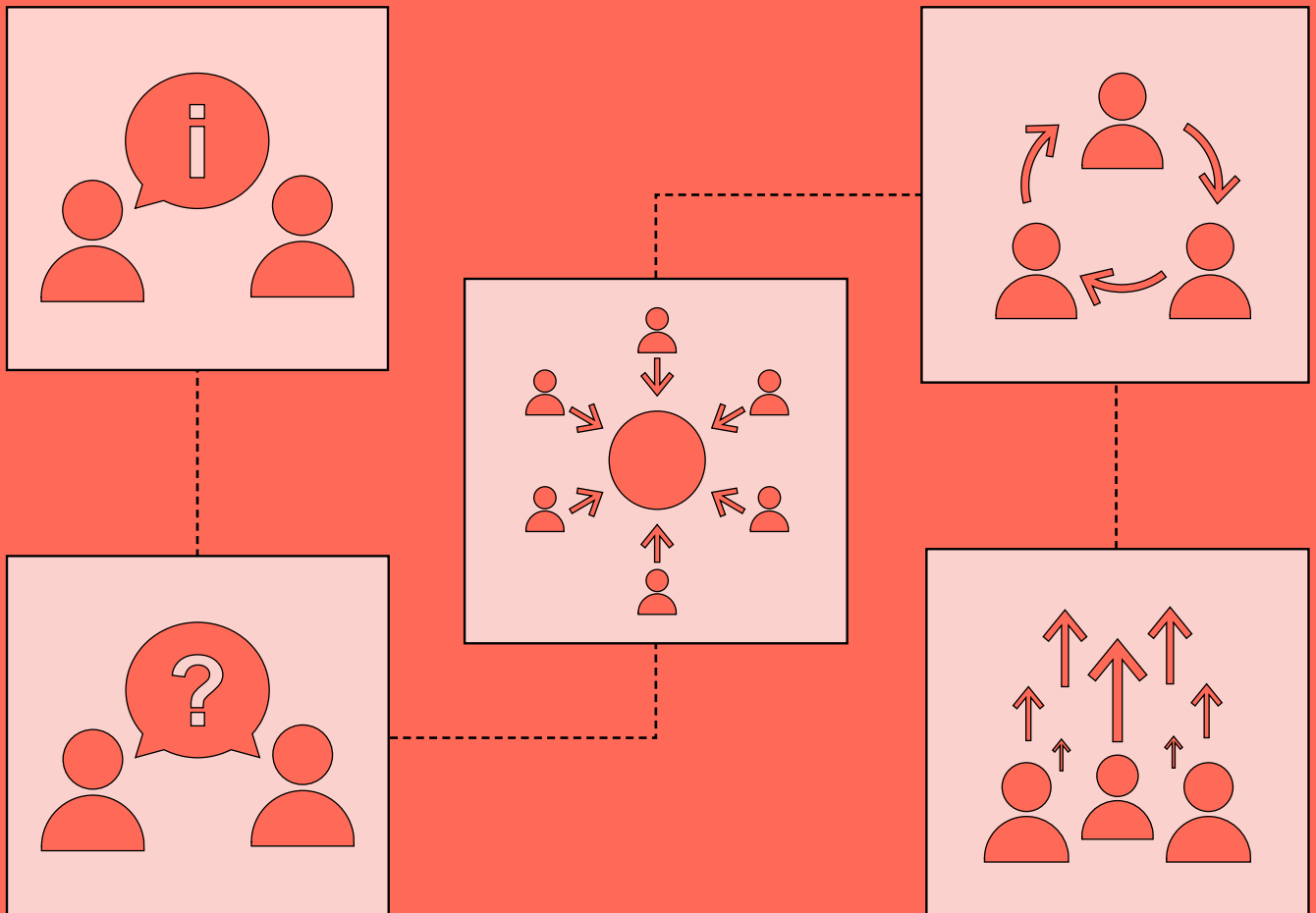


Participatory data stewardship

A framework for involving
people in the use of data

September 2021



Contents

- 3 Executive summary
- 6 Foreword
- 7 How to read this report
- 9 Introduction
- 17 Mechanisms for participatory data stewardship
- 48 Who to involve when developing participatory mechanisms
- 49 Benefits of effective participation for the design, development and use of data-driven systems
- 53 Conclusion
- 55 Methodology
- 56 Appendix
- 58 Acknowledgements

Executive summary

Well-managed data can support organisations, researchers, governments and corporations to conduct lifesaving health research, reduce environmental harms and produce societal value for individuals and communities. But these benefits are often overshadowed by harms, as current practices in data collection, storage, sharing and use have led to high-profile misuses of personal data, data breaches and sharing scandals.

These range from the backlash to Care.Data,¹ to the response to Cambridge Analytica and Facebook's collection and use of data for political advertising.² These cumulative scandals have resulted in 'tenuous' public trust in data sharing,³ which entrenches public concern about data and impedes its use in the public interest. To reverse this trend, what is needed is increased legitimacy, and increased trustworthiness, of data and AI use.

This report proposes a **'framework for participatory data stewardship'**, which rejects practices of data collection, storage, sharing and use in ways that are opaque or seek to manipulate people, in favour of practices that empower

1 Triggler, N. (2014). 'Care.data: How did it go so wrong?' *BBC News*. 19 Feb. Available at: <https://www.bbc.co.uk/news/health-26259101> [Accessed 6 Jul. 2021].

2 Fruchter, N., Yuan, B. and Specter, M. (2018). 'Facebook/Cambridge Analytica: Privacy lessons and a way forward'. *Internet Policy Research Initiative at MIT*. Available at: <https://internetpolicy.mit.edu/blog-2018-fb-cambridgeanalytica/>.

3 Centre for Data Ethics and Innovation. (2020). *Independent report: Addressing trust in public sector data use*. GOV.UK. Available at: <https://www.gov.uk/government/publications/cdei-publishes-its-first-report-on-public-sector-data-sharing/addressing-trust-in-public-sector-data-use> [Accessed 15 February 2021].

people to help inform, shape and – in some instances – govern their own data.

As a critical component of good data governance, it proposes **data stewardship** as the responsible use, collection and management of data in a participatory and rights-preserving way, informed by values and engaging with questions of fairness.

Drawing extensively from Sherry Arnstein's 'ladder of citizen participation'⁴ and its more recent adaptation into a spectrum,⁵ this new framework is based on an analysis of over 100 case studies of different methods of participatory data stewardship.⁶ It demonstrates ways that people can gain increasing levels of control and agency over their data – from being informed about what is happening to data about themselves, through to being empowered to take responsibility for exercising and actively managing decisions about data governance.

Throughout this report, we explore – using case studies and accompanying commentary – a range of mechanisms for achieving participatory decision-making around the design, development and use of data-driven systems and data-governance frameworks. This report provides evidence that involving people in the way data is used

4 Arnstein, S. (1969). 'A Ladder of Citizen Participation'. *Journal of the American Institute of Planners*, 35(4), pp.216-224. Available at: <https://www.tandfonline.com/doi/abs/10.1080/01944366908977225>

5 Patel, R. and Gibbon, K. (2018). 'Why decisions about the economy need you'. *The RSA*. Available at: <https://www.thersa.org/blog/2017/04/why-decisions-about-the-economy-need-you> [Accessed 15 February 2021].

6 Patel, R. and Peppin, A. (2020). 'Exploring principles for data stewardship'. Ada Lovelace Institute. Available at: <https://www.adalovelaceinstitute.org/project/exploring-principles-for-data-stewardship/> [Accessed 16 Feb. 2021].

can support greater social and economic equity, and rebalance asymmetries of power.⁷

It also highlights how examining different mechanisms of participatory data stewardship can help businesses, developers and policymakers to better understand which rights to enshrine, in order to contribute towards the increased legitimacy of – and public confidence in – the use of data and AI that works for people and society.

Focusing on participatory approaches to data stewardship, this report provides a complementary perspective to Ada's joint publication with the AI Council, *Exploring legal mechanisms for data stewardship*, which explores three legal mechanisms that could help facilitate responsible data stewardship.⁸

We do not propose participatory approaches as an alternative to legal and rights-based approaches, but rather as a set of complementary mechanisms to ensure public confidence and trust in appropriate uses of data, and – in some cases – to help shape the future of rights-based approaches, governance and regulation.

7 Kapoor, Astha and Whitt, Richard S. (2021). *Nudging Towards Data Equity: The Role of Stewardship and Fiduciaries in the Digital Economy*. February 22. Available at SSRN: <https://ssrn.com/abstract=3791845> or <http://dx.doi.org/10.2139/ssrn.3791845>

8 Ada Lovelace Institute. (2021). *Exploring legal mechanisms for data stewardship*. Available at: <https://www.adalovelaceinstitute.org/report/legal-mechanisms-data-stewardship/>

Foreword

Companies, governments and civil-society organisations around the world are looking at ways to innovate with data in the hope of improving people's lives through evidence-based research and better services. Innovation can move quickly, but it is important that people are given opportunities to shape it. This is especially true for areas where public views are not clearly settled.

This new report from the independent research institute and deliberative body the Ada Lovelace Institute provides practical examples of how to engage people in the governance of data through participatory data stewardship. The report shows that there are choices about when and how to bring people into the process – from data collection, to linkage, to data analysis.

One of the most frustrating experiences for people is when they are told they will have the power to shape something, but find in fact that consultation is very limited. To mitigate this the report helpfully distinguishes between different kinds of involvement, with a spectrum ranging from 'inform' all the way to 'empower'.

Participatory data stewardship provides a practical framework and case studies, to demonstrate how citizens can participate in shaping the way that data is being used. I hope businesses, policymakers and leaders of organisations will take inspiration from it, and generate a new set of use cases that we can continue to share in the future. Data innovation creates new opportunities and challenges that can take us beyond agreed social conventions. To make the most of the opportunities it is therefore imperative that people's voices are heard to shape how we use data.

Hetan Shah

Vice-Chair, Ada Lovelace Institute

Chief Executive, The British Academy

How to read this report

If you're a policymaker wishing to understand more about participatory approaches to data stewardship

- For an overview of the benefits of participatory approaches to data governance for people and society go to [p. 49](#).
- See [p. 14](#) for an illustration of the participatory framework.
- Read through each of the approaches, starting with the summary table on [p. 17](#), noting that each have their pros and cons, and that the most appropriate approach to take will vary depending on context, use of the data, the type of data being governed, and those who are most likely now and in future to be impacted by the data use and governance.
- Read the case studies on [pp. 22–46](#) for examples of how these mechanisms have been used in the real world, their participatory implications and complexities and critiques.

If you're responsible for data governance in your organisation and seeking to embed participatory approaches

- Go to [p. 49](#) for an overview of the benefits of taking a participatory approach to data stewardship.
- See [p. 48](#) for the key questions to answer about **who** to involve in the process.
- Go to [p. 17](#) for an indicative list of the range of approaches and tools data stewards can use to best to involve people in data governance.
- Read the case studies on [pp. 22–46](#) for examples of how these mechanisms have been used in the real world, their participatory implications and complexities and critiques.

If you're a civil-society organisation supporting the use of data for individual and collective benefit

- See [p. 14](#) for an illustration of the participatory framework.
- Read through each of the approaches, starting with the summary table on [p. 17](#), noting that each have their pros and cons, and that the most appropriate approach to take will vary depending on context, use of the data, the type of data being governed, and those who are most likely now and in future to be impacted by data use and governance.
- Read the case studies on [pp. 22–46](#) for examples of how these mechanisms have been used in the real world, their participatory implications and complexities and critiques.
- See [p. 48](#) for the key questions to answer about **who** to involve in the process.

If you're using or aware of mechanisms for participatory data stewardship

- See [p. 14](#) for an illustration of the participatory framework, and let us know whether it's a useful model.
- Read the case studies for each mechanism on [pp. 22-46](#), and let us know what you think about the implications for participation, and the summary of their complexities and critiques.
- See the table in the appendix listing 30 examples and get in touch to let us know if we are missing something.

Introduction

To understand how participatory data stewardship can bring unique benefits to data governance, we need to understand what it is and how it can be used.

1. Why participatory data stewardship matters

Organisations, governments and citizen-led initiatives around the world that aspire to use data to tackle major societal and economic problems (such as the COVID-19 pandemic) face significant ethical and practical challenges, alongside ‘tenuous’ public trust in public- and private-sector data sharing.⁹ To overcome these challenges, we will need to create robust mechanisms for data governance, and participatory data stewardship has a distinct role to play in their development.

Traditionally, data governance refers to the framework used to define who has authority and control over data and how that data may be used.¹⁰ It commonly includes questions around data architecture, security, documentation, retention, and access. Many organisations that use data or build data-driven systems implement a data-governance framework to guide their decision-making.

Institutions including data cooperatives, data trusts, data-donation systems and trusted research environments are designed to govern the use of beneficiaries’ data. Currently, private-sector data-governance approaches often do not address the concerns of the beneficiaries of data, and do not encourage those using that data to consider how their choices can best support the needs of those who will be affected by their decisions.

⁹ Centre for Data Ethics and Innovation. (2020).

¹⁰ Olavsrud, T. (2021). ‘Data governance: A best practices framework for managing data assets.’ *C/O*. <https://www.cio.com/article/3521011/what-is-data-governance-a-best-practices-framework-for-managing-data-assets.html>

‘Involving beneficiaries can encourage responsible innovation and improve data quality’

The report provides evidence that involving people (‘beneficiaries’) in the design, development and deployment of data governance frameworks can help create the checks and balances that engender greater societal and economic equity, can help to rebalance asymmetries of power, and can contribute towards increased public confidence in the use of data.¹¹

The term beneficiaries includes ‘data subjects’, who have a direct relationship with the data in question as specified in the GDPR,¹² and also encompasses those impacted by the use of data (e.g. workers, underrepresented and excluded groups) even if they are not themselves data subjects. In other words, we use the term ‘beneficiary’ to encompass anyone who might be affected by the use of data beyond simply the data subjects – those who have the *potential* to benefit from participatory data stewardship – and this helps to move beyond a compliance-based approach to a model that is underpinned by social license.

Beneficiaries can include:

- the data subjects – the people to whom the data directly relates, for instance, when processing biometrics data
- people within the wider public – for instance, those who might have an interest in how data is governed or used ethically, as well as those who might have lived experience of an issue or disadvantage
- people at risk of being oversurveilled, underrepresented or missing from the data itself, e.g. migrant populations, members of Indigenous communities, people from racialised minority groups, people with mental health conditions and transgender people
- stakeholders working in technology or related organisations, or members of a global supply chain. e.g. those engaged in collecting and processing data, or who have an interest in data to secure their own collective workplace rights.

In addition, involving beneficiaries can encourage responsible innovation and improve data quality, as the beneficial feedback loop below illustrates.

11 Kapoor, Astha and Whitt, Richard S. (2021). *Nudging Towards Data Equity: The Role of Stewardship and Fiduciaries in the Digital Economy*. 22 February. Available at SSRN: <https://ssrn.com/abstract=3791845> or <http://dx.doi.org/10.2139/ssrn.3791845>

12 The Information Commissioner’s Office defines the data subject as ‘the identified or identifiable living individual to whom personal data relates’. Available at: <https://ico.org.uk/for-organisations/data-protection-fee/legal-definitions-fees/#subject>

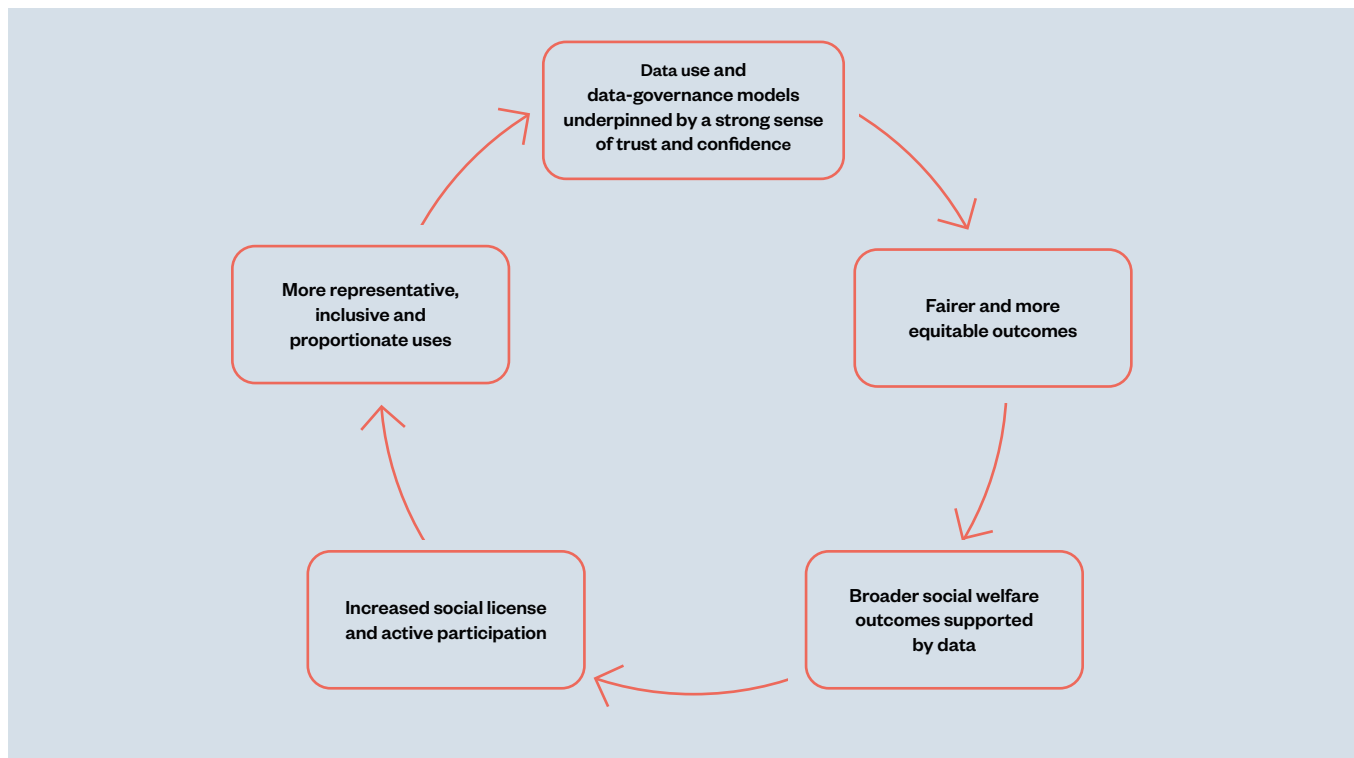


Figure 1: Effective participatory approaches generate a beneficial feedback loop

We outline the benefits of effective participation in the design, development and use of data and data-governance frameworks later in the report (see [p. 49](#)).

2. What do we mean by ‘stewardship’?

Stewardship is a concept that embodies the responsible planning and management of common resources.¹³ To apply this concept of stewardship to data, we must first recognise that data is not a ‘resource’, in the same way that forests or fisheries are finite but renewable resources. Rather, it is a common good that everyone has a stake in, and where the interests of beneficiaries should be at the heart of every conversation.

¹³ The use of stewardship to describe the governance of common resources was foundational to the work of Nobel Prize-winning economist Elinor Ostrom, who developed design principles for collective governance. Though often focused on shared natural resources like pastures, forests or fisheries, applying Ostrom’s principles to data can help us to think through thorny challenges of doing good with data. See Ostrom, E. (2015). *Governing the Commons*. Cambridge: Cambridge University Press. The Ada Lovelace Institute has mapped Ostrom’s principles against real-world examples where data is stewarded for social causes, or on behalf of data subjects, see: <https://docs.google.com/spreadsheets/d/1hAN8xMJuxobjARAWprZjtcZgqllwOIF7hf2UisiRBYU/edit#gid=1786878085>.

‘We understand data stewardship as key to protecting the data rights of individuals and communities’

The Ada Lovelace Institute has developed the following working definition of data stewardship:

‘The responsible use, collection and management of data in a participatory and rights-preserving way.’¹⁴

We understand data stewardship as key to protecting the data rights of individuals and communities, within a robust system of data governance, and to unlocking the benefits of data in a way that’s fair, equitable and focused on societal benefit. We contend that the principles and values that underpin stewardship can help to realise aspects of trustworthy and responsible data collection and use.

In this report, the term ‘data steward’ is used to describe the role of the individuals and organisations processing and using data on behalf of beneficiaries. In particular, stewards are responsible for involving the people who have a stake in and are impacted by data use and processing. That involvement is based on a relationship of trust and a social mandate to use data (often described in the legal context as a trust-based ‘fiduciary’ relationship – where the data steward has a responsibility to put people and society’s interests ahead of their own individual or organisational interests).

Data stewardship can operate throughout the data lifecycle (see figure 2 below). In the age of ‘datafication’,¹⁵ data stewardship operates not only in relation to data collection, processing, organising, analysis and use, but also in the design and development of data-driven systems, including those that aim to predict, prescribe, observe, evaluate and sometimes influence human and societal behaviour. This makes data stewardship a complex task, but also points to its potential to engender better outcomes for people and society.

14 See Ada’s post, *Disambiguating data stewardship*, for more on the development of this definition. Available at: <https://www.adalovelaceinstitute.org/blog/disambiguating-data-stewardship/>

15 ‘Datafication’ is a term that describes the process of rendering human and social behaviour in quantifiable ways, in turn releasing it as a new form of economic and social value. See the Ada Lovelace Institute’s report *The data will see you now* for a detailed explanation of how this operates in relation to health data. Available at: <https://www.adalovelaceinstitute.org/report/the-data-will-see-you-now/>

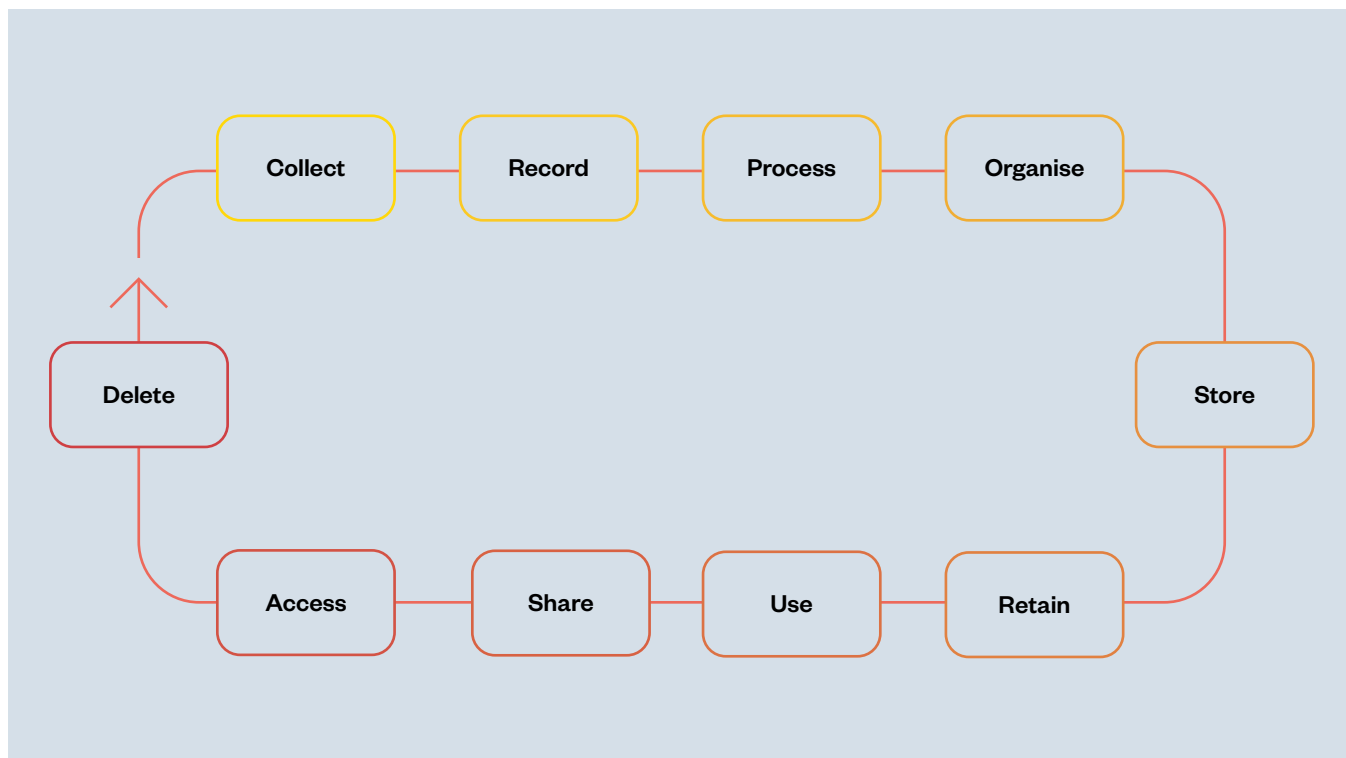


Figure 2: The data lifecycle

3. Introducing a framework for participatory data stewardship

Participation in its most fundamental sense is the involvement of people in influencing the decisions that affect their lives. In the technology-mediated environment that many of us currently inhabit, the ways data is used can preclude meaningful participation.

The Ofqual exam results algorithm that made predictions about A-levels in the UK, or the cookie notices and terms and conditions that ‘nudge’ towards uninformed consent at the expense of individual data rights, demonstrate how this can disempower people.¹⁶

In response to these conditions, we have developed a framework for participatory data stewardship (see figure 3 below) to demonstrate how it is possible to empower beneficiaries to affect the design, development and use of data-driven systems.

¹⁶ Office for Statistics Regulation. (2021). *Ensuring statistical models command public confidence*. Available at: <https://osr.statisticsauthority.gov.uk/publication/ensuring-statistical-models-command-public-confidence/> [Accessed 19 Aug 2021].

The framework is based on Sherry Arnstein's 'ladder of citizen participation',¹⁷ which illustrates that there are different ways to enable participation (see figure 5 in appendix), and its more recent adaptation into a spectrum of power,¹⁸ which represents the possible outcomes of informing, consulting, involving, collaborating and empowering people (see figure 6 in appendix).

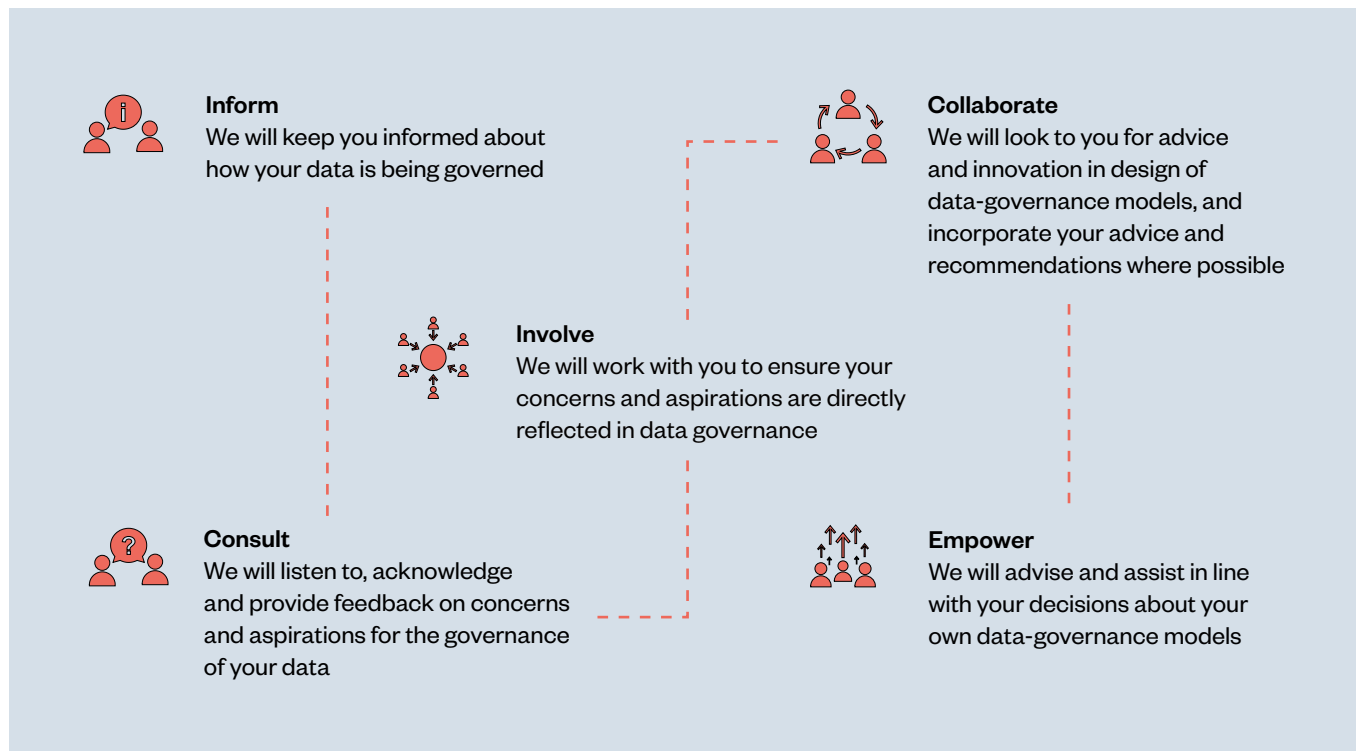


Figure 3: Framework for participation in data stewardship

We propose this framework to support thinking about how different modes of participation in and about data governance can enable beneficiaries to have increasing power and agency in the decisions made about their (and by extension others') data. Moving through the spectrum, the level of power afforded to beneficiaries increases from being the recipients of 'information' through to being 'empowered' to act with agency.

17 Arnstein, S. (1969). 'A Ladder of Citizen Participation'. *Journal of the American Institute of Planners*, 35(4), pp.216-224. Available at: <https://www.tandfonline.com/doi/abs/10.1080/01944366908977225>

18 Patel, R. and Gibbon, K. (2018). 'Why decisions about the economy need you'. *The RSA*. Available at: <https://www.thersa.org/blog/2017/04/why-decisions-about-the-economy-need-you> [Accessed 15 February 2021].

The framework first shows how participatory data stewardship mechanisms can seek to achieve meaningful transparency, responding to people's rights to be **informed** about what is happening or can happen with data about them.

Next, the framework describes mechanisms and processes that can build towards understanding and responding to people's views (**consult and involve**) in decision-making about data.

Ultimately, the vision is to realise conditions where people can **collaborate** actively with designers, developers and deployers of data-driven systems and (to the extent that is possible) are **empowered**. When this happens, beneficiaries' perspectives form a central part of the design of data governance, which then builds confidence and capacity for people to continue to participate in the data-governance process.

The next section of the report outlines the range of participatory mechanisms at data stewards' disposal.

It is important first to note that while participatory approaches may take different forms or have different intended outcomes, they are most often not mutually exclusive, and can in fact often complement each other. There is no single 'right' way to do participation, and effective participation is not a 'one-off' solution or mechanism.

The complex issues raised by data governance can't be solved by a 'one-size-fits-all' or an 'off-the-shelf' approach. Beneficiaries can participate at different stages in the data lifecycle – from collection, storage, cleaning and processing of data, as well as its use and deployment – and there are different types, approaches, methods or means of participation that afford very different levels of power.

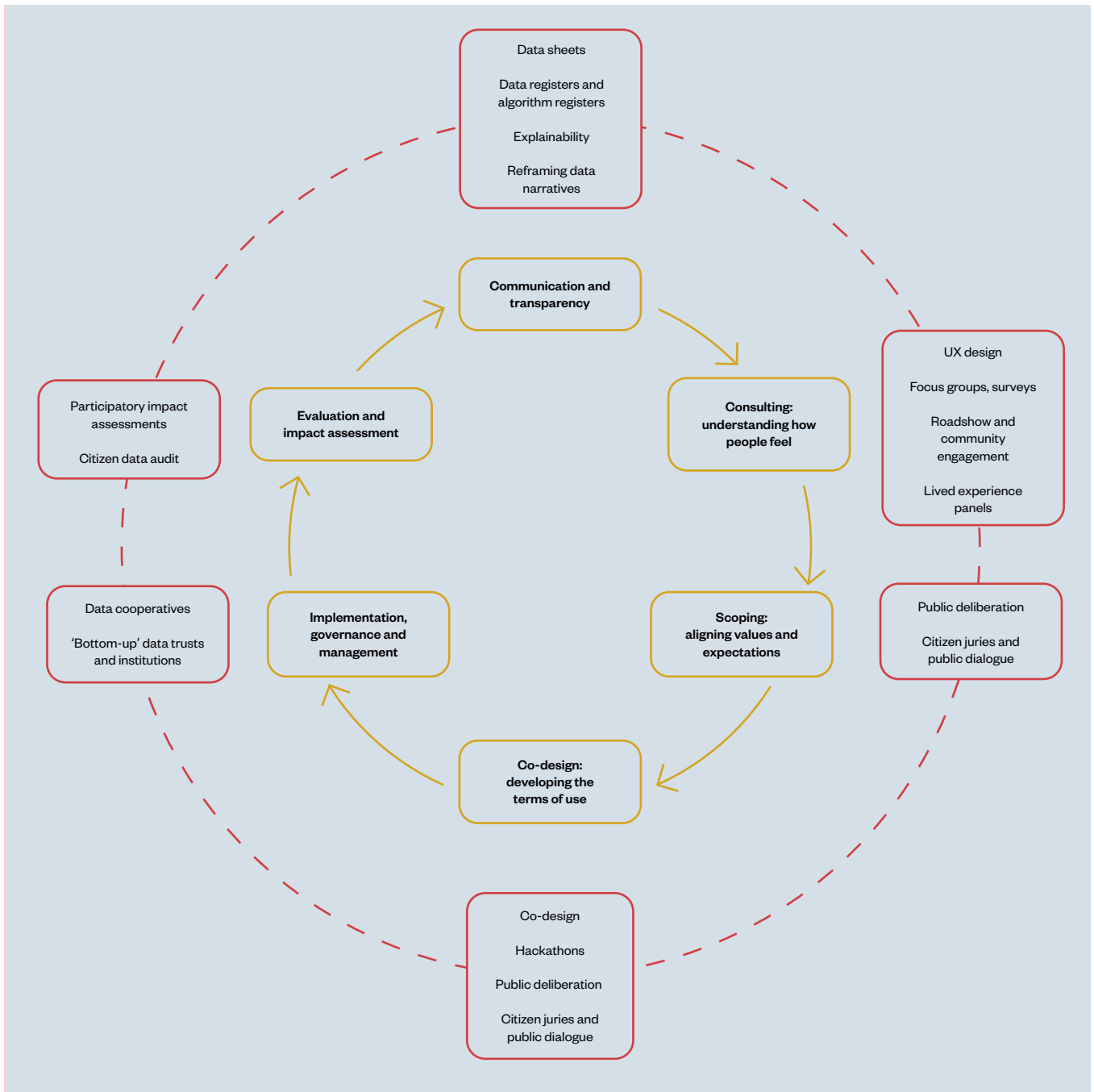


Figure 4: Purposes of different participatory mechanisms


Mechanisms for participatory data stewardship


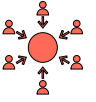


There are a wide range of different participatory mechanisms, methods and activities that can be used to support better data governance, depending on purpose and context (see figure 4).

These are summarised in table 1 below, and followed by a detailed description and analysis of each mechanism in detail, which explains how it works and explores the benefits, complexities and challenges through real-world case studies to help understand how these models operate in practice. These models form an indicative but not an exhaustive list, which we expect will iterate and evolve over time.

The various mechanisms and accompanying case studies are also illustrative of the creative potential of participation, and the range of approaches and tools at data stewards' disposal in thinking about how best to involve people in the use of data.

Table 1: Participatory mechanisms

Purpose	Description (drawn from Arnstein's ladder of participation)	What people can expect (drawn from the IAP2 Participation Spectrum)	Relevant participatory mechanisms, methods and activities	Case studies
INFORMING (p. 19) 	'A one way flow of information'	'We will keep you informed on how your data is being used'	<ul style="list-style-type: none"> • Meaningful and radical transparency • Explainability – the process of enabling a data-driven system to be explained in human terms • Mechanisms such as model cards and data sheets • Rethinking and reframing how we talk about data, so it is more accessible and inclusive 	<ul style="list-style-type: none"> • Enabling 'armchair audit' through open AI and algorithm registers (Helsinki and Amsterdam), administered by Saidot, 2020 • Making sense of the societal value of data by reframing 'big-data mining' and 'data as oil' narratives

Purpose	Description (drawn from Arnstein's ladder of participation)	What people can expect (drawn from the IAP2 Participation Spectrum)	Relevant participatory mechanisms, methods and activities	Case studies
CONSULTING <p>(p. 28)</p> 	<p>'Inviting people's opinions, through attitude surveys, neighbourhood meetings and public hearings'</p>	<p>'We will listen to, acknowledge concerns and aspirations, and provide feedback on how public input influenced the data-governance framework'</p>	<ul style="list-style-type: none"> Community networks User experience (UX) testing and co-design Surveys and public attitudes research Community engagements and consultations 	<ul style="list-style-type: none"> Community engagement and consultation by CityVerve (Manchester), multi-sector consortium led by Manchester City Council, 2015
INVOLVING <p>(p. 32)</p> 	<p>'Allow citizens to advise, but retain for powerholders the continued right to decide'</p>	<p>'We will work with you to ensure your concerns and aspirations are directly reflected in data governance... we will provide feedback on how public input influenced these decisions'</p>	<ul style="list-style-type: none"> One-off and institutionalised public deliberation and deliberative democracy initiatives Lived experience panels Horizon scanning, design thinking and futures thinking 	<ul style="list-style-type: none"> Institutionalised public deliberation, the Danish Board of Technology, 1986–2011
COLLABORATING <p>(p. 38)</p> 	<p>'Enables people to negotiate and engage in trade-offs with powerholders'</p>	<p>'We will look to you for advice and innovation in design of data-governance frameworks... and incorporate your advice and recommendations to the maximum extent possible'</p>	<ul style="list-style-type: none"> Public deliberation and deliberative democracy initiatives Bottom-up 'data-governance initiatives' managed by an independent fiduciary (e.g. data trusts, and data-sharing contracts that build in collaboration) Participant panels and data donation mechanisms 	<ul style="list-style-type: none"> Global mental health data bank, the Wellcome Trust with Sage Bionetworks, 2020 Large-scale data donation for research, UK Biobank, 2006–present
EMPOWERING <p>(p. 44)</p> 	<p>'Citizens obtain the majority of decision-making seats or full managerial power'</p>	<p>'We will provide advice and assistance as requested in line with your decisions for designing/developing your own data-governance framework'</p>	<ul style="list-style-type: none"> Data governance rules shaped and routinely reviewed by beneficiaries and data donors Voting on governance boards of data-access initiatives Ownership and/or control of data cooperatives Setting terms of data licensing and access, with permissions overseen by citizens 	<ul style="list-style-type: none"> Citizen-driven, collaborative data management and governance, Salus Coop (Spain), 2017–present

‘Transparency and explainability are distinct mechanisms that can contribute to informing beneficiaries meaningfully as to how data is used’

Inform

‘Informing’ people about data use and governance involves a one-way flow of information from those who use, gather, deploy and analyse data, to data subjects or ‘beneficiaries’. This information flow can be direct or indirect (through an intermediary, such as a data trust).

1. Meaningful transparency and explainability

Transparency and explainability are distinct mechanisms that can contribute to informing beneficiaries meaningfully as to how data is used. Transparency provides people with the necessary information and tools to be able to assess the use of data-driven systems and how their data is managed. Features of transparency include openness and clarity about the purpose, functions and effectiveness of these systems and frameworks. Transparency is enshrined in the GDPR as the ‘right to an explanation’, under Articles 13 and 14, as well as in Article 15 as the ‘right to access’ of data subjects for solely automated decision-making.¹⁹

An alternative way of expressing this one-way flow of information is as ‘explainability’. In the context of AI decision-making, explainability can be understood as the extent to which the internal mechanics of a machine or deep-learning system can be explained or made understandable in human terms. Enabling ‘explainability’ of complex uses of data (such as algorithmic or data-driven systems) has developed into its own established field of research, with contributions from the Alan Turing Institute and the Information Commissioner’s Office among many.²⁰

However, to improve outcomes for beneficiaries, transparency and explainability must move beyond simply informing people about how their own individual data has been used, or how a data-driven system has affected a decision about them, and towards the data beneficiary being able to influence the outcome of the data use through their increased understanding. In other words, it must be meaningful, and there must be a recognition that the rights of transparency and explainability can extend beyond the individual data subject.

19 Wachter, S. (2021). *A right to explanation*. Alan Turing Institute. Available at: <https://www.turing.ac.uk/research/impact-stories/a-right-to-explanation> [Accessed 15 February 2021].

20 See website from the Alan Turing Institute on Project ExplAIIn. Leslie, D. (2019). *Project ExplAIIn*. The Alan Turing Institute. Available at: <https://www.turing.ac.uk/news/project-explain> [Accessed 15 February 2021].

What is *meaningful* transparency and explainability?

What meaningful transparency and explainability look like will depend on the type of data used, the intent and purpose for its use, and the intended audience of the explanation.

Some researchers have distinguished between **model-centric explanations** (an explanation of the AI model itself) for general information-sharing and broader accountability purposes; and more **subject-centric explanations** (explanations of how a particular decision has impacted on a particular individual, or on certain groups).²¹

Subject-centric explanations are usually the first step towards increased accountability. The UK's Information Commissioner's Office (ICO) has identified six main types of explanation:²²

- **Rationale:** the reasons that led to a decision, delivered in an accessible and non-technical way.
- **Responsibility:** who is involved in the development, management and implementation of an AI system, and who to contact for a human review of a decision.
- **Data:** what data has been used in a particular decision and how.
- **Fairness:** steps taken across the design and implementation of an AI system to ensure that the decisions it supports are generally unbiased and fair, and whether or not an individual has been treated equitably.
- **Safety and performance:** steps taken across the design and implementation of an AI system to maximise the accuracy, reliability, security and robustness of its decisions and behaviours.
- **Impact:** steps taken across the design and implementation of an AI system to consider and monitor the impacts that the use of an AI system and its decisions has or may have on an individual, and on wider society.

21 Ruiz, J. (2018). 'Machine learning and the right to explanation in GDPR'. *Open Rights Group*. Available at: <https://www.openrightsgroup.org/blog/machine-learning-and-the-right-to-explanation-in-gdpr/>. [Accessed 15 February 2021].

22 ICO. (2020). *What goes into an explanation?* Available at: <https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence/part-1-the-basics-of-explaining-ai/what-goes-into-an-explanation/> [Accessed 5 Jul. 2021].

One example of meaningful transparency in practice is the use of ‘algorithm registers’ implemented by the cities of Amsterdam and Helsinki

In addition to explainability, meaningful transparency refers to initiatives that seek to make information about an AI system both visible and interpretable to audiences who may have different levels of competency and understanding of technical systems.²³ These initiatives identify clear ‘targets’ of transparency, such as an algorithm, the ‘assemblage of human and non-human actors’ around algorithmic systems, or the governance regime that oversees the management and use of that system.²⁴

As our article *Meaningful transparency and (in)visible algorithms* demonstrates, meaningful transparency initiatives are ones that ‘build publicness’ by aligning AI systems with values traditionally found in the public sector, such as due process, accountability and welfare provision.²⁵ These initiatives provide people with the necessary tools and information to assess and interact with AI systems.

One example of meaningful transparency in practice is the use of ‘algorithm registers’ implemented by the cities of Amsterdam and Helsinki. These registers provide a list of data-driven systems put into use by these cities, and provide different levels of information aimed at audiences with differing competencies and backgrounds. These registers not only disclose key facts about the use of data-driven systems, but also enable members of the public to perform their own independent ‘armchair audit’ into how these systems may affect them.

Other examples of emerging transparency mechanisms include model cards, which aim to document key details about an AI system’s intended uses, features and performance.²⁶ Data sheets are a similar mechanism that aim to summarise essential details about the collection, features, and intended uses of a dataset.²⁷ To date, data sheets and model cards have been used primarily by AI researchers and companies as a

23 Ananny M, Crawford K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*. 20(3):973-989. doi:10.1177/1461444816676645

24 M. Kaminski. (2020). Understanding Transparency in Algorithmic Accountability. Forthcoming in *Cambridge Handbook of the Law of Algorithms*, ed. Woodrow Barfield, Cambridge University Press (2020), U of Colorado Law Legal Studies Research Paper No. 20-34, Available at SSRN: <https://ssrn.com/abstract=3622657>

25 L. Stirton, M. Lodge. (2002). Transparency Mechanisms: Building Publicness into Public Services. *Journal of Law and Society*. <https://doi.org/10.1111/1467-6478.00199>

26 See, for example the model cards from Google Cloud: <https://modelcards.withgoogle.com/about>

27 Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J.W., Wallach, H., Daumeé III, Hal and Crawford, K. (2018). *Datasheets for Datasets*. Available at: <https://arxiv.org/abs/1803.09010>.

means of transferring information about a dataset or model between teams within or outside of an organisation. While, to date, they are not intended to provide information to members of the public, they could be repurposed.

Case study: Enabling ‘armchair audit’ through open AI and algorithm registers (Helsinki and Amsterdam), administered by Saidot, 2020²⁸

Overview:

Helsinki and Amsterdam were among the first cities to announce open artificial intelligence (AI) and algorithm registers. These were founded on the premise that the use of AI in public services should adhere to the same principles of transparency, security and openness as other city activities, such as public bodies’ approaches to procurement.

The infrastructure for open AI registers in both countries is being administered by the Finnish commercial organisation, Saidot. They are aiming to build a scalable and adaptable register platform, to collect information on algorithmic systems and share them flexibly with a wide range of stakeholder groups.

Participatory implications:

These AI and algorithm registers are openly available, and any individual is able to check them. This illuminates the potential for ‘armchair auditing’ of public-service algorithmic systems. Accessing the register reveals information about the systems that are reported on, the information that is being provided, and the specific applications and contexts in which algorithmic systems and AI are being used.

Saidot conducted research and interviews with clients and stakeholders to develop a model that serves the wider public, meaning it’s accessible not only to tech experts but also to those who know less about technology, or are less interested in it. The approach adopts a layered model, enabling any individual to find and discover more information based on their level of knowledge and interest.

28 Dimitrova, A., (2020). ‘Helsinki and Amsterdam with first ever AI registries’ *TheMAYOR.eu* Available at: <https://www.themayor.eu/en/a/view/helsinki-and-amsterdam-with-first-ever-ai-registries-5982> [Accessed 15 Feb. 2021].

To develop meaningful transparency and explainability of data-driven and AI systems, it is necessary to enable beneficiaries to gain insights into the decision-making processes that determine their goals

Complexities and critiques

Given the relatively recent emergence of these approaches, more needs to be understood about the effectiveness of these systems in enabling 'meaningful transparency' in the way defined above. Do these registers, for example, genuinely enable non-specialist people to become armchair algorithmic auditors? And are they complemented by the implementation of, and access to, independent audits, which are also important transparency mechanisms?²⁹

To develop meaningful transparency and explainability of data-driven and AI systems, it is necessary to enable beneficiaries to gain insights into the decision-making processes that determine their goals, the impact of their outcomes, how they interact with social structures and what power relations they engender through continued use. This may sound straightforward, but these are not static pieces of information that can be easily captured; they require ongoing investigation using a multitude of sources.

Lack of access to complete, contextual information about complex uses of data is often due to information being spread across different organisational structures and under different practices of documentation. Any information architecture set up to promote transparency must contend with the way data-governance frameworks relate to the real world, tracking the shifting influences of technology, governance and economics, and the public and private actors embedded within them. This requires an articulation of *how* to access information as well as *what* information to access. Answering the 'how' question of transparency will require addressing the conditions around who executes it, with what motives and purposes.³⁰

Some critics have argued that relying exclusively on transparency and explainability as mechanisms for engaging and involving people in data governance risks creating an illusion of control that does not exist in reality. This might distract from addressing some more harmful data practices that engender unfairness or discriminatory outcomes.³¹

29 Dimitrova, A. (2020).

30 Safak, C. and Parker, I. (2020). 'Meaningful transparency and in(visible) algorithms'. *Ada Lovelace Institute*. Available at: <https://www.adalovelaceinstitute.org/blog/meaningful-transparency-and-invisible-algorithms/>

31 Edwards, L. and Veale, M. (2018). 'Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?' *IEEE Security & Privacy*, 16(3), pp.46–54. Available at: <https://arxiv.org/pdf/1803.07540.pdf>

Technologists and policymakers frequently reach for analogies and metaphors, in efforts to communicate with diverse publics about data

Others have highlighted that increased transparency about data use and management forces policymakers to be more explicit and transparent about the trade-offs they are choosing to make. In this context, it has been argued that explainability doesn't just pose a technical challenge, but also a policy challenge about which priorities, objectives and trade-offs to choose.³²

2. Rethinking and reframing the way we communicate about data

Another core component that informs the way we collectively conceptualise, imagine and interact with data are the narratives we build to describe how data operates in societies.

Narratives, fictional or not, have a profound effect on shaping our collective understandings of all aspects of our world, from the political or economic to the technological. They rely on a range of devices: news articles, visual images, political rhetoric and influential people all contribute to the prevalence of a particular narrative in the public consciousness.

Central to many narratives is the use of metaphors, and that's particularly true when it comes to technology and data, where the abstract nature of many concepts means that metaphors are used to conceptualise and understand processes and practices.³³

Technologists and policymakers frequently reach for analogies and metaphors, in efforts to communicate with diverse publics about data. Sociologists Puschmann and Burgess have explored how metaphors are central to making sense of the often abstract concept of data.³⁴ If we are to move to more participatory mechanisms for data stewardship and governance, we will need to understand and unpick how these specific metaphors and narratives shape data practices.

32 Coyle, D., & Weller, A. (2020). "Explaining" machine learning reveals policy challenges'. *Science*, 368 (6498), pp.1433-1434. Available at: <https://doi.org/10.1126/science.aba9647>

33 Van den Boomen, M. (2014). *Transcoding the digital: How metaphors matter in new media*. Instituut voor Netwerkcultuur. Available at: <https://dspace.library.uu.nl/handle/1874/289883>

34 Puschmann, C. and Burgess, J. (2014). 'Big Data, Big Questions | Metaphors of Big Data', *International Journal of Communication*, 8(0), p. 20. Available at: <https://ijoc.org/index.php/ijoc/article/view/2169>

Current narratives create policies and technologies that too-often minimise the potential societal benefit of data while facilitating commercial incentives

Why narratives about data matter

How we talk about data influences the way we think about it. How we think about data influences how we design data policies and technologies. Current narratives create policies and technologies that too-often minimise the potential societal benefit of data while facilitating commercial incentives. To reimagine narratives, we need to understand and challenge the metaphors and framings that are already prominent.

Data has been called everything from ‘the new oil’, to water or radioactive waste.³⁵ Metaphors like this are used to make the concept of data more tangible, more open and more accessible to the public, but they ascribe qualities to data that aren’t necessarily inherent, and so can have the effect of mystifying or creating disempowering perceptions of data. Metaphors can frame issues in particular ways, highlighting certain qualities while obscuring others, and this can shape perspectives and belief according to the intentions of those setting the narrative.

Common understandings created through metaphor are all-too-often misunderstandings. For example, as ‘oil’, data’s economic value is brought into focus, but its non-rivalrous quality (the fact that it can be used simultaneously by different people for different purposes) is hidden.³⁶ These misunderstandings about the ways we describe data can reveal much about the contemporary social and cultural meaning we give to it.³⁷ For example, when economic value is highlighted, and common good is obscured, the market motivations to make ‘commodity’ the meaning of data are revealed.

Such narratives also suggest that practices about data are, like forces of nature, unfixed and unchangeable – rather than social practices undertaken by people and power holders. And it’s important to note that

35 Rajan, A. (2017). ‘Data is not the new oil’, *BBC News*. Available at: <https://www.bbc.com/news/entertainment-arts-41559076>; Lupton, D. (2013). ‘Swimming or drowning in the data ocean? Thoughts on the metaphors of big data’, *This Sociological Life*, Available at: <https://simplysociology.wordpress.com/2013/10/29/swimming-or-drowning-in-the-data-ocean-thoughts-on-the-metaphors-of-big-data/>; Doctorow, C. (2008). ‘Cory Doctorow: why personal data is like nuclear waste.’ *The Guardian*. Available at: <http://www.theguardian.com/technology/2008/jan/15/data.security> (All accessed: 15 February 2021).

36 Coyle D., Diepeveen S., Wdowin J., et al. (2020). *The value of data: Policy Implications*. Bennett Institute for Public Policy Cambridge, the Open Data Institute and Nuffield Foundation. Available at: https://www.bennettinstitute.cam.ac.uk/media/uploads/files/Value_of_data_Policy_Implications_Report_26_Feb_ok4noWn.pdf

37 Lupton D. (2013). ‘Swimming or drowning in the data ocean? Thoughts on the metaphors of big data’, *This Sociological Life*, 29 October. Available at: <https://simplysociology.wordpress.com/2013/10/29/swimming-or-drowning-in-the-data-ocean-thoughts-on-the-metaphors-of-big-data/> (Accessed: 15 February 2021).

'data' itself is not the source of value – but rather actors' positioning and abilities to make use of data is what brings value.

But focusing on one single metaphor and its pitfalls reveals only part of the power of narratives. Researching the metaphors that exist across a range of prominent data narratives, Puschmann and Burgess identify two interconnected metaphorical frames for data: data as a *natural force* to resist or control; and data as a *resource to be consumed*. As a *natural force*, data 'exists' in the world – like water, gravity or forests – just waiting to be discovered and harnessed by humans. As a *resource to be consumed*, data is conceptualised as a commodity that can be owned and sold, then used – just like oil or coal.

These two metaphorical frames promote the idea that data is a value-neutral commodity, and so shape data practices to be extractive and promote commercial and competitive incentives. Many other sociocultural meanings exist, or could exist, manifested through metaphor. Recognising and understanding them allows us to interpret these sociocultural meanings, and address how they shape data policies, processes and practices.

Case study: Making sense of the societal value of data by reframing 'big-data mining' and 'data as oil' narratives

Overview:

One of the major metaphorical frames that embodies data as a *natural force* or *resource to be consumed* is that of 'big-data mining', a narrative that emerged in a range of corporate literature in the 1990s. This framed big data as a new 'gold mine' waiting to be tapped, from which companies could gain wealth if they knew how to extract it.³⁸

Positioning 'data' as a resource from which insight can be 'mined' perpetuates the conceptual model that economic value can be extracted from data, leveraging the culturally embedded analogy that gold ore mined from the earth

38 Kerssens, N. (2019). 'De-Agentializing Data Practices: The Shifting Power of Metaphor in 1990s Discourses on Data Mining', *Journal of Cultural Analytics*, 1(1), p. 11049. Available at: <https://dspace.library.uu.nl/handle/1874/387149>

(for example, in the American Midwest) delivered untold wealth for colonialist settlers and ignores the unequal power dynamics of the extractive practices that underpin these narratives.

Participatory implications:

The reality is that data does not exist in the natural world, lying dormant until it is discovered and its value extracted, and the focus this metaphor places on wealth value obscures the societal value data can have if stewarded well. To overcome these framings it's necessary for those affected by data to participate in rethinking and reframing how data is conceptualised, to bring data narratives (and therefore practices) in greater alignment with public expectations and values.

Following existing metaphors like oil and mining, the natural world remains a good conceptual foundation to develop alternative metaphors. One option is to see data as a greenhouse gas, where its uses create harmful byproducts that must be limited.³⁹ Other metaphors might recognise that oil, gold and other natural commodities are not perceived and treated in the same way as forests, rivers and sunlight. These more 'ecological' metaphors have been posed as alternatives that could promote concepts and practices where data is stewarded more sustainably and less extractively.

Complexities and critiques – why reframing doesn't go far enough

Reframing old narratives, and creating new narratives is no easy task. Narratives are slow to develop and permeate, and require critical mass and collective understanding to gain the traction needed to change our deeply conditioned mental models of the world. Introducing new metaphors is challenging when so many, often-contradictory, ones already exist. Moreover, creating new metaphors from the 'top down' would continue to represent asymmetries of power by perpetuating a model where data stewards – governments, corporations and other institutions – continue to impose their world view on data subjects, especially those already marginalised. Rethinking the narratives and metaphors that shape how data is conceptualised must therefore be part of a participatory process.

39 Tisne M. (2019). 'Data isn't the new oil, it's the new CO2'. *Luminate Group*. Available at: <https://luminategroup.com/posts/blog/data-isnt-the-new-oil-its-the-new-co2> (Accessed: 15 February 2021).

A range of data
stewardship mechanisms
build consultation
into their design
and development

Consult

Understanding public attitudes to different uses of data

A range of data stewardship mechanisms build consultation into their design and development. Consultation can take place with individuals, groups, networks or communities, to enable people to voice their concerns, thoughts and perspectives. Consultation activities can take a range of forms but often involve the use of quantitative and qualitative methods including public-opinion ‘attitude’ surveys, neighbourhood meetings and public hearings.

While the principles of user-led design aim to understand a narrower set of perspectives – those of the intended *user* of the data, rather than the views of beneficiaries or the public – they can contribute to consultation’s aim by ensuring designers and developers of data-driven systems understand people’s aspirations and concerns.

What is meaningful consultation?

The process of seeking to understand people’s attitudes to different uses and approaches to data, commonly referred to as consultation, has undergone a shift in acceptance. It was dismissed by Arnstein as a tokenistic activity, because it carried with it no meaningful promise of effecting power shifts or changes.⁴⁰ However, eighteen years later, the Gunning Principles – relating to consultation in two landmark British court cases of Coughlan (2001) and Brent London Borough *ex parte* Gunning (1989) – sought to make a distinction between *ineffectual* consultation of the type Arnstein had in mind and *effective* consultation, by setting out guiding principles.

40 Arnstein, S. (1969). ‘A Ladder of Citizen Participation’. *Journal of the American Institute of Planners*, 35(4), pp.216-224. Available at: <https://www.tandfonline.com/doi/abs/10.1080/01944366908977225>

In the United Kingdom, the Gunning Principles have formed a strong legal foundation for assessing the legitimacy of public consultations

The Gunning Principles of Consultation:⁴¹

- **Consultations should be undertaken when proposals are still at a formative stage:** Bodies cannot consult on a decision already made. Organisations need to have an open mind during a consultation and a willingness to listen to alternatives to any proposals they have put forward.
- **There must be sufficient information to permit ‘intelligent consideration’ by the people who have been consulted:** People involved in the consultation need to have enough information to make an intelligent choice and input in the process.
- **Adequate time for consideration and response:** Timing is crucial – participants must have sufficient time and a supportive environment to consider issues and make an informed response.
- **Responses must be conscientiously taken into account:** Bodies must be able to demonstrate (with reference to evidence) that they have listened to and fully considered responses from those they engaged prior to making any final decision. They must also provide motivation as to why certain responses were taken into consideration and others were not and for what reasons.

In the United Kingdom, the Gunning Principles have formed a strong legal foundation for assessing the legitimacy of public consultations, and are frequently referred to as a legal basis for judicial review decisions (the process by which UK law reviews the validity of decisions made by public bodies). They also act as a valuable mechanism through which to understand how impactful and meaningful a consultation process is likely to have been.

In addition to the principles above, a range of consultation-based approaches have sought to acknowledge significant asymmetries of power in relation to the perspectives of people at risk of being disproportionately impacted by the use of data, and/or who are at risk of being excluded or underrepresented by data-driven systems. This is especially important in a context where ‘publics’ themselves are diverse and not equal.

41 The Consultation Institute. (2018). *The Gunning Principles – Implications* (2018). Available at: <https://www.consultationinstitute.org/the-gunning-principles-implications/> [Accessed 15 Feb. 2021].

The use of lived-experience panels (groups that work to shape an initiative, drawing from their own personal experience of an issue, such as racial or social injustice) seek to ensure structural inequalities and dynamics of power and privilege are considered in the design of data-driven systems and governance frameworks. These can form an important part of good and effective consultation. Examples include the Ada Lovelace Institute's approach to understanding issues of power and privilege impacting on people from minority-ethnic backgrounds, on issues of access and gender diversity;⁴² and the emphasis on lived-experience panels by the Wellcome Trust, by engaging with people with mental health conditions to inform the establishment of a global mental-health databank (see case study, p.39).⁴³

Globally, much has been made of the importance of co-creation and enabling agency over data in ways that are led and co-designed *with* a range of Indigenous communities rather than designed without and imposed on the communities,⁴⁴ and these initiatives draw on similar insights. There is good practice emerging across New Zealand and Canada in this context that demonstrates how to move away from the invisibility of Indigenous peoples in data systems towards a 'people-and-purpose orientation to data governance' that promotes active involvement and co-creation of systems in ways that are culturally and societally relevant and sensitive.⁴⁵

-
- 42 Patel, R. and Peppin, A. (2020). 'Making visible the invisible: what public engagement uncovers about privilege and power in data systems'. *Ada Lovelace Institute*. Available at: <https://www.adalovelaceinstitute.org/blog/public-engagement-uncovers-privilege-and-power-in-data-systems/> [Accessed 5 Jul. 2021].
- 43 Wolpert, M. (2020). 'Global mental health databank pilot launches to discover what helps whom and why in youth anxiety and depression'. *LinkedIn*. Available at: <https://www.linkedin.com/pulse/global-mental-health-databank-pilot-launches-discover-wolpert/?trackingId=uh3v1yO AQ3WGMK1Ssvq9JA%3D%3D> [Accessed 5 Jul. 2021].
- 44 Carroll, S.R., Hudson, M., Holbrook, J., Materechera, S. and Anderson, J. (2020). 'Working with the CARE principles: operationalising Indigenous data governance'. *Ada Lovelace Institute*. Available at: <https://www.adalovelaceinstitute.org/blog/care-principles-operationalising-indigenous-data-governance/> [Accessed 5 Jul. 2021].
- 45 Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodríguez-Lonebear, D., Rowe, R., Sara, R., Walker, J.D., Anderson, J. and Hudson, M., (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal*, 19(1), p.43. Available at: <https://datascience.codata.org/articles/10.5334/dsj-2020-043/>

Case study: Community engagement and consultation by CityVerve (Manchester), multi-sector consortium led by Manchester City Council, 2015⁴⁶

Overview:

The CityVerve project was a Smart City demonstrator, funded by the UK Government and Innovate UK, in the UK city of Manchester. It was a two-year project to develop and test new Internet of Things (IoT) services at scale, adding sensors to equipment throughout the city that collected and shared data across a network. It combined health and social-care data, environment data and transport data to analyse energy use and monitor air quality in cities, and to provide integrated and personalised healthcare feedback to individuals and clinicians about people's symptoms.

A key aim of CityVerve was to build a 'platform of platforms' to gather and collate data across Manchester about the needs of a UK city. This was built by telecoms company Cisco, and the data was stored by BT's cloud-based 'Internet of Things Datahub'.

Participatory implications:

The CityVerve team used a creative approach to consultation, which attracted 1,000 workshop attendees. They reached a further 11,300 with art and performance, commissioning local artist, Naho Matsuda, to bring the project to life through an innovative digital installation.

The team adopted user-design approaches and developed a local communities' platform, which was built around interests and activities relating to the IoT, and also ran community forums to provide people with an introduction to networked data. One of the project partners, Future Everything, led human-centred design workshops with participants to discuss and explore IoT technology pilots and design use cases for the data being collected – highlighting the extent to which 'consultation' can be creative and innovative, helping to inform and build the capacity of those who participate.

46 University of Manchester Digital Futures. (2015). *The CityVerve Project*. Available at: <http://www.digitalfutures.manchester.ac.uk/case-studies/the-cityverve-project/> [Accessed 19 Aug. 2021].

Complexities and critiques

A critique of consultative approaches and user-design approaches is that, too often, the feedback loop between the people's input and the decisions made about design and implementation fails to operate effectively or clearly, negating participants' ability to have or exercise 'real power'. When this happens, consultation can serve to legitimate decisions that would have been made regardless, or can design in a manner that is 'leading', exclude marginalised perspectives or fail to capture the full complexity and nuance of an issue. Critics also express concern about the process misleading participants into thinking they have greater power and agency over the terms and conditions of the data governance initiative than they actually have, creating a potential misalignment between expectations and the reality of the consultation's likely impact.

The controversy around Google affiliate Sidewalks Labs' efforts to develop Toronto's Waterfront is a good example of this.⁴⁷ The now-abandoned proposal to deploy a civic data trust as a participatory mechanism was demonstrated to be insufficient and inadequate in reassuring members of the public of the acceptability of the partnership between Sidewalks Lab and Waterfront Toronto. Instead of assuming that the partnership itself was acceptable and proposing engagement within that context, Waterfront Toronto could have considered engaging publics on their expectations prior to entering into it – and then sought to embed consultation within the partnership.

Involve

Beneficiaries advise through deliberation

The process of 'involvement' positions beneficiaries in an advisory role, which helps inform early-stage decision-making by data stewards and fiduciaries. These initiatives convene non-specialist beneficiaries alongside specialists and stakeholders, with a view to informing key moments in the public-policy landscape. They seek to advise and better inform government and regulatory bodies on the conditions for the

⁴⁷ Hawkins, A.J. (2020). 'Alphabet's Sidewalk Labs shuts down Toronto smart city project'. *The Verge*. Available at: <https://www.theverge.com/2020/5/7/21250594/alphabet-sidewalk-labs-toronto-quayside-shutting-down>.

acceptability of uses of AI and data (particularly where the use of data is contested, controversial and where clear regulatory norms have yet to be established).

Often, enabling an advisory role for beneficiaries in shaping approaches to data governance draws on the values, norms and principles of deliberative democracy. Through the established methodologies of deliberation, beneficiaries and key stakeholders are provided with access to privileged information and equipped at an early stage with sufficient time and information to play a central role in shaping approaches to data governance.⁴⁸

Doing this well can take time (through a long-form process, convening and re-convening people over weekends, weeks or months at a time), considerable resources (human and financial – to support people to contribute through the process), and the skills to make complex issues transparent and understandable.

This is particularly important given that questions about how data is used and governed can be deeply societally relevant, but also opaque for non-specialists unless considerable time and effort is made to address questions around how uses of data-driven systems interact with the complex policy contexts in which they take place. For instance, in a deliberative process seeking advice on the parameters of health-data sharing, beneficiaries will need to be informed about the use of data, the proposal for a particular data governance mechanism, and also about the implications of its use for health-data outcomes and the impact on patients.⁴⁹

48 Solomon, S., & Abelson, J. (2012). 'Why and when should we use public deliberation?' *The Hastings Center report*, 42(2), p.17-20. Available at: <https://doi.org/10.1002/hast.27>

49 Understanding Patient Data and Ada Lovelace Institute. (2020). *Foundations of Fairness: Where Next For NHS Health Data Partnerships?* Available at: <https://understandingpatientdata.org.uk/sites/default/files/2020-03/Foundations%20of%20Fairness%20-%20Summary%20and%20Analysis.pdf>

There has been a proliferation of one-off initiatives in this space, which include:

- Understanding Patient Data's citizen juries on commercial health data partnerships, in partnership with the Ada Lovelace Institute⁵⁰
- US-based TheGovLab's Data Assembly⁵¹
- the Royal Society of Arts (RSA) citizen juries in partnership with DeepMind (an AI company owned by Google) on AI explainability⁵²
- the Information Commissioners' Office and Alan Turing Institute's Explainable AI citizen juries⁵³
- the Ada Lovelace Institute's Citizens' Biometrics Council, with input from stakeholders such as the Information Commissioner's Office and the Biometrics Commissioner⁵⁴
- a range of UK Research and Innovation (UKRI)-funded initiatives under the auspices of the long-established public-dialogue programme Sciencewise, including a public dialogue on location data with the Ada Lovelace Institute.⁵⁵

In addition to and beyond the one-off initiatives, calls have grown for the 'institutionalisation' of participatory and deliberative approaches to data and AI governance. A recent OECD report, *Catching the Deliberative Wave* defines 'institutionalisation' in two ways. The first is 'incorporating deliberative activities into the rules of public decision-making structures and governance arrangements in a way that is legally constituted so as to establish a basic legal or regulatory framework to ensure continuity regardless of political change'. The second is 'regular and repeated processes that are maintained and sanctioned by social norms, which are important for ensuring that new institutions are aligned with societal values'.⁵⁶

50 Understanding Patient Data and Ada Lovelace Institute. (2020). Foundations of Fairness: <https://understandingpatientdata.org.uk/sites/default/files/2020-03/Foundations%20of%20Fairness%20-%20Summary%20and%20Analysis.pdf>

51 TheGovLab. (2020). *The Data Assembly*. Available at: <https://thedataassembly.org/> [Accessed 5 Jul. 2021].

52 Balaram, B. (2017). 'The role of citizens in developing ethical AI'. *The RSA*. Available at: <https://www.thersa.org/blog/2017/10/the-role-of-citizens-in-developing-ethical-ai> [Accessed 5 Jul. 2021].

53 ICO (2019). *Project explAI In Interim report*. Available at: <https://ico.org.uk/media/2615039/project-explain-20190603.pdf> [Accessed 5 Jul. 2021].

54 The Ada Lovelace Institute. (2021). *The Citizens' Biometrics Council*. Available at: <https://www.adalovelaceinstitute.org/project/citizens-biometrics-council/> [Accessed 5 Jul. 2021].

55 Peppin, A. (2021). 'How charting public perspectives can show the way to unlocking the benefits of location data.' *Ada Lovelace Institute*. Available at: <https://www.adalovelaceinstitute.org/blog/public-perspectives-unlocking-benefits-location-data/> [Accessed 5 Jul. 2021].

56 OECD. (2020). *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave*. Available at: [innovative-citizen-participation-new-democratic-institutions-catching-the-deliberative-wave-highlights.pdf](https://www.oecd.org/publications/innovative-citizen-participation-new-democratic-institutions-catching-the-deliberative-wave-highlights.pdf)

Media commentators on data and AI have proposed that a ‘council of citizens’ should form part of the basis of the institutional decision-making structure that should be enabled to help regulate algorithmic decision-making.⁵⁷ While these kinds of approaches may be relatively novel in the field of AI and data governance specifically, they have precedents in the assessment of emerging technologies more broadly.

Case study: Institutionalised public deliberation, the Danish Board of Technology, 1986–2011⁵⁸

Overview:

One of the highest-profile forms of institutionalised participatory technology assessments was the deliberative consensus conference model implemented by the Danish Board of Technology, which has now been disbanded. The Danish Board of Technology had a statutory duty to inform citizens and politicians on new technology implications, received an annual subsidy and delivered an annual report to the Danish Parliamentary Committee on Science and Technology. It was abolished by law in 2011, and its successor is the Danish Board of Technology Foundation.

The model combined the knowledge of technology experts with the perspectives of non-specialists. The experts helped inform citizen-led reports that summarised the citizens’ agreement or disagreement on questions of how a technology should be developed, their potential risks, future potential applications and appropriate mechanisms through which the effects of the technology on society might be measured. They then shared their consensus reports with the Danish Parliament and the media, which resulted in wide debate and reporting on the findings.

The Danish Parliament received reports directly from consensus conferences on topics as contentious as food irradiation and genetic modification. Deliberators in both instances proposed that government funding should not be spent on those technologies. These recommendations were subsequently endorsed by the Danish Parliament.⁵⁹

57 Carugati, F. (2020). ‘A Council of Citizens Should Regulate Algorithms’. *Wired*. Available at: <https://www.wired.com/story/opinion-a-council-of-citizens-should-regulate-algorithms/> [Accessed 15 Feb 2021].

58 Jørgensen, M.S. (2020). ‘A pioneer in trouble: Danish Board of Technology are facing problems’. *EASST*. Available at: <https://easst.net/easst-review/easst-review-volume-311-march-2012/a-pioneer-in-trouble-danish-board-of-technology-are-facing-problems/> [Accessed 15 Feb 2021].

59 Grundahl, J. (1995). ‘The Danish consensus conference model.’ In Joss, S.& Durant, J. (eds) *Public Participation in Science: The Role of Consensus Conferences in Europe*. Science Museum: London. Available at: <https://people.ucalgary.ca/~pubconf/Education/grundahl.htm>

Participatory implications:

The consensus conference as a method of technology assessment was introduced in many western European countries, the USA and Japan during the 1990s. This assured a route for participatory governance of technologies in an advisory capacity by non-specialist perspectives, while simultaneously ensuring that these perspectives reached power-holders, including parliamentarians.

While the Danish Board of Technology model no longer operates with a statutory footing, due to reductions in public funding, it has a legacy through the establishment of a non-profit foundation in shaping policy on emerging technology. As a case study, it highlights the potential policy and cultural institutionalisation of participatory advice and input in influencing policy decisions about data governance.

Complexities and critiques

Critics of public deliberation acknowledge the cost, resource and time involved, and these can make it difficult to embed public perspectives in a fast-moving policy, regulatory and development field such as technology.

Other critiques acknowledge that requirements that expect citizens to deliberate over some time may impose structural constraints on who is empowered or able to participate. This is a critique identified in a UK Government report on the use of public engagement for technological innovation, for instance.⁶⁰

Not all data governance questions and issues will demand this level of engagement – on questions where there is a high level of established public support and mandate, meaningful transparency and consultation-based approaches may be adequate.

There are organisations working on technology issues that have sought to adapt the traditional deliberation model to incorporate the potential for a more rapid, iterative and online approach to deliberation, as well as to complement it with approaches to engagement that consider

⁶⁰ GOV.UK, (2021). *The Use of Public Engagement for Technological Innovation*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/955880/use-of-public-engagement-for-technological-innovation.pdf.

Embedding public deliberation thoughtfully and effectively requires time and resourcing, and this can be in tension with the imperative to work rapidly in developing and designing data-driven systems

and engage with the lived experiences of underrepresented groups and communities. Examples include the rapid, online Sounding Board model prototyped in 2017 by UK Government-funded programme Sciencewise,⁶¹ with a view to informing rapidly moving policy cycles, and a recent prototype undertaken during the COVID-19 pandemic crisis by the Ada Lovelace Institute in partnership with engagement agencies Traverse, Involve and Bang The Table.⁶²

The Royal Society of Arts (RSA)'s landmark work on building a public culture of economics is an example of how deliberative processes have been complemented with community engagement and broader outreach work with underrepresented communities.⁶³ This initiative complemented a deliberative Citizens' Economic Council with an economic inclusion community outreach programme ('On the Road') that worked in locations identified as high on the index of multiple deprivation.⁶⁴

Embedding public deliberation thoughtfully and effectively requires time and resourcing (often over months rather than weeks), and this can be in tension with the imperative to work rapidly in developing and designing data-driven systems, or in developing policy and regulation to govern the use of these systems. An example of when urgent decision-making demands a more iterative and agile approach to assembling data infrastructures is the necessarily rapid response to the COVID-19 pandemic. In these instances, deliberative exercises may not always appear to be expedient or proportionate – but might on balance be valuable and save energy and effort if a particular use of data is likely to generate significant societal concern and disquiet.

61 Patel, R. (2015). 'Digital Public Engagement – Lessons from the Sounding Board'. *Involve*. Available at: <https://www.involve.org.uk/resources/blog/opinion/digital-public-engagement-lessons-sounding-board> [Accessed 5 Jul. 2021].

62 Ada Lovelace Institute. (2020). *Confidence in a crisis? Building public trust in a contact tracing app*. Available at: <https://www.adalovelaceinstitute.org/report/confidence-in-crisis-building-public-trust-contact-tracing-app/> [Accessed 5 Jul. 2021].

63 The RSA. (2017). *The Citizens' Economic Council 'On The Road'*. Available at: <https://www.thersa.org/projects/archive/economy/citizens-economic-council/the-council2>

64 The RSA. (2017).

While deliberation can be exclusively advisory, when embedded in a data access, sharing or governance framework, it can also have potential to navigate tensions and trade-offs.

Collaborate

Stewards and beneficiaries collaborate in governing data

The process of collaboration in the context of public involvement in data and AI can be understood as enabling people to negotiate and engage in trade-offs with powerholders and those governing data about specific aspects of decision-making.

1. Deliberation embedded across data access, sharing and governance frameworks

While deliberation can be exclusively advisory, when embedded in a data access, sharing or governance framework, it can also have potential to navigate tensions and trade-offs. There has been increased interest in the question of whether (and to what extent) it is feasible to *embed* deliberation as part of data-sharing, databank, or data-trust models, through the use of 'bottom-up data trusts' and 'civic data trusts' for instance.⁶⁵

There are two aspects to this potential. Firstly, deliberative approaches focus on understanding societal benefits and harms, exhibiting potential for enabling participation of data beneficiaries as collectives, rather than as individuals (who are rarely in a position to negotiate or engage in trade-offs as they relate to their data rights).

These approaches have the potential to enable collective consent around the particular uses of data governance, including engaging those most likely to be directly impacted – in turn helping to address a central challenge associated with relying exclusively on individual consent. Consent-based models have been critiqued for failing to encourage citizens to consider the benefits of data use for wider society, as well as being designed in ways that can feel coercive or manipulative – rather than genuinely seeking informed consent.

65 Delacroix, S., and Lawrence, N. D. (2019). 'Bottom-up Data Trusts: disturbing the 'one size fits all' approach to data governance', *International Data Privacy Law*, Volume 9, Issue 4, pp. 236–252. Available at: <https://doi.org/10.1093/idpl/ipy014>

Secondly, unlike other mechanisms, deliberative approaches have the potential to be embedded at different points of the data lifecycle – shaping collection, access, use and analysis of data.

The Ada Lovelace Institute has not identified any pilots that have been successful, to date, in embedding deliberative approaches specifically in the governance of data-sharing partnerships. However, there is potential for developing future pilots drawing on and iterating from models in related areas, as the following case study – a pilot project initiated by the Wellcome Trust – demonstrates.

Case study: Global mental health databank, the Wellcome Trust with Sage Bionetworks, 2020⁶⁶

Overview:

The Wellcome Trust has funded a pilot project with young people across India, South Africa and the United Kingdom with a view to co-designing, building and testing a mental-health databank that holds information about mental and physical health among people from 17 to 24 years old. The goal is to create a databank that can be used in the future to help improve the ways mental health problems like anxiety and depression for young people living in high and low-income countries around the world are treated and prevented. The Wellcome Trust expects the databank to help researchers, scientists and young people answer the questions, 'What kinds of treatments and prevention activities for anxiety and depression really work, who do they work for, and why do they work among young people across different settings?'

Participatory implications:

This initiative aims to test participatory approaches to involving young people, including through the use of lived-experience panels and citizen deliberation, working directly with lived-experience advisers and young people experiencing mental health conditions through the process of creating the databank.

66 Taylor, K. (2020). 'Developing a mental health databank'. *Medium*. Available at: <https://medium.com/wellcome-digital/developing-a-mental-health-databank-99c25a96001d> [Accessed 15 Feb. 2021].

The Wellcome Trust's mental-health databank pilot signals the likely direction of travel for many participatory approaches to data governance. Although it remains to be seen what its effect or impact will be, its ambitions to embed deliberative democracy approaches into its design, and involve young people experiencing mental health conditions in the process, are likely to be replicable.⁶⁷

Complexities and critiques

There are currently few examples of participatory approaches to govern access and use of data, and therefore few studies of the effectiveness of these methods in practice. Early insights suggest that this approach may not be appropriate for all data-sharing initiatives, but may be particularly appropriate for those that require a thoughtful, measured and collaborative approach, usually involving the collection and use of data that is highly sensitive (as with the Wellcome Trust initiative, which relates to mental-health data).

A recent report from public participation thinktank Involve finds that there are three potential stages at which deliberative participation could be feasible in the design of data-sharing or data-access initiatives (what we describe in this report as a data-governance framework), which warrant further experimentation, testing and evaluation. These are:⁶⁸

1. **Scoping** – aligning purpose and values of the data-sharing or data-access initiative, prior to its establishment.
2. **Co-design** – developing criteria, framework or principles to ensure its decision-making process meets the needs and expectations of wider stakeholders and the public, and developing policy on the distribution of value (i.e who benefits from the use of the data, and whether there is public or wider societal value that comes from it).
3. **Evaluation** – deploying participatory approaches to ensure that the intended impact of the access initiative has been met, that outcomes and potential have been maximised, and to ensure adequate accountability and scrutiny around the claims for data access and use. This reflection on how the initiative has worked can be continual as well as retrospective.

67 Sage Bionetworks. (2020). *Collaborating with youth is key to studying mental-health management*. Available at: https://www.eurekalert.org/pub_releases/2020-11/sb-cwy111720.php [Accessed 5 Jul. 2021].

68 Lansdell, S. and Bunting, M. (2019). *Designing decision making processes for data trusts: lessons from three pilots*. The Involve Foundation (Involve). Available at: <http://theodi.org/wp-content/uploads/2019/04/General-decision-making-report-Apr-19.pdf>

In data donation, an individual actively chooses to contribute their data for the purposes of wider societal gains

2. Data-donation and data-altruism models

Another example of a collaborative approach to data governance is the model of ‘data donation’, which has at its basis the premise that those who *contribute or donate their data* will expect to see certain terms and conditions realised around the process of donation. Usually, a term or condition will revolve around a clear articulation of ‘public benefit’ or wider societal gain.

In contrast to more ‘extractivist’ approaches to data governance, where approaches to gathering and mining data about people can take place surreptitiously, data-donation mechanisms are a route through which individuals can explicitly agree to share their data for wider societal and collective benefits under clear terms and conditions.⁶⁹

In data donation, an individual actively chooses to contribute their data for the purposes of wider societal gains, such as medical research. It is a non-transactional process, so donors do not expect an immediate direct benefit back as a consequence of donating. If data donors and volunteers are not satisfied that their expectations are being met, for example if an AI or machine-learning process failed to meet donor expectations, they have the right to ‘withdraw’ or delete their data, or otherwise to be able to take action, and many contributors may choose to do so during the lifespan of their participation in the process.

Data-donation initiatives can be part of a research study, where data subjects voluntarily contribute their own personal data for a specific purpose. Another model, which is increasingly gaining popularity, is that data subjects might opt to share data that has already been generated for a different purpose. For example, people who use wearable devices to track their own activities may choose to share their data with a third party).

Emerging evidence suggests that encouraging people to share data for pro-societal purposes is both a strong motivator and a key basis for public confidence in the effectiveness of the approach. A recent University of Bristol research study into the psychology of data donation found that the strongest predictor of the decision to donate data was to

69 Bietz, M., Patrick, K. and Bloss, C. (2019). ‘Data Donation as a Model for Citizen Science Health Research’. *Citizen Science: Theory and Practice*, 4(1), p.6. Available at: <http://doi.org/10.5334/cstp.178>

serve society, and that knowing the consequences and potential benefits of donating data was a critical factor influencing people's decisions to participate. Here, data governance becomes a collaborative endeavour, where the legitimisation, consent and pro-societal motives of data donors become central to both the viability and the effectiveness of the approach.

A good example of this type of initiative is UK Biobank, which is a large-scale biomedical database that provides accredited researchers and industry organisations with access to medical and genetic data generated by 500,000 volunteer participants specifically for the purpose of effecting good health outcomes.

Case study: Large-scale data donation for research, UK Biobank, 2006–present⁷⁰

Overview:

UK Biobank has blood, urine and saliva samples from 500,000 volunteers whose health has been tracked over the past decade, as part of a large-scale 'data donation' research initiative. This has enabled it to gather longitudinal data about a large sample size of the population, helping to answer questions about how diseases such as cancer, stroke and dementia develop. It has also formed the basis for intervening in response to COVID-19, with data about positive coronavirus and GP/hospital appointments added, and 20,000 contributors sharing blood samples for pandemic response purposes.⁷¹ A range of third-party organisations including those in academia and industry can apply for different layered levels of access, paying a subscription fee to be able to access this data.

Participatory implications:

UK Biobank is a data-donation model that has monitoring of data access and use in place through internal and external auditing mechanisms. It is a national resource for health research, with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses. It operates on the basis that contributors 'opt in' to sharing their data for research purposes.

70 UK BioBank. (2021). *Explore your participation in UK Biobank*. Available at: <https://www.ukbiobank.ac.uk/explore-your-participation> [Accessed 15 Feb. 2021].

71 UK BioBank. (2021). *Explore your participation in UK Biobank*. Available at: <https://www.ukbiobank.ac.uk/explore-your-participation> [Accessed 15 Feb. 2021].

Data donors are free to opt out and stop sharing their data at any point without needing to provide a reason. If someone opts out, UK Biobank will no longer contact the participant or obtain further information, and any information and samples collected previously would no longer be available to researchers. They would also destroy samples (although it may not be possible to trace all distributed sample remnants) and continue to hold information only for archival audit purposes.

Complexities and critiques

While data donation has the potential to support longitudinal, evidence-based research, a major critique and challenge of this mechanism has been the extent to which self-selection can result in the underrepresentation of marginalised groups and communities. Research demonstrates that there is a consistent trend of underrepresentation of minority populations in biobanks, which undermines their value.⁷²

A number of dynamics contribute towards this, including the lack of financial support for 'data donors', or assumptions that, for instance, all data donors will have the necessary resources to contribute, or have a confirmed place of residence. Further structural inequalities compound underrepresentation, for instance, differential levels of trust in the effectiveness of data-driven systems, uses of data and in researchers.

Data from self-selecting biobank models, while useful, can therefore be at risk of perpetuating unequal outcomes when it comes to the use of social policy mechanisms, or excluding underrepresented groups and communities from data-driven health policy. This persistent phenomenon of exclusion of underrepresented groups and communities from datasets is often described as the 'missing data' problem. This missing data can entrench and perpetuate inadvertent bias and discrimination by failing to identify differential impacts.⁷³

72 Kim P, Milliken EL. (2019). 'Minority Participation in Biobanks: An Essential Key to Progress: Methods and Protocols'. *Methods in Molecular Biology*. 1897: pp.43-50. Available at: https://www.researchgate.net/publication/329590514_Minority_Participation_in_Biobanks_An_Essential_Key_to_Progress_Methods_and_Protocols

73 Groenwold, R., & Dekkers, O. (2020). 'Missing data: the impact of what is not there'. *European Journal of Endocrinology*, 183(4), E7-E9. Available at: <https://eje.bioscientifica.com/view/journals/eje/183/4/EJE-20-0732.xml>.

Other challenges include a lack of clarity to data donors about exactly how their data has been used, and how to ‘opt out’ of data donation. Increasingly, data donation’s incentive structures warrant critical scrutiny – there is always a risk of creating coercive or perverse incentive structures in corporate environments (for instance, private-sector providers such as insurers requesting ‘donations’ of wearable data in exchange for lower premiums or annual fees).

Another potential issue is the lack of clarity about the exact terms on which the data is ‘donated’, for instance, some people feeling that they are *expected* to share their data to access a healthcare service, when in reality there is no such expectation, or limited awareness about rights to opt out of data sharing. This is a particular challenge with ‘opt out’, rather than ‘opt in’ models of data donation (presently widespread practice in the UK National Health Service), as recent societal discussion and debate about the UK’s new centralised data-sharing infrastructure, and opt-out mechanisms under the General Practice Data for Planning and Research (GPDPR) proposals have highlighted.⁷⁴

Empower

Beneficiaries actively make decisions about data governance

Empowering data beneficiaries enables them to exercise full managerial power and agency, and take responsibility for exercising and actively managing decisions about data governance – specifically, how data is accessed, shared, governed and used. In this model, the dynamic of power is shifted away from the data steward towards the data beneficiary who makes the decision, advised where necessary by appropriate specialist expertise.

⁷⁴ Machirori, M. and Patel, R. (2021). ‘Turning distrust in data sharing into “engage, deliberate, decide”’: *Ada Lovelace Institute*. Available at: <https://www.adalovelaceinstitute.org/blog/distrust-data-sharing-engage-deliberate-decide/> [Accessed 5 Jul. 2021].

Examples of these approaches are relatively rare, but they do exist, are increasingly emerging and include the following:

- shared control and ownership of data (through, for instance, data cooperatives)⁷⁵
- electoral mechanisms for beneficiary involvement (such as voting on boards)
- setting terms and conditions for licensing and data access
- shaping the rules of the data-governance framework.

The following case study of the Salus Cooperative (known as Salus Coop), based in Spain, illustrates how beneficiaries have been enabled to make decisions actively about the governance of their, and others', data – through corporate governance, but also through processes such as setting license terms.

Case study: Citizen-driven, collaborative data management and governance, Salus Coop (Spain), 2017–present⁷⁶

Overview:

Salus Coop is a non-profit data cooperative for health research (meaning in this context not only health data, but also lifestyle-related data that has health indicators, such as number of steps taken), founded in Barcelona by members of the public in September 2017. It set out to create a citizen-driven model of collaborative governance and management of health data 'to legitimize citizens' rights to control their own health records while facilitating data sharing to accelerate research innovation in healthcare'.

Salus meets the definition of a data cooperative, as it provides clear and specified benefits for its members – specifically a set of powers, rights and constraints over the use of their personal health data – in a way that also benefits the wider community by providing data for health research. Some of these powers and rights would be provided by enforcement of the GDPR, but Salus is committed to providing them to its members in a transparent and usable fashion.

⁷⁵ For more on data cooperatives see chapter 2 in *Exploring legal mechanisms for data stewardship* (Ada Lovelace Institute, 2021)

⁷⁶ Data Collaboratives. (n.d.). *Salus Coop*. Available at: <https://datacollaboratives.org/cases/salus-coop.html> [Accessed 15 Feb. 2021].

Participatory implications:

Together with citizens Salus has developed a 'common good data license for health research' through a crowd-design mechanism, which it describes as the first health data-sharing license in the world. The Salus common-good license applies to data that members donate and specifies the conditions that any research projects seeking to use the member data must adhere to. The conditions are:

- **Health only:** The data will only be used for health-related (i.e. treatment of chronic disease) research.
- **Non-commercial:** Research projects will be promoted by entities of general interest such as public institutions, universities and foundations only.
- **Shared results:** All research results will be accessible at no cost.
- **Maximum privacy:** All data will be anonymised and unidentified before any use.
- **Total control:** Data donors can cancel or change the conditions of access to their data at any time.

Salus describes itself as committed to supporting data donors' rights and ensuring they are upheld, and requiring researchers interacting with the data to ensure that:

- individuals have the right to know under what conditions the data they've contributed will be used, for what uses, by which institutions, for how long and with what levels of anonymisation
- individuals have the right to obtain the results of studies carried out with the use of data they've contributed openly and at no cost
- any technological architecture used allows individuals to know about and manage any data they contribute.

Critiques and complexities

Models such as data cooperatives can provide security for beneficiaries and data subjects that supports data sharing for beneficial purposes. However, because they exclusively expect or engender greater levels of active participation in managing and shaping a data-sharing regime or process, they are at risk of excluding those who may wish to actively participate but find the costs onerous. For example, potential data donors may feel they lack the time, levels of knowledge about the regulatory landscape, or financial and social capital.

This means that these approaches are rarely appropriate for all beneficiaries and data subjects, and so cannot claim to be fully inclusive or representative. Data cooperatives can, therefore, struggle to generate the scale and level of participation from data subjects that they might hope for, but they can nevertheless help broaden out the range of intelligence informing data governance.

Another critique relates to the financial sustainability of these approaches and models. There is limited financing available that would absorb the project and start-up costs associated with data cooperative models, especially where they need to meet regulatory requirements and constraints. This can present a financial and a regulatory burden that is a barrier to setting up a data cooperative. In contrast to shareholder-controlled companies, cooperatives cannot give equity to investors, as they are owned by, and give return on investment to, their respective members. Therefore, cooperatives (and governance models similar to cooperatives) require significant, external financial support from governments, foundations and research grants if they are to succeed.

When designing a participatory process to meet a specific objective, the choice about who to involve matters as much as which types of involvement or participation mechanisms are used

Who to involve when developing participatory mechanisms

What all these different methods and approaches have in common is that they seek to involve beneficiaries in the use, development and design of data and AI, and that they involve some element of sharing power with those beneficiaries in the way that data-driven systems and governance frameworks themselves are designed.

When designing a participatory process to meet a specific objective, the choice about who to involve matters as much as *which* types of involvement or participation mechanisms are used. These choices will be dependent on context, but – as outlined on [p. 48](#) – the range of actors who can be defined as beneficiaries is broad, and extends beyond those designing and deploying data-driven systems and governance frameworks, to those affected by them.

When participatory mechanisms are introduced, key questions developers of data-driven systems and governance frameworks should answer about **who** to involve (who their beneficiaries are) will be:

1. **Who has a stake in the outcomes** that emerge?
2. Who is most likely to be **directly affected and impacted**, either benefiting or being adversely impacted?
3. Who is most likely to be **overrepresented and/or underrepresented** in the data?

These three key questions are informed by a recognition that the data stewards' responsibility is not just to manage data itself effectively, but also to recognise that data often relates, either directly, or indirectly to people (beneficiaries). As well as recognising the rights of data subjects, and the potential benefits and harms of data use to beneficiaries, data stewards need to understand that when data *omits* or excludes people, it has the potential to have harmful consequences. This can happen, for instance, by discriminating against or underrepresenting some people's interests and concerns. This means that participation can be as much about including or involving those who *do not* have a direct relationship with the data as assembled, as those who do.

Benefits of effective participation for the design, development and use of data-driven systems

Early and continuous engagement with beneficiaries, and those most likely to be affected by the use and deployment of data-driven systems can help inform decisions about the design of those systems. These decisions will create better outcomes for those designing, developing and deploying data-driven systems and governance frameworks, as well as for people and society.⁷⁷

Beneficial outcomes for designers, developers and deployers of data-driven systems and governance frameworks

Because Participatory approaches encourage interactions with a range of views, perspectives, norms and lived experiences that might not otherwise be encountered in the development and design of data-driven systems and governance frameworks. This minimises the risks of groupthink, unconscious biases and misalignments between intended and actual outcomes. Benefits of effective participation for designers and developers include:⁷⁸

- **Better understanding of ethical and practical concerns and issues:** Enabling developers and designers to better understand ethical concerns and challenges from the public, and better understand public perspectives, values, trade-offs and choices.⁷⁹ Participatory data stewardship can also inform and affect the *quality* of data embedded within a system.

77 ESRC website. (n.d). *Why public engagement is important*. Available at: <https://esrc.ukri.org/public-engagement/public-engagement-guidance/why-public-engagement-is-important/> [Accessed 15 February 2021]

78 There are various definitions of values-led and instrumental rationales for public engagement including: Stirling, A. (2012). 'Opening Up the Politics of Knowledge and Power in Bioscience'. *PLoS Biology* 10(1). Available at: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001233>

79 The Involve Foundation ('Involve'), (2019). *How to stimulate effective public engagement on the ethics of artificial intelligence*. Available at: <https://www.involve.org.uk/sites/default/files/field/attachemnt/How%20to%20stimulate%20effective%20public%20debate%20on%20the%20ethics%20of%20artificial%20intelligence%20.pdf> [Accessed 15 February 2021].

- **More considered design of systems and frameworks informed by diverse thinking:** Improved decision-making about development and design that reflects and has taken account of a diversity of experiences and perspectives.⁸⁰
- **The anticipation and management of risk in the development and design of systems or frameworks:** The ability to manage risk in development and design, particularly those systems or frameworks that are complex and controversial because of sensitive data, circumventing and addressing the risk of ‘techlash’. Participation can also reduce the long-term costs for technology developers and designers.⁸¹
- **Higher-quality data-governance frameworks:** Across corporate data governance, there is often a lack of documentation or knowledge about the context in which a particular dataset was collected, for example, what levels of consent it has or what were the data subject’s intended uses. When data generation is opaque and invisible to the data subject, its legitimacy as a data source is frequently either assumed through terms and conditions, or ignored entirely. This can lead to downstream violations of contextual integrity of that data. Embedding participatory data stewardship involves shifting institutional incentives in corporate practice, to prioritise improved data quality over data quantity (a ‘less is more’ ethos), with the benefits of clearer, higher quality and fit-for-purpose datasets.

Beneficial outcomes for trustworthy data governance

Participatory approaches to data governance also play a central role in shaping technology’s legitimacy, where legitimacy is defined as the ‘reservoir of support that allows governments to deliver positive outcomes for people.’⁸² In the case of data, the principle of legitimacy extends beyond public bodies such as governments and regulators, to those designing, developing and implementing data-governance frameworks – this is developers and deployers’ ‘social license to build’. All these public bodies, companies and individuals can use data to deliver positive outcomes for people, but only with public support.

80 Page, S. (2011). *Diversity and complexity*. Princeton, N.J: Princeton University Press.

81 Clarke, R. (2015). *Valuing Dialogue: Economic Benefits and Social Impacts*. London: Sciencewise Expert Resource Centre. Available at: <https://sciencewise.org.uk/wp-content/uploads/2018/12/Valuing-dialogue-2015.pdf> [Accessed 15 February 2021].

82 Centre for Public Impact. (2018). *Finding Legitimacy*. Available at: <https://www.centreforpublicimpact.org/assets/documents/Finding-a-more-Human-Government.pdf> [Accessed 15 February 2021].

Participatory approaches to data governance can help engender some societally beneficial outcomes

We redefine legitimacy in the context of data, therefore, as the **broad base of public support that allows companies, designers, public servants and others to use data, and to design and develop data-governance frameworks**. The absence of this legitimacy has been illustrated in a series of public, controversial events that have taken place since 2018. Examples include the Cambridge Analytica/Facebook scandal, in which political microtargeting was misused,⁸³ or the scandal that emerged when the NHS shared sensitive patient data with private research company DeepMind, with limited oversight or constraints.⁸⁴ More recently, governments' efforts to implement and deploy digital contact tracing across the world have also met with considerable public debate and scrutiny.⁸⁵

Participatory approaches to data governance can help engender some societally beneficial outcomes, but we do not propose that they replace legal and rights-based approaches, such as those embedded in the General Data Protection Regulation (GDPR) or across broader data protection and governance frameworks, which are analysed in detail in our companion report *Exploring legal mechanisms for data stewardship*.⁸⁶ Rather, they work hand in hand with rights-based and legal approaches to ensure public confidence and trust in appropriate uses of data. Some participatory approaches will and can, help shape the *future of rights-based approaches, governance and regulation*.

83 Fruchter, N., Yuan, B. and Specter, M. (2018). *Facebook/Cambridge Analytica: Privacy lessons and a way forward*. Internet Policy Research Initiative at MIT. Available at: <https://internetpolicy.mit.edu/blog-2018-fb-cambridgeanalytica/>.

84 ICO. (2018). *Royal Free - Google DeepMind trial failed to comply with data protection law*. Available at: <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2017/07/royal-free-google-deepmind-trial-failed-to-comply-with-data-protection-law/>.

85 Ada Lovelace Institute. (2020). *Exit through the App Store? Rapid evidence review*. Available at: <https://www.adalovelaceinstitute.org/evidence-review/covid-19-rapid-evidence-review-exit-through-the-app-store/> [Accessed 5 Jul. 2021].

86 Ada Lovelace Institute. (2021). *Exploring legal mechanisms for data stewardship*. Available at: <https://www.adalovelaceinstitute.org/report/legal-mechanisms-data-stewardship/> [Accessed 5 Jul. 2021].

Outcomes from different approaches to participatory data governance can include:

- enabling people and society to **better understand and appreciate the wider public benefit** that comes from ‘donating their data’ in a given context
- enabling designers of data-governance framework to **understand the boundaries of data use and management: what people feel is acceptable and unacceptable** for the access, use and sharing of their data in a given context
- enabling the use of **collective intelligence shared by data stewards and beneficiaries**, to improve the quality of data protection and data governance when gathering and using personal or sensitive data, or data about marginalised groups
- **strengthening accountability**, by opening opaque data uses up to democratic, civic scrutiny and constructive criticism (particularly in contexts where the use of data is likely to have a significant impact on individuals or diverse groups, or where there is a risk of bias and self-selection in the use of datasets)
- building **public confidence in the way data is gathered and used, particularly where third parties access data**, by ensuring people are able to oversee and hold to account decisions about rights to access, share and control that data
- **tightening the feedback loops that exist** between those who are stewards and fiduciaries of data, and those to whom the data relates (either directly, or indirectly).

These are all potentially valid *outcomes* that emerge from embedding participatory data stewardship mechanisms. The mechanisms to enable these outcomes can be different – if they are to be impactful, they are quite likely to vary depending on context, use of the data, the type of data being governed, and those who are most likely now and in future to be impacted by the data use and governance.

Conclusion

This report highlights the range and plurality of participatory data stewardship mechanisms through which people and society are able to influence and shape data governance processes, and the potential benefits for beneficiaries, public bodies and those designing and deploying data-driven systems and governance frameworks.

By examining real-world case studies, it has set out some of the ways in which those managing, controlling and governing the use of data can avoid ‘going through the empty ritual of participation’ described by Arnstein – and shift away from practices that are harmful, disempowering or coercive of people, towards practices that promote a collaborative and co-creative approach to working together.

The framework for participation on p. 14 demonstrates the multiple and overlapping goals participation can have, in the context of the design, development and deployment of data-driven systems and data-governance frameworks. The range of different mechanisms that serve different goals and purposes can complement each other and are mutually reinforcing.

Effective participation must start with the goal of informing people about their data – with *transparency* (the right to be informed and the right to take meaningful action based on what one knows), where the data steward shares information about what is happening to people’s data in ways that enable them to take action.

For participation to become meaningful, it must extend beyond transparency towards mechanisms that aim to understand, interact with, be advised by and then respond to people’s views (*consult and involve*) in decision making about data– shifting power dynamics in that process.

The end goal of participation is to realise a set of conditions where people have the potential to be *empowered*, so that their perspectives form part of the design of data governance, are responded to, and in turn build confidence and capacity for beneficiaries to continue to participate in the data governance process.

As this report makes apparent, there is considerable creative potential for developers, designers, beneficiaries and users of data-driven systems and data-governance frameworks to work together in shaping improved mechanisms for stewardship. The aim is to create a *virtuous cycle*, where participation effects substantial change in practice, and public confidence and trust generated by participatory approaches effect improved outcomes in the use of data for people and society.

Methodology

This report and framework has been developed and informed by extensive mixed-methods qualitative and case study-based research that encompasses:

- a case study-based analysis of over one hundred examples of participatory data sharing⁸⁷
- a literature review of data metaphors and narratives
- a desk-based review and synthesis of grey and academic literature on existing models of public engagement in technology and data
- the deliberations of the Ada Lovelace Institute and Royal Society's Legal Mechanisms for Data Stewardship working group.⁸⁸

87 Patel, R., Gibbon, K. and Peppin, A. (2020). *Exploring principles for data stewardship*. [online] Ada Lovelace Institute. Available at: <https://www.adalovelaceinstitute.org/project/exploring-principles-for-data-stewardship/> [Accessed 16 Feb. 2021].

88 The Ada Lovelace Institute. (2021). *Exploring legal mechanisms for data stewardship*. Available at: <https://www.adalovelaceinstitute.org/report/legal-mechanisms-data-stewardship/>

Appendix

Sherry Arnstein's 'ladder of citizen participation' sets out how institutions might ascend the metaphorical rungs in involving people in critical decisions made about what affects their lives. The lower rungs on the ladder represent nudging or manipulating people to act in a particular way, or merely attempting to tell people about a decision, and the higher rungs grant greater levels of power and involvement in shaping those decisions.

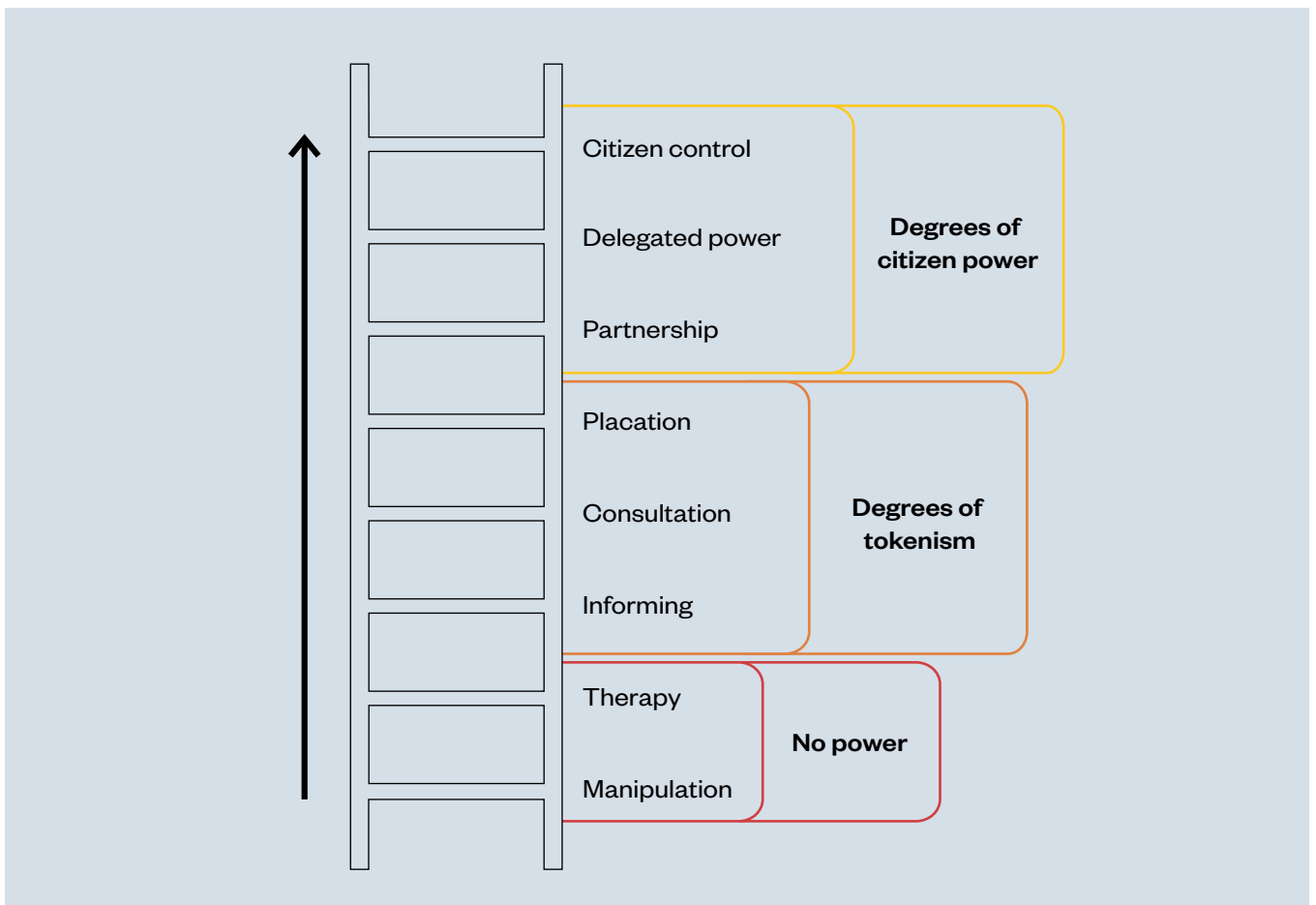


Figure 5: Arnstein's 'Ladder of Citizen Participation'

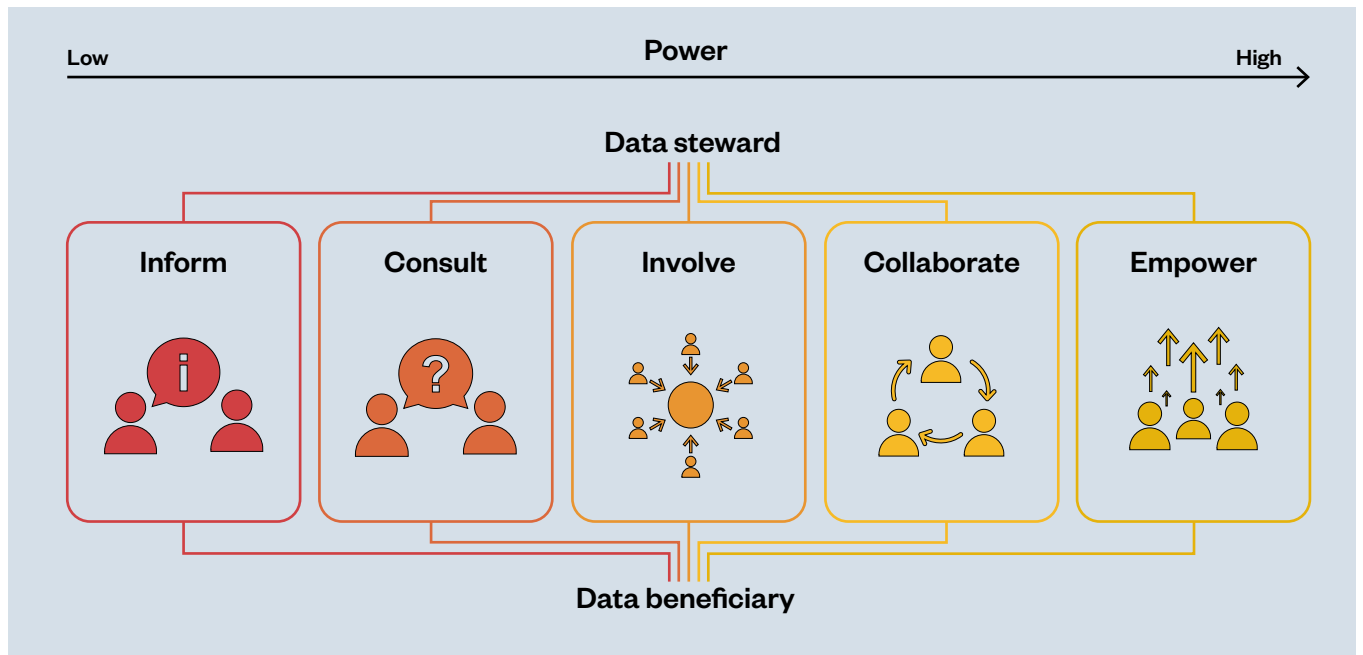


Figure 6: The participation spectrum

Acknowledgements

We would like to thank the following colleagues for taking time to review a draft of this paper, and offering their expertise and feedback:

- Alix Dunn, Computer Says Maybe and Ada Lovelace Institute Board member
- Dr Allison Gardner, Keele University
- Andrew Strait, Ada Lovelace Institute
- Astha Kapoor, The Aapti Institute
- Carly Kind, Ada Lovelace Institute
- Professor Diane Coyle, Bennett Institute for Public Policy, Cambridge University
- Hetan Shah, British Academy and Ada Lovelace Institute Board member
- Jack Hardinges, Open Data Institute
- Dr Jeni Tennison, Open Data Institute
- Kasia Odrozek, Mozilla Foundation
- Katie Taylor, Wellcome Trust
- Professor Lina Dencik, The Cardiff Data Justice Lab
- Lisa Murphy, NHSx
- Dr Mahlet Zimeta, Open Data Institute
- Miranda Marcus, Wellcome Trust
- Dr Miranda Wolpert, Wellcome Trust

We are also grateful to techUK and to the Cardiff Data Justice Lab, who both organised workshops to facilitate input into this research, and for attendance at those workshops by a range of technology companies, policymakers, civil-society organisations and academics.

This report was lead authored by Reema Patel, with substantive contributions from Aidan Peppin, Valentina Pavel, Jenny Brennan, Imogen Parker and Cansu Safak.

About the Ada Lovelace Institute

The Ada Lovelace Institute was established by the Nuffield Foundation in early 2018, in collaboration with the Alan Turing Institute, the Royal Society, the British Academy, the Royal Statistical Society, the Wellcome Trust, Luminata, techUK and the Nuffield Council on Bioethics.

The mission of the Ada Lovelace Institute is to ensure that data and AI work for people and society. We believe that a world where data and AI work for people and society is a world in which the opportunities, benefits and privileges generated by data and AI are justly and equitably distributed and experienced.

We recognise the power asymmetries that exist in ethical and legal debates around the development of data-driven technologies, and will represent people in those conversations. We focus not on the types of technologies we want to build, but on the types of societies we want to build.

Through research, policy and practice, we aim to ensure that the transformative power of data and AI is used and harnessed in ways that maximise social wellbeing and put technology at the service of humanity.

We are funded by the Nuffield Foundation, an independent charitable trust with a mission to advance social well-being. The Foundation funds research that informs social policy, primarily in education, welfare and justice. It also provides opportunities for young people to develop skills and confidence in STEM and research. In addition to the Ada Lovelace Institute, the Foundation is also the founder and co-funder of the Nuffield Council on Bioethics and the Nuffield Family Justice Observatory.

Find out more:

Website: [Adalovlaceinstitute.org](https://adalovlaceinstitute.org)

Twitter: [@AdaLovelaceInst](https://twitter.com/AdaLovelaceInst)

Email: hello@adalovlaceinstitute.org



Permission to share:
This document is published
under a creative commons
licence: CC-BY-4.0

ISBN 978-1-8382567-6-0