

Data trusts: international perspectives on the development of data institutions

WORKING PAPER 2

28 JANUARY 2021



Data
Trusts
Initiative

Data trusts: international perspectives on the development of data institutions

Working Paper 2

Summary

Data acquires value through use; it can create tailored projects and services for individuals, streamline organisational processes, and improve the delivery of public administration. The magnitude of the potential economic and social benefits that follow demand action to create frameworks and mechanisms that enable data sharing for public benefit. At the same time, these shifting patterns of data generation and use create new vulnerabilities for individuals, communities and society.

With the ambition of harnessing the value of data for economic and social benefit, while preventing citizens being exposed to new forms of data-enabled harm, governments across the world have looked to data trusts as a new model of data institution. In creating these data trusts, lessons can be drawn from comparing the approaches and legal models that are being established across different jurisdictions.

A workshop convened by the Data Trusts Initiative on 28 January 2021 set out to explore how discussions about data trusts are progressing across different jurisdictions and the insights that these international perspectives offer for the future development of data institutions.¹

Key findings

A prerequisite for the establishment of data trusts is that individuals have rights over data about them that can be held in trust. Countries and regions of the world differ in the types of right that are available under current legislative frameworks. Building on these local legislative foundations, a range of different approaches to collective data governance challenges have emerged. These initiatives highlight a range of operational issues that data trusts must navigate, including:

- Business models for long-term sustainability;
- Technical architectures for data access or management;
- Methods of participatory governance;
- Perceptions of trustworthiness;
- Distribution of benefits from trust activities.

Across jurisdictions, emerging debates about the development of data trusts point to the importance of building on local frameworks to address new governance challenges; tailoring governance structures to context and purpose; and managing tensions between individual and collective interests.

¹ The Data Trusts Initiative would like to thank all those that participated in the 28 January workshop. This note summarises discussions at the workshop. It is not intended as a verbatim record and does not reflect an agreed position by workshop participants.

1. Background: data institutions and data trusts

The role of data institutions in data governance

Individuals, communities and organisations are generating increasing volumes of data from a greater variety of digitally-mediated daily activities. This data acquires value through use; it can create tailored projects and services for individuals, streamline organisational processes, and improve the delivery of public administration. The magnitude of the potential economic and social benefits that follow demand action to create frameworks and mechanisms that enable data sharing for public benefit.

At the same time, these shifting patterns of data generation and use create new vulnerabilities for individuals, communities and society. Individuals risk the exposure of sensitive information, as seemingly innocuous data is shared and used in new ways, with unanticipated consequences; some communities have experienced algorithmically-enabled forms of discrimination, as the biases that shape society are replicated in the digital realm; and societies are grappling with the ways in which digital-mediation of public life reinforces social divisions.

Across the world, different national and international approaches to data governance have emerged to help tackle these challenges. These existing laws already provide a variety of data rights that seek to prevent harms arising from data misuse, but – in the jurisdictions where such rights do exist – exercising them can demand considerable knowledge, time, and energy. Changing patterns of data use also create pressure points in current governance approaches, as legislative frameworks struggle to cope with a data environment where data can be used and re-used in different – and unanticipated – ways.

Achieving the promise of the digital economy will require robust new data governance institutions that allow data sharing – helping develop new data-enabled products and services – while protecting our rights and freedoms.

‘Data institution’ is an umbrella term; it refers to a suite of organisations and institutional structures that steward data on behalf of individuals, communities and organisations.² The nature of the data stewardship that is required by these institutions depends on the purpose of the data sharing activity that they oversee, the nature of the risks and benefits involved, and the suite of stakeholders potentially affected. Such stewardship might seek to promote individual interests, address shared or societal areas of need, manage the vulnerabilities accruing to different parties as a result of data sharing, or empower individuals or groups to have a voice in how data about them is used.

The form and function of data institutions varies. For example, existing data institutions include:

- National statistical agencies – public bodies or independent agencies – that play a role in collecting and stewarding datasets of importance to public policy, reporting to national governments or parliaments.³
- Biobanks that provide repositories for research data, providing resources for many forms of medical research, typically based on corporate or charity governance structures.⁴

² See, for example: <https://theodi.org/project/sustainable-data-institutions/>

³ Many national statistical agencies are independent of government, and are accountable to regulators or parliaments. For example, The UK’s Office for National Statistics is a non-ministerial department that reports to the UK Parliament.

⁴ For example, UK BioBank is incorporated as a limited company and is registered as a charity, with its board acting as charity trustees.

-
- Data cooperatives that support individuals or organisations to pool their data into a single institution, with that institution then being controlled by its members.⁵

While today's data governance frameworks are able to support a range of data sharing activities, recent years have shown their limitations when seeking to prevent the vulnerabilities that stem from daily digital interactions, or when supporting groups to promote data use for a collective purpose. For example, reliance on consent as the basis for data sharing through contracts presupposes time, resources and the ability to negotiate conditions for data use that most do not have, while assuming that individual consent is sufficient to address the relational aspects of data use. Adding to these limitations, the power asymmetries that characterise today's digital environment make it difficult for individuals acting alone to have their voices heard in debates about data use.

In response, new forms of data institution are emerging. These include new data cooperatives, data unions, and data trusts.

Data trusts: a data institution for the 21st century?

A data trust is a mechanism for individuals to pool the data rights created by current legislation into an organisation – a trust – in which trustees make decisions about data use on their behalf (see Box 1). Complementing the regulatory regimes that already exist in many countries, these new 'bottom-up' institutions would seek to manage the vulnerabilities to which individuals and groups are exposed as a result of shifting patterns of data use, and to empower individuals to influence the terms by which data about them may be used.

(Box 1) Characteristics of a data trust

Independent stewardship: In exercising data rights held in trust, trustees have a fiduciary responsibility to the beneficiary. These require that trustees act with undivided loyalty in pursuing the best interests of the trust's beneficiaries and create safeguards that help ensure trustees operate independently of other interests when managing the trust.

Institutional safeguards: In common law jurisdictions, trust law offers well-established processes for holding trustees to account, with the overseeing Court offering a route to remedial action against a trustee considered to be in breach of the terms of a trust. In the event of a claim, it is for the trustees to demonstrate that they have sought to promote the beneficiaries' interest with appropriate degrees of impartiality, prudence, transparency and undivided loyalty.

Collective action: Like 19th Century 'land societies' empowered people to get the right to vote, by pooling resources to acquire a piece of freehold land, data trusts could help empower otherwise disenfranchised groups. With access to the bargaining power associated with the aggregation of data in the trust, the trustee would be better-placed to negotiate terms of data use – and respond to undesirable uses of data – than any lone individual.

With the ambition of harnessing the value of data for economic and social benefit, while preventing citizens being exposed to new forms of data-enabled harm, governments across the world have

⁵ Some examples are explored at <https://foundation.mozilla.org/en/initiatives/data-futures/data-for-empowerment/what-is-a-data-cooperative/>

looked to data trusts as a new model of data institution. Shifting conversations about data trusts from theory to practice requires action in key areas to clarify their conceptual foundations and identify ways of working (see Box 2).

In pursuing this action, lessons can be drawn from comparing the approaches and legal models that are being established across different jurisdictions.

(Box 2) Data trusts: from theory to practice⁶

Further action is needed to lay the foundations for data trusts that are accessible, reliable and effective, including research to better understand their conceptual underpinnings and ways of working, and pilot projects to identify best practice in data trust design.

Areas to address in the next wave of data trust design and development include:

Conceptual clarity: How do data trusts fit in the wider data governance landscape, and which core capabilities must sit at the heart of any real-world data trust?

Incentivisation: What forms of governance can align individual and community incentives? And what mechanisms can enable trusts to negotiate trade-offs between these, when necessary?

Accountability: Beyond those already provided by trust law, what institutional safeguards would help ensure a data trust represents and operates for the benefit of those it purports to serve?

Participation, inclusion and digital equity: Which interventions can help make data trusts accessible to all in society, and ensure that all those in a trust can meaningfully contribute to decision-making?

Finance and sustainability: Which business models can help ensure the continued sustainability of a data trust?

Implementation issues: Which use cases can help clarify how data trusts would work in practice?

Developing data trusts: understanding jurisdictional issues and opportunities

A prerequisite for the establishment of data trusts is that individuals have rights over data about them that can be held in trust. Countries and regions of the world differ in the types of right that are available under current legislative frameworks. For example:

- In the EU, personal data rights – including rights of access, erasure and portability between providers – are provided under the General Data Protection Regulation.
- In Canada, data protection laws exist at both federal and provincial levels, with different frameworks for public sector data and private sector data. Adding to this framework, proposed legislative changes through the Consumer Privacy Protection Act would bring de-identified data within the scope of data governance frameworks and restrict the reasons for which this can be shared.

⁶ Adapted from Data Trusts Initiative (2020) Data trusts: from theory to practice: Working Paper 1 ([available at www.datarusts.uk](http://www.datarusts.uk))

- In South Africa, data rights are created through both data protection legislation and constitutional provisions, based in human rights law, as well as database laws.
- In the US, few personal data rights exist through federal law. At state-level, there are examples of data protection-like legislation. In California, for example, the California Consumer Privacy Act seeks to give individuals control over the type of information that is collected about them by businesses through rights including the right to delete personal information or opt out of the sale of this information.

In addition to these frameworks for personal data, there exist legal restrictions on different uses of data, and different approaches to managing ‘raw’ and ‘derived’ datasets. Copyright and intellectual property legislation, for example, sets out terms of acceptable use for some forms of data. In some areas – the EU, for example – new legislative frameworks also seek to promote the use of data for social or altruistic purposes, by creating regulatory frameworks to promote such uses,⁷ or to prevent data use for undesirable purposes.⁸

Building on these local legislative foundations, a range of different approaches to collective data governance challenges have emerged across jurisdictions (Table 1). These initiatives highlight a range of operational issues that data trusts must navigate, including:

- Business models for long-term sustainability;
- Technical architectures for data access or management;
- Methods of participatory governance;
- Perceptions of trustworthiness;
- Distribution of benefits from trust activities.

Table 1. Attempts to establish alternative data institutions

Driver’s Seat is a driver-owned cooperative that seeks to empower gig workers to access data from their rideshare or delivery work and use this data to improve their work patterns and earnings. It also sells insights about mobility patterns to local government to inform planning decisions.⁹ This is one example of an emerging model of cooperative data governance that provides a repository for data, with a user interface that allows members to access insights from that data. Similar models are being explored by sectors including utilities and finance.¹⁰

OPAL (Open Algorithms) is a non-profit that supports private sector data sharing for public good purposes. OPAL acts as a form of data intermediary – it uses privacy-preserving technologies to enable researchers and policymakers to access data sources held by companies, without releasing commercially or personally sensitive insights. Use cases of OPAL technologies in telecoms are showing how these technologies can be used alongside participatory governance methods to analyse data in ways that inform public policy while supporting different stakeholders to have a say over what forms of data access are – or are not – acceptable.¹¹

⁷ The EU’s draft Data Governance Act is an example of such a framework. <https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-european-data-governance-data-governance-act>

⁸ See, for example recent calls by regulators to ban targeted advertising on social media:

<https://www.forbes.com/sites/emmawoolacott/2021/02/11/european-regulator-calls-for-ad-targeting-ban/>

⁹ Further information about this initiative can be found at: <https://www.driversseat.co>

¹⁰ Further information about some of this work can be found at: <https://www.cooperative.com/news/Pages/nreca-annual-meeting-2020-data-strategy-can-help-co-ops.aspx>

¹¹ For further information, see: <https://www.opalproject.org/about-opal>

Sidewalk Labs is perhaps the best-known proposal for a data trust in Canada.¹² As part of a waterfront development project, developers proposed an urban data trust to govern data collected from across the city’s infrastructure. Under this proposal, data would be deidentified at source, then governed by a public authority according to Canadian public data legislation. However, this model of operation did not account for citizen concerns about surveillance, consent, and accountability or oversight of data generated by people’s daily activities. This project illustrates how the term ‘data trust’ is being co-opted to describe a variety of data governance activities, including activities that do not reflect the aspirations that data trusts act as a tool for empowering communities. The failure of the Sidewalk Labs project also highlights the importance of designing data trusts in ways that respond to citizen needs and concerns, creating space for ‘bottom-up’ engagement.

UK Biobank stewards health data donated by 500,000 individuals for use in research. Operating under the GDPR, UK Biobank acts as a data controller, managing data for which consent has been given when participants engaged with the project. Biobank’s work is overseen by governance structures that include a trust, which aims to improve the diagnosis and treatment of serious illnesses.

PlaceFund is a non-profit organisation, which operates as a geospatial data trust promoting data use to help address challenges relating to insecure property rights, unsustainable land use and climate change. Core to PlaceFund’s ambition is to create an operational data trust that collects and maintains geospatial data on a sustainable basis, making this accessible to the communities that need it.

AirBnB is currently exploring alternative approaches to its governance and allocation of data rights. The company proposes to create a trust-like structure, which would distribute equity benefits from the company’s success to the trust’s members. This structure would be complemented by an oversight board, the members of which would be user-elected and which would consider how benefits from the company’s growth should be allocated.¹³

The extent to which data trust projects are able to draw directly from the framework of trust law in establishing their ways of working varies across jurisdictions (see Table 2 for examples). Where trust law – as understood in common law jurisdictions – is not available, there exist alternative legal frameworks that can foster independent stewardship and enable collective action, while providing institutional safeguards.

Table 2: Legal frameworks for data trusts in different countries	
Country	Legal framework for data trusts
Canada	Canada’s legal system has both common law and civil law elements. Discussions about the legal frameworks for data trusts are at a nascent stage. Different provinces can make use of different frameworks for establishing data trusts and in some areas purpose trusts can be established for pursuing philanthropic, if not necessarily charitable, objectives. In Quebec, which is a civil law jurisdiction, a Quebec Trust can be established for any non-charitable purpose and can hold property or other rights.

¹² Some context about the Sidewalk Labs project is available at: <https://medium.com/sidewalk-talk/why-were-no-longer-pursuing-the-quayside-project-and-what-s-next-for-sidewalk-labs-9a61de3fee3a>

¹³ For further information, see: <https://www.bloombergquint.com/gadfly/airbnb-host-endowment-could-transform-sharing-economy>

	The legal framework provided by these trusts functions to ensure they operate independently of outside influence and that trustees can be held to account for their decisions in managing the trust.
India	In late 2019, India’s Ministry of Electronics and Information Technology established a committee tasked with designing a framework for the regulation of non-personal data. In July, this Committee of Experts on Non-Personal Data published its recommendations, putting data trusts front-and-centre in its proposals. ¹⁴ This proposes a policy framework in which a data principal – the individual, community or organisation to which the personal data relates – would exercise its data rights “through an appropriate community data trustee”. That data trustee should be “the closest and most appropriate representative body for the community concerned”, and could be a government entity, citizens’ group, or a university, for example.
UK	In the UK, trust law has long been used to manage the rights and responsibilities associated with different assets. A Trust is formed when a person in whom a set of resources is vested—the Trustee—is compelled to hold and manage those resources either for the benefit of another person(s)—the beneficiaries—or for some legally enforceable purpose(s) <i>other than the Trustee’s own</i> . ¹⁵ Bound by a fiduciary obligation of undivided loyalty, data trustees would exercise the data rights conferred by existing regulations (such as the EU’s General Data Protection Regulation) on behalf of the trust’s beneficiaries. Acting as an independent intermediary between data subjects and data collectors, the data trustee would leverage the bargaining power associated with the aggregation of data or rights in the trust to negotiate terms of data use in accordance with the trust’s terms
US	<p>The US has a history of agricultural cooperatives and labour unions in which individuals have worked together to negotiate rights and working conditions. This legacy of employee-ownership could point to areas in which data trusts could emerge as a means of empowering individuals or local communities in decisions about data use.</p> <p>The Delaware Statutory Trust Act allows for the establishment of trusts in Delaware that can operate across the US, but that are subject to the jurisdiction of the Delaware Chancery Court. The Act allows those establishing the trust flexibility in their structure and governance, including their specification of fiduciary duties and ways of working, within the well-established guardrails set by Delaware case law. If those establishing a trust wish to direct the work of the trustees – limiting their discretion over how to manage the trust’s assets – there exists an option to create a sponsor body; an LLC owned by the beneficiaries of the trust, through which those beneficiaries can influence the work of the trustees.</p>

¹⁴ This report is available at: <https://ourgovdotin.files.wordpress.com/2020/07/kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf>

¹⁵ Delacroix, S. and Lawrence, N. (2019) Bottom-up data Trusts: disturbing the ‘one size fits all’ approach to data governance, *International Data Privacy Law*, Volume 9, Issue 4, Pages 236–252, <https://doi.org/10.1093/idpl/ijp014>

Insights from cross-jurisdictional comparisons

Adapting local frameworks to new governance challenges

The term 'data trust' is now widely used to describe a collection of different approaches to data governance. The Data Trusts Initiative takes as its starting point that a data trust is a mechanism for individuals to pool their data rights into an organisation that: provides independent stewardship of those data rights; embeds fiduciary responsibilities in its ways of working; operates within a frameworks of institutional safeguards; and facilitates collective action.

Thinking in terms of these core characteristics can help identify relevant legal frameworks across different jurisdictions. It can also highlight areas where further clarity is needed, if data trusts are to operate effectively in these different jurisdictions, including:

- What rights can be asserted in different jurisdictions, how do these rights interact with each other, and how they can be settled in a data trust;
- What value can data trusts add to the local data governance environment, based on its history, stakeholders and needs;
- The purposes for which data trusts might be better suited than other legal mechanisms for data stewardship;
- Why or how different stakeholders might be incentivised to – or disincentivised from – creating a data trust, and what value data trusts create for those stakeholders.

Managing tensions between individual and collective interests

Data trusts offer a means for individuals and communities to enact their data rights and direct patterns of data use to match their interests. In many circumstances, it is reasonable to expect that the interests of those engaging in a data trust will be in tension: individuals might differ in their priorities, or the interests of individuals will not align with those of the community. These tensions will influence the nature of fiduciary duties that need to be built into trust governance and trustee ways of working. A trustee upholding their fiduciary duty to act for the best interests of an individual, for example, would behave differently to one prioritising the needs of a community or acting for wider social interest.¹⁶ The design of fiduciary duties that reflect the needs of user-led governance frameworks is an important area for further attention in the development of data trusts.

If data trusts are to be successful in practice, they will require structures and processes that can align these differing interests or make trade-offs in ways that are acceptable to trust members. In addition to considering the interests of their beneficiaries, data trusts will also need to grapple with the strong commercial incentives that shape the behaviour of data-holding organisations. Further work is required to better understand the value proposition that data trusts offer individuals, communities and organisations working with data, and how this value proposition can align the interests of these different parties.

When developing this value proposition, history may offer insights into the action needed to create sustainable funding and resourcing strategies for providers of public goods. The 19th century, for example, saw a range of institutional innovations in the UK that laid the foundations for long-term social infrastructures for public benefit.

Tailoring governance structures to context and purpose

¹⁶ For a discussion of the tensions that can arise, see <http://www.jenitennison.com/2020/12/27/individual-collective-community.html>

The next phase of development for data trusts will require real-world pilot projects. Creating such projects will require close engagement with potential beneficiary communities to identify their needs, understand their expectations around data use, and design governance structures that meet these requirements. In so doing, these projects can better identify how core data trust characteristics – such as fiduciary duties for data trustees – might operate in practice.

Jurisdictions differ in the rights they provide, and communities will differ in their ambitions for data use. Pilot projects will need to adapt to these differing circumstances, operating within local regulatory frameworks and user demands. In some contexts, there may be opportunities to build on existing civic institutions to create new data institutions. In the US, for example, companies that have built electricity infrastructure and social credit unions could benefit from creating data trusts to govern the data they hold and direct its use for community benefit.¹⁷

Coming years will likely bring a variety of experiments around new forms of community-focused data governance. While pursuing these innovations, those leading such projects have a responsibility to consider what happens when these projects fail. Data institutions stewarding common data assets, for example, must consider how those assets can be maintained in the long-term, should the institution disappear, while those working with personal data must consider the protections that need to be in place. While trust law provides more safeguards than other legal mechanisms in this respect – the ability for the courts to appoint a caretaker trustee in the event a trustee fails in their duties, for example – those creating data trusts must carefully consider what frameworks for performance monitoring, evaluation and sustainability should be in place.

¹⁷ This example is explored further at: <http://ide.mit.edu/news-blog/blog/qa-sandy-pentland-data-cooperatives>

About the Data Trusts Initiative

The Data Trusts Initiative is an interdisciplinary programme that pursues research at the interface of technology, policy and the law to better understand the role data trusts can play in addressing the challenges of data governance in the 21st century.

Supported by a donation from the Patrick J McGovern Foundation, the Data Trusts Initiative will fund research and engagement activities to clarify the conceptual foundations of data trusts and explore potential models of operation. By building a community of researchers and social entrepreneurs, the Initiative will shift discussions about data trusts from principle to practice.

For further information about the Initiative, and opportunities for funding in some of the areas discussed in this note, please visit www.datatrusts.uk