Alternative Data Governance Approaches:
Global Landscape Scan and Analysis

Authors: Stefan Baack and Madeleine Maxwell for Mozilla Insights.

Co-authoring regional researchers: Afef Abrougui, Beatriz Botero Arcila, Tetyana Bohdanova, Claude Migisha, Marilia Monteiro, Natalie Pang.

Table of contents

Preface

Ideas about new forms of data governance are on the rise, suggesting alternatives to big tech monopolies. Could this help us *reimagine, reconstitute,* and *rebalance* skewed power dynamics?

How data is controlled and governed is central to Mozilla's understanding of internet health and trustworthy AI. We are committed to working with others to reimagine data governance as a building block toward a [trustworthy AI ecosystem](#).

The Data Futures research presented here is an open study by Mozilla Insights that examines approaches to data stewardship and alternative data governance (in theory and in practice).

This research will help inform decisions to support new technologies and infrastructure, as well as support further cross-disciplinary thinking as part of Mozilla's new Data Futures Lab. The Data Futures Lab will resource projects and foster networks across the ecosystem of actors seeking to realize data governance alternatives which redirect agency, ownership, value, and power back to the people from whom these elements are commonly extracted. We believe that informed and collaborative experimentation in the data governance field, motivated beyond the consolidation of power and profit, should be a priority and at the center of our work in this space. The more people deploy trustworthy models to govern and handle data in the future, the bigger impact it will have on the health of the internet and the AI field we are trying to change.

To kick things off with the Data Futures Lab, we will initiate convenings and we'll fund initiatives to design, build, test, or scale future products that can unleash the creative, financial, and political potential of more equitable data futures.

As part of this broader effort, [this research series](#) will be expanded periodically with more publications throughout 2020. We are committed to working in the open, collaborating with others, and sharing our findings widely to advance and complement a growing community of thought.

Summary

Mozilla's Insights team led an international landscape scan and analysis to surface data governance initiatives that manage and share benefits from data in new and interesting ways.

Rather than rely on terminology that is imprecise or not commonly known, we chose to categorize initiatives based on a set of key characteristics. We intentionally sought out projects, services and companies that present clear alternatives to dominant digital business models, like those of Facebook, where a handful of large corporations amass large amounts of data to gain a competitive edge and freely commercialize personal data by securing all control over data.

We looked for projects that do at least one of the following:

- Shift agency between data collectors, data subjects, and other beneficiaries in a meaningful way.
- Share the benefits of data between various parties rather than concentrating most or all of the value within a single organization.
- Manage data in ways that represent multiple interests (the data collectors, data subjects, or other beneficiaries of the initiatives).

This research had three goals:

1. Create a database of initiatives worldwide that align with one or more of the three criteria outlined above.
2. Describe the core characteristics of these initiatives.
3. Understand how 'alternative data governance' is understood in different regions around the world.

The analysis is based on a public survey with 70 responses, reviews of databases by other research groups, and findings from seven regional researchers working in local languages. In total, we reviewed and categorized 110 initiatives worldwide across Africa, the Middle East, Asia, Europe, the Americas, and Oceania.

The database can be found at airtable.com/shrn9jnFOQByon2i7.

Key distinguishing characteristics

We analyzed the initiatives in our database according to four key characteristics about what they claim to achieve, whom they aim to benefit, some basic attributes of the data, as well as who has access to it.

1. What is the primary benefit?	
While most initiatives have more than one stated benefit, we identified one primary purpose (out of five) for each.	1. **Increasing data availability:** Gathering data that wasn't available before and sharing it publicly or within groups, e.g. to increase government transparency. 2. **Increasing data accessibility**: Making data that is already available somewhere easier to use or digest for a broader audience, e.g. by providing a unified interface to access and search through data from various sources. 3. **Giving data subjects or interest groups more control over data**: Initiatives that ensure rights, provide insight into what third-parties have access to one's data, or forms of consent management like automatically adjusting privacy settings on various third-party services. 4. **Gaining benefits from data sharing**: Initiatives that directly reward individuals for sharing their data, e.g. by monetizing it; or where groups or communities share data among themselves for mutual benefit. 5. **Enabling privacy compliance**: Initiatives that help organizations comply with certain privacy regulations.
2. Who are the primary and secondary beneficiaries?	
Who is the initiative trying to serve?	This can overlap with 'data subjects' (the people whose personal data is used) but it does not have to. Often, initiatives have both primary and secondary beneficiaries. For example, many health sector initiatives have both patients and medical researchers as beneficiaries, though they differ in whether the patients or researchers are the primary or secondary beneficiaries.
3. What type of data is it?	
Whose data is it and how was it gathered?	We logged whether the data being collected or governed is personal or non-personal, disclosed or observed. Only some projects deal with personal data, and only in some cases do individuals actively contribute or share data to the initiative.

4. Who has access to the data?	
Who can access the data that is being collected or governed?	This can be individuals or beneficiary communities, but could also be third-parties (e.g. when an initiative gives individuals more control over who can access their data) or instead is simply openly available to the public (open access).

Our five types of data governance initiatives

By correlating the primary benefits, the primary beneficiaries and sectors of the initiatives in our database, we identified five prototypical types of data governance initiatives.

1. **Data donation initiatives**	These act as mediators between two beneficiaries and facilitate the data donations of one beneficiary to another without an expectation of remuneration. Often, these initiatives connect patients and medical researchers and allow patients to donate their data.
2. **Individual control initiatives**	These are initiatives focusing on giving individuals more control over their personal data via privacy protections and easy ways to control how one's data is shared.
3. **Group protection and empowerment initiatives**	These are dedicated to protecting the rights and interests of particular groups or communities, such as Indigenous data stewardship initiatives.
4. **Public-facing data collection initiatives**	These are dedicated to collecting data and making it publicly available.
5. **Data exchange networks**	These are dedicated to pooling data held by numerous organizations who agree to create a shared resource.

Alternative data governance around the world

Local researchers focused on five regions: Eastern Europe (Russian and Ukrainian), the Middle East and North Africa (French and Arabic), Sub-Saharan Africa (French and English), Latin America (Spanish and Portuguese), and Southeast Asia (English and Malay language).

Reachers all found that 'alternative data governance' (in discourse and in practice) is relatively uncommon beyond Western Europe, the United States, and Canada. They found that it is mostly associated with open data or public data collection initiatives. Furthermore, they identified a number of likely prerequisites for alternative data governance approaches to be of greater local relevance in individual countries, including the existence of data protection laws, prevalence of open data initiatives, traditions with cooperatives in other contexts, expectations over whether governments or the private or civic sectors are responsible for holding and managing data on citizens, the maturity of markets for data and digital ads, and the prominence of data privacy movements. Having more pressing and immediate human rights and digital rights concerns were also cited as a likely barrier to local innovation with data governance, especially in Africa and the Middle East.

How is 'alternative data governance' understood?

We chose the term 'alternative data governance' for this research because we assessed that it would be easier to apply in practice among people who are not immersed in discourse on data, as opposed to a term like 'data stewardship' that has fluctuating definitions even among its proponents. By data governance, we simply mean how data itself is governed in various ways (eg. as a data trust or a cooperative). However, judging from the responses we received to a public survey, findings made by regional researchers, as well what we observed in other databases by research groups, it appears that 'alternative data governance' is often understood in much broader terms. For instance, quite often, initiatives that were suggested for inclusion in our database were not primarily about managing data, but about making data openly accessible to affect governance decisions. For example, projects that collect data to challenge or improve on official data sources, or initiatives that enable the collection and management of data in new ways for different purposes without actually prescribing specific governance approaches.

It seems to us that empirical investigations in this emerging field are likely to be clouded to some degree by lack of shared understandings of key terminology, particularly across different languages. Many simply still do not consider data governance to be of

central importance, even when they are actively engaged in data privacy and digital rights advocacy or in developing civic technology platforms. Beyond a relatively small circle of people and organizations discussing alternative data governance as having the potential to solve power inequities of the internet, numerous people across different nationalities consulted for this research indicated that they do not currently consider this a core priority to their interests.

Areas for further research

We identified a number of key gaps between theory and practice that would be good topics for further research:

- A literature review commissioned by Mozilla's Insights team in parallel with the empirical research presented here identified [7 data governance approaches](#) often discussed in scholarly writing. We could not easily match these approaches to the initiatives in our database as categories. On the one hand, this is because a number of approaches overlap and lack firm characteristics. On the other, it is because the majority of initiatives we reviewed offer scant information about their own data governance approach. It would be interesting to apply the characteristics we describe in this report to the different data governance approaches prominent in scholarly discussions around alternative data governance.

- Additional research into the ecosystem of organizations and structures that support advancements in data governance would complement findings from this review and help to explain regional differences in practice.

- Further investigations into the cultural and social factors that influence understandings of alternative data governance is important. Much of the theoretical literature is about legal mechanisms and systems of accountability. This needs to be expanded, e.g. by asking epistemological questions of what gets highlighted (by drawing from feminist data practices and related work); by investigating the values of decision-makers and founders as well as of communities in this space.

Ten examples of initiatives

Below we have summarized ten initiatives that stood out to us as being particularly unique.

Initiative	Description	Region	Why it stands out
Amazonia Socioambiental	Initiative to map threats to the Amazon and Indigenous territories.	Latin America	Gives voice to Indigenious communities and shows the connection between Indigenous rights and environmental destruction.
eRouška (emask)	COVID-19 contact tracing app that preserves privacy.	Europe	An initiative started by a non-profit with support from the Czech Ministry of Health.
Our Brain Bank	A patient-led movement designed to move glioblastoma from terminal to treatable, powered by patients.	Multiple	Combines community support with data collaboration between patients and researchers focusing on a rare type of cancer.
PySyft & PyGrid	Open source software libraries to enable secure, privacy-preserving machine learning and data science.	Global	Aims to change industry standards and enable the use of AI technology while preserving privacy.
Glimpse	Glimpse Protocol connects companies to consumers while respecting the privacy of individuals as required by GDPR and CCPA regulation.	Europe	Enables brands to reach precise audiences without collecting their data while complying with data privacy regulation.
British Columbia First Nations' Data Governance Initiative	In the Canadian province of British Columbia, the initiative supports First Nations with the technological and human resource capacity to govern and own community data.	North America	Supports the self-determination and well-being of an Indigenous community by enabling it to own, control, access, and possess information about its people.
MIDATA	MIDATA shows how data can be used for the	Global	The non-profit cooperative operates a data platform, acts as a trustee for data

	common good, while at the same time ensuring citizens' control over their personal data.		collection, and guarantees the sovereignty of citizens over the use of their data.
DECODE	DECODE creates open data commons from data produced by individuals and devices, enabling citizens to control and share their data for the common good on terms that are fair and transparent.	Europe	The combination of smart rules running on distributed ledger technologies will produce a platform that is fully decentralized and allows flexible, extensible data governance.
Driver's Seat Cooperative	A driver-owned rideshare cooperative, empowering gig workers and local governments to make informed decisions with insights from their data.	North America	Collects and sells mobility data to city agencies so they can make better transportation planning decisions. When the Driver's Seat Cooperative profits from data sales it shares the wealth by distributing dividends to the driver-owners.
Raval Data Commons	A data commons created by various local actors, in the district of Raval in Barcelona, Spain.	Europe	Participatory governance, data about and from citizens, managed by citizens.

1. Introduction

Mozilla's Insights team led an international landscape scan and analysis to surface innovative data governance projects and products anywhere in the world (and in any language) that manage and share benefits from datasets in new and interesting ways.

From data coops and data commons projects, to data trusts, personal data storage solutions, and more, we were interested in identifying approaches that differ from dominant digital business models around data, whose data governance model has been described by Stuart Mills (2019) as 'Laissez Faire data ownership.' Data collectors like Facebook acquire all data ownership rights via user consent in exchange for 'free' services, essentially meaning that all value is extracted by the data collectors alone. This approach to data governance is prone to (1) over-restrictive hoarding of data, managing data solely within organizational silos; and (2) over-sharing of data in ways that expose data subjects (both individuals and communities) to undue risks without their knowledge or consent.

In contrast to this approach, we sought out alternative data governance structures that seek to promote different ways of sharing data, particularly between different stakeholders and organizations, while giving more power to the people and communities represented or affected by the data. Our investigation is centered on whether and how the value and benefits derived from data can be shared more equitably as opposed to reinforcing the leader-takes-all logic that is the typical default for capital-backed network technologies.

The goal of this report is twofold. First, to present a database of initiatives that exemplify alternative data governance approaches around the world and describe some of their core characteristics. Second, this report asks how alternative governance approaches are understood by organizations and activists close to Mozilla as well as by activists, governments, and entrepreneurs in different regions around the world. Therefore, we do not rely on a particular definition of 'alternative data governance' followed by a systematic search based on that definition. Instead, we provided a rough outline of what we mean by 'alternative data governance' and invited survey responses, reviewed other databases, and hired regional researchers to collect relevant initiatives from beyond North America and Europe (see next section for more details).

This report is part of a broader research project to help inform Mozilla's understanding of and strategies for supporting alternative data governance innovations. It complements a [literature review](#) on alternative data governance commissioned in parallel, which provides definitions and delineations of data governance approaches as discussed in literature and among scholars and activists. Together, these two reports

form a foundation for upcoming analyses of the data governance ecosystems and enabling structures, research about the demand for such approaches, as well as qualitative research focusing on people founding and building alternative data governance initiatives.

The guiding questions in analyzing the initiatives collected in our database were:

- In what ways is control over and value from data being shared, as compared to the dominant model exemplified by big tech platforms like Facebook?

- What differences do we see between approaches in different parts of the world?

The database can be found at airtable.com/shrn9jnFOQByon2i7.

2. Approach and scope

Over a period of six weeks from May to June 2020, an international team of nine researchers conducted a landscape scan aiming to surface a diverse set of data governance approaches, initiatives, and products from around the world. The goal was not to gather a comprehensive database, but rather to get a sense of what types of data governance initiatives exist and which ones appear to be more common in different regions around the world.

First, we defined the scope of the landscape scan, conducting a first preliminary literature review of "data governance." Then we decided what information to gather about each initiative.

We logged initiatives that do at least one of the following:

1. **Shift agency** between data collectors, data subjects, and other beneficiaries in a meaningful way compared to the dominant model of data governance described above.
2. **Share the benefits** of data between various parties rather than concentrating most or all of the value within a single organization.
3. Manage data in ways that **represent multiple interests** (the data collectors, data subjects, or other beneficiaries of the initiatives).

We collected initiatives in three steps. First, building on the preliminary literature review and existing team knowledge, we added examples from Mozilla's broader network or found via desk research using key search terms and adapting language to local/regional contexts. These searches were not extensive or systematic, but were meant to serve as a starting point for designing the database. Second, we distributed a survey in seven languages (English, French, Arabic, Spanish, Portuguese, Russian, and Ukrainian) within our networks, mainly via Twitter, which yielded 70 responses. The survey asked 'What data governance projects inspire you?' and described alternative data governance similar to how we described it in the introduction above. Respondents provided the name and URL of the project as well as a description of what makes the initiative unique.

Third, we involved seven regional researchers: two for Sub-Saharan Africa, two for Latin America, one for Eastern Europe, one for Southeast Asia, and one for Middle East and North Africa (see list of names in the acknowledgements). They conducted searches in various languages, spoke to local activists, entrepreneurs, or members of

relevant government initiatives, and helped distribute the (translated) survey. During the research process, we had three group calls with the researchers: first to brief them on the shared terms and understandings, second to discuss preliminary findings, and third to share final results and discuss regional particularities. Each researcher added initiatives to the database and six of them concluded their work with a short report summarizing their findings. These reports formed the basis for our summary of regional trends.

Finally, we also reviewed two pre-existing databases of alternative data governance initiatives: GovLab's (2020) Data Collaboratives Explorer and a preliminary version of the Data Stewardship Explorer by the Data Economy Lab of the Aapti Institute (2020). Each of these databases was built on different definitions and categories from the ones in our database. We therefore reviewed every initiative in these databases and only added them if they fulfilled one of the selection criteria outlined above. Finally, we also included initiatives collected by Anouk Ruhaak (2020), a Mozilla Fellow researching data trusts at the time of writing.

For each initiative, we captured a wide range of available information to understand the aims, scope, and context. The information we gathered fell into four categories: basic information about the initiative itself (e.g. geographic base or legal status), the beneficiaries of the initiative, information about the data that is produced or managed by the initiative, and finally some context related information like whether and how the initiative relates to the COVID-19 pandemic. These four overarching categories helped us assess how the collected initiatives distribute agency and power among the involved parties and how this differs from dominant data governance approaches. After compiling information against this defined set of characteristics for each initiative, we conducted an inductive analysis to understand key distinguishing features between the initiatives gathered.

Limitations

As mentioned, this is not a comprehensive database, but a landscape scan to get a sense of the range and reach of alternative data governance initiatives in different parts of the world.

We categorized initiatives based primarily on what they claim to do and who they claim to benefit. Notably absent, however, is a categorization of their governance model or approach. This is partly due to the fact that understanding the definitions of alternative data governance approaches was commissioned as a separate study to this one

completed in parallel with ours. The primary reason for the absence of data governance approaches, however, is the paucity of information that is publicly available around how decisions are made as part of individual initiatives, and what mechanisms and systems exist to ensure accountability. Very few of the initiatives we reviewed provided any information about decision making processes and lines of accountability, and even among those exceptions, information often remains basic, e.g. mentioning an elected board and annual review without much specificity, or emphasizing security measures like data anonymization rather than governance approaches. This indicates that there is a gap between theoretical discussions of data governance approaches — which tend to focus on legal frameworks for decision making and other responsibilities — and practice, where the majority of initiatives we collected do not prioritize sharing this information. However, we hope that our analysis and the categories we suggest could help add nuance to the definition and delineation of alternative data governance approaches in future research.

The second limitation is that we formulated our core concepts and goals based on discussions in North America and Europe, which turned out to limit our access to, and findings in other regions. As a result, we received very few survey responses in languages other than English, even though we translated our survey. We could have done more to distribute the survey via additional communication channels. But the poor response might also support the findings of our regional researchers: that terms like 'alternative data governance,' 'data stewardship,' 'data trust' etc. do not resonate in non-English languages. They are difficult to translate and largely unknown (see a discussion of regional findings below). This does not mean that there is a complete lack of interesting initiatives related to the governance of data in those regions: it may simply be that a different conceptualization of our research interests would result in more findings.

Third, we need to emphasize that this study is deliberately 'naive' in the sense that we categorized initiatives based on their own claims, e.g. how they claim to benefit particular groups and what impact they seek. We did not critically evaluate these claims or study the 'actual' impact of their services (e.g. by studying who is using these services for what purposes). This report aims to provide a basis for such critical evaluations in follow-up investigations, e.g. by allowing us to compare claims with actual impact.

Finally, we decided to exclude official (i.e. government initiated) open data portals from the database. While they might arguably fulfill some of our selection criteria, we think this space is already well covered by others.

3. Database overview

In total, we assessed 110 initiatives from 37 countries in Africa, Asia, Europe, the Americas, Middle East and North Africa, and Oceania. Of the initiatives we reviewed, the majority are established initiatives (84%) with the remaining 14% being pilots (we considered them pilots when we saw any clear indication that the project is in an early stage). The two largest groups of initiatives belong to the health sector (25%) and technology sector (19%). The technology sector in our categorization primarily contains initiatives that develop technology solutions that can be applied across other sectors. The other initiatives are scattered across a range of sectors, from transportation to advertising, agriculture, and more. (Figure 1).
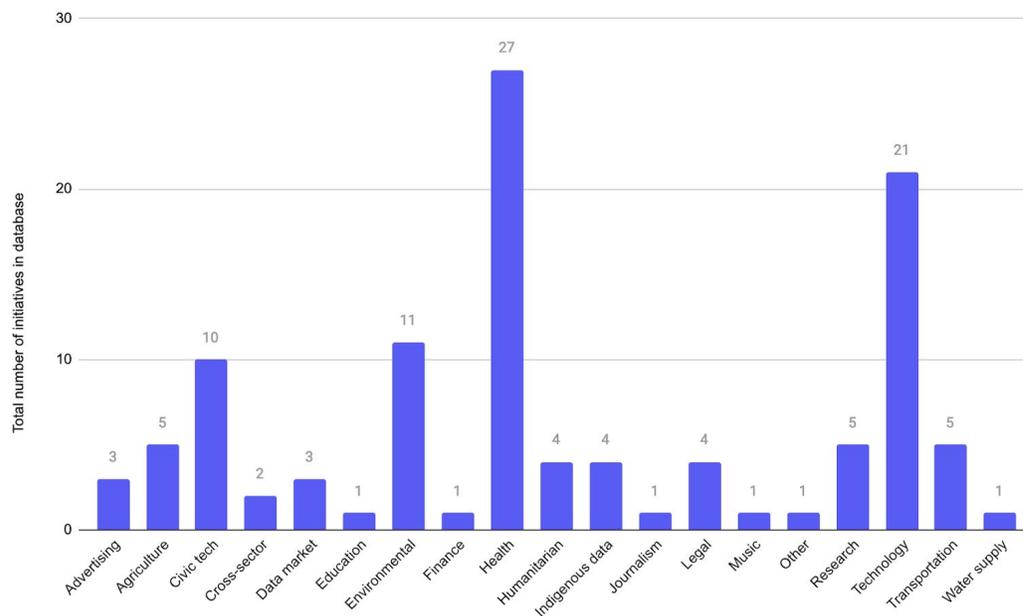

Figure 1: Different sectors represented in the database.

The database consists of 55% non-profit organizations, 20% for-profit initiatives, and 17% government related initiatives. 7%, are cooperatives (Figure 2). We list cooperatives separately although they can be for-profit or non-profit due to their relevance for alternative data governance approaches.
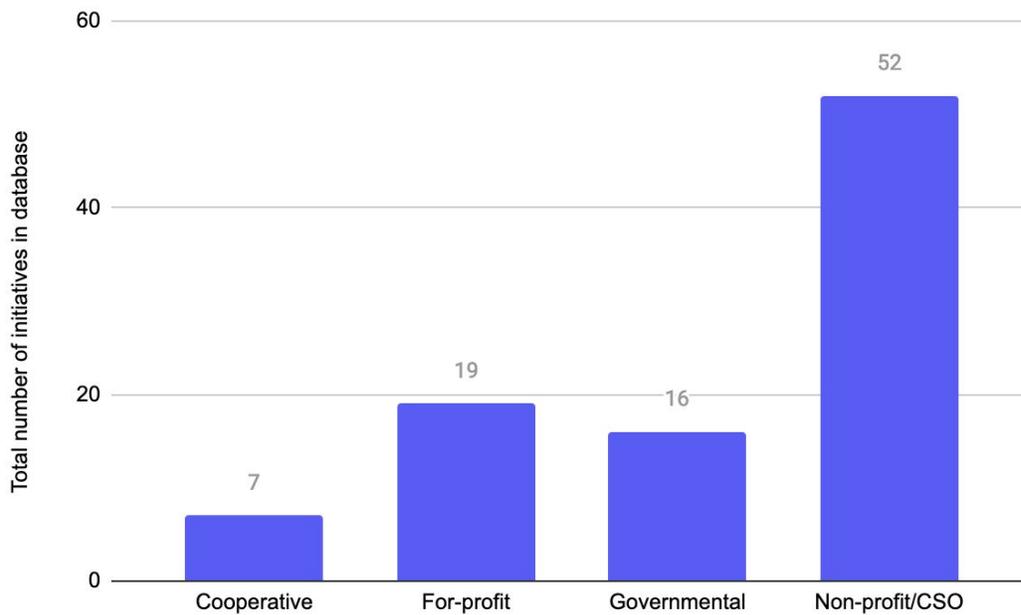
Figure 2: Legal status of initiatives in our database. Note that for some initiatives, we could not clearly determine the legal status and excluded them from this figure.

Filtered by sectors, for-profit companies are most common in the technology sector, but are otherwise relatively uncommon compared to non-profits. Non-profit organizations are most common, and also cover more sectors than for-profits, governments, or cooperatives (Figure 3).
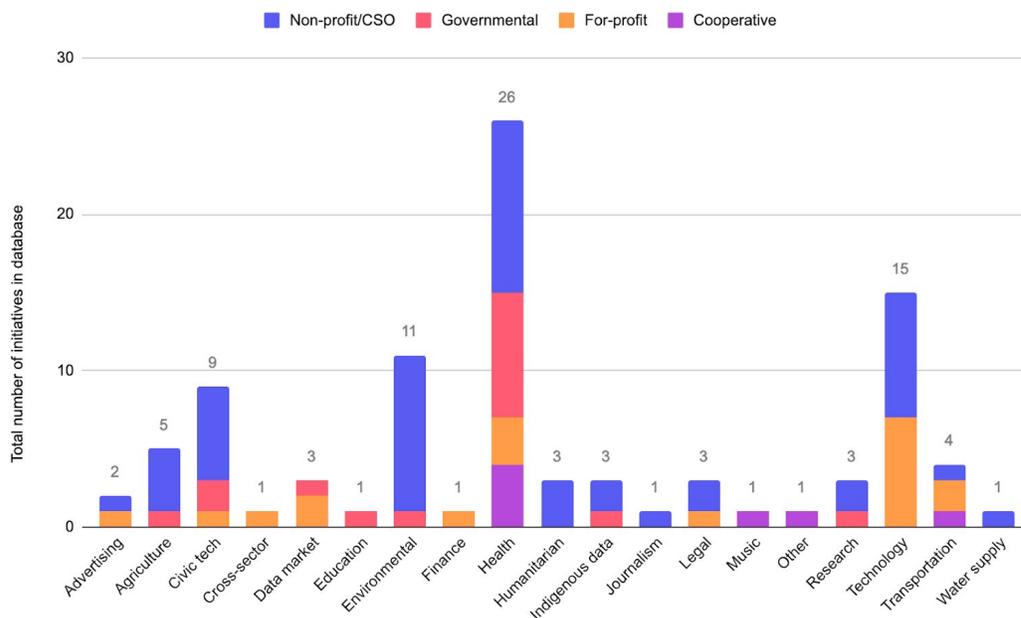

Figure 3: Legal status by sector. Initiatives where we could not determine the legal status are excluded.

Through this research, we also identified supporting entities. These organizations aren't implementing new data governance approaches themselves, but support other actors in the ecosystem to do so in a variety of ways. This supportive ecosystem was not the primary focus of this research, but we will outline findings from a rough analysis in the concluding section.

4. Analysis

4.1 Key distinguishing features

The initiatives in our database were separated along four key distinguishing features: the primary benefit, the beneficiaries of the initiatives, the type of data being governed, and to whom the data is accessible. These features were chosen to assess what the initiatives claim to achieve, who is supposed to benefit, and some basic characteristics of the data that is being governed.

What is the primary benefit?

First, we captured the primary benefit that the initiatives in our database claim to offer their beneficiaries with regards to data. Note that this is not necessarily identical to the overall goal of an initiative, which can be much broader than the immediate benefits provided by the services or tools offered. We chose this narrower focus to better understand how the alternative data governance approaches employed by the initiatives distribute the benefits of data differently from the dominant online services with their 'Laissez Faire data ownership' model (see introduction).

We identified five primary benefits with the following subcategories in mind:

What is the primary benefit? (the five categories explained)	
1.) Increasing data availability This refers to instances where data is gathered that was not available before, e.g. via crowdsourcing. Making new data available is usually not a goal in itself but rather has various functions, including:	*Government transparency*: Typically these are data collection initiatives that gather (or publish) information about politicians or government institutions that would not otherwise be available.
	Environmental transparency: These are initiatives that provide an overview of a particular geography. We refer to 'environment' broadly to include both ecological and social environments. An example of the latter is Strava Metro, a company in the US that collects movement data from users via a smartphone app and uses it to collaborate with officials to improve infrastructure for cyclists and pedestrians. An example of an overview of an ecological environment is InfoAmazonia an initiative that maps the Amazon region to generate public awareness of environmental problems.
	Connecting particular data subjects with data collectors: This is most common in the health sector, where initiatives offer a way for patients to privately give personal data to medical researchers, whose

	work ultimately might benefit the patients who contributed data. An example is UK Biobank, a "national and international Health resource" that collects health data from volunteer participants and shares them with approved researchers.
2) Increasing data accessibility This refers to initiatives that aim to make pre-existing datasets easier to access and use. Similar to data availability, making data more accessible usually has various secondary purposes:	*Improved access to data from diverse sources*: Initiatives that provide access to data from many different sources in one unified and searchable portal. An example is Datos Salvador, an open data portal managed by citizens in El Salvador.
	Encourage and facilitate data sharing: Initiatives that provide a platform that facilitates and encourages particular communities or organizations to share their data with others. An example is AG Data Commons, which aims to provide easy access to and discoverability of data relevant to agricultural research funded by the United States Department of Agriculture.
	Making data machine readable: Initiatives that collect and clean data, for example, to make it searchable. Often these are initiatives relate to government transparency. An example is Justice Lab in India, which collects, standardizes, and republishes legal data.
	Making data intelligible: Often initiatives that map, visualize, or otherwise make data easier to grasp for a broader public. An example is A Tu Servicio from Uruguay, which doesn't just make health data open, but digestible with visualizations and more.
3) Giving data subjects or interest groups more control over data Often technology-driven initiatives that give individuals or groups more control over their privacy and how their data is shared. Secondary purposes may include:	*Ensuring data rights*: Initiatives that ensure individuals or groups that a product fulfills certain standards when it comes to data privacy, security, and portability. An example is Ag Data Transparent, which provides a certificate for companies selling equipment to farmers to indicate that certain principles like data portability are met.
	Enabling data portability for individuals: Initiatives that enable individuals to take their data from one service to another. An example is CommonHealth, an Android app that lets people collect and manage their personal Health data and share it with the health services
	Insights into third-party data sharing: Initiatives that enable individuals to see and better understand what data various third-parties (like social media networks) have about them and how they share it. An example is JoinData from The Netherlands, a tool for farmers to better control what data about their farm is shared with

	whom (e.g. by showing what data is collected by their increasingly digitized equipment).
	Consent management: Initiatives that help data subjects managing their privacy settings on various other services. An example is [Jumbo Privacy](#), a smartphone app that, among other things, helps users to manage their privacy settings on various other services like Facebook.
	Easier management of data about own person/community/group: Initiatives that make it easier for individuals or groups to manage their own data, especially with whom it is shared. An example is [MyData's Commons Prototype](#), a proposal to standardize a 'Decentralized Identity' technology that would give individuals a unified view to manage their data and share it (for example with various health authorities).
4) Gaining benefits from data sharing Initiatives that encourage data sharing of individuals or groups in exchange for various benefits. We can distinguish two different types:	*Individual benefits*: Often, this involves monetizing one's personal data to third parties, as with the [Streamr's Data Unions prototype](#). Other examples are about getting other services as a result from data sharing, as with [CoverUS](#) from the US, where individuals can share their health data to earn rewards and save on health costs.
	Data sharing networks: Groups or communities decide to share data among themselves for mutual benefit and in some cases regulate membership access via fees or other instruments. An example is the [Idaho Health Data Exchange](#), a network of medical institutions that exchange data to optimize patient care across members (e.g. by showing how many patients are in what facility, how many beds are available etc.).
5) Enabling privacy compliance *Services or tools that automate or make compliance with privacy regulations easier.*	*Technology solutions and services*: Initiatives that help groups and organizations comply with certain privacy standards. An example is [Glimpse](#), a protocol that "connects companies to consumers while respecting the privacy of individuals as required by laws like GDPR." Or tools that automate data collection and processing in ways that respect privacy. An example is [PySyft](#), a Python library that seeks to enable more secure and private machine learning.

Who are the primary and secondary beneficiaries?

What is the intended target audience, i.e. who is the initiative trying to serve? This roughly falls into three categories: individuals; groups and communities; or an unspecified, general public.

Many of the initiatives we collected have a primary and a secondary group of beneficiaries and the relationship between these beneficiaries helps us understand the initiative more fully.

Many health initiatives aim to benefit both patients and medical researchers, but an important difference is whether the patients or the researchers are the primary beneficiary group. For example, SavvyCooperative benefits patients by offering them paid gigs to share data about their experiences with healthcare providers, while UK BioBank has researchers as a primary beneficiary because health data from volunteer participants is shared with them to aid their discovery process.

Through their websites, their features, and in some cases also their governance structures, health initiatives primarily serving patients tend to emphasize user benefits, community aspects, privacy, and security, while those prioritizing researchers tend to emphasize the breadth and scope of the patient data and more prominently highlight research papers relying on this data.

Initiatives with multiple beneficiaries will often make their data openly available to the public, but at the same time clearly have an intended target audience. For example, the European Open Science Cloud is a hub for hosting research data "to support EU science." The target audience are researchers from European universities, but the data it hosts is open access and thus may benefit journalists or civil society actors as well.

What type of data is it?

The type of data that the initiative governs is another distinguishing characteristic we logged in our database. Whether data is personal or non-personal, observed or disclosed, can be relevant to questions of consent and multiple other aspects of data governance. Here, we distinguished between four types, inspired by the GovLab's (Verhulst, Young, and Srinivasan 2017) database of data collaboratives:

Disclosed Personal Data: This is data that data subjects actively share themselves. For instance, initiatives in the health sector that encourage patients to	**Disclosed Non-Personal Data**: This is data that is actively shared by individuals, institutions, or companies but does not

share data for research via a questionnaire or other means.	contain any personal information, e.g. citizen science data.
Observed Personal Data: These are instances where data is not actively provided by data subjects, but where they agree to share data traces somehow. For example, by allowing browsing history to be recorded.	**Observed Non-Personal Data**: This is data that is observed and collected but does not contain personally identifiable information. An example could be agricultural or environmental data.

Who has access to the data?

Finally, initiatives also vary in terms of how widely or publicly accessible data is, which is also a key characteristic that has bearing on data governance decisions:

Individual: Typically, this is for initiatives that give users more control over their own data. In these cases, only the data subject or data contributor has access to their own data.
Beneficiaries: The data is only made available to the particular beneficiary group of the initiative. This is typical for many initiatives in the health sector that made the data of one beneficiaries group (often patients) available to another (medical researchers).
Open Access: Data is made openly available to everyone.
Third-parties: Initiatives like personal data stores provide individuals with a way to control who can access their personal data, i.e. the data is ultimately shared with various entities who are neither data subjects nor beneficiaries of the initiative.

4.2 Types of data governance initiatives

To identify prototypical data governance initiatives, we analyzed how the primary benefits, the primary beneficiaries, and the sectors of the initiatives correlate.
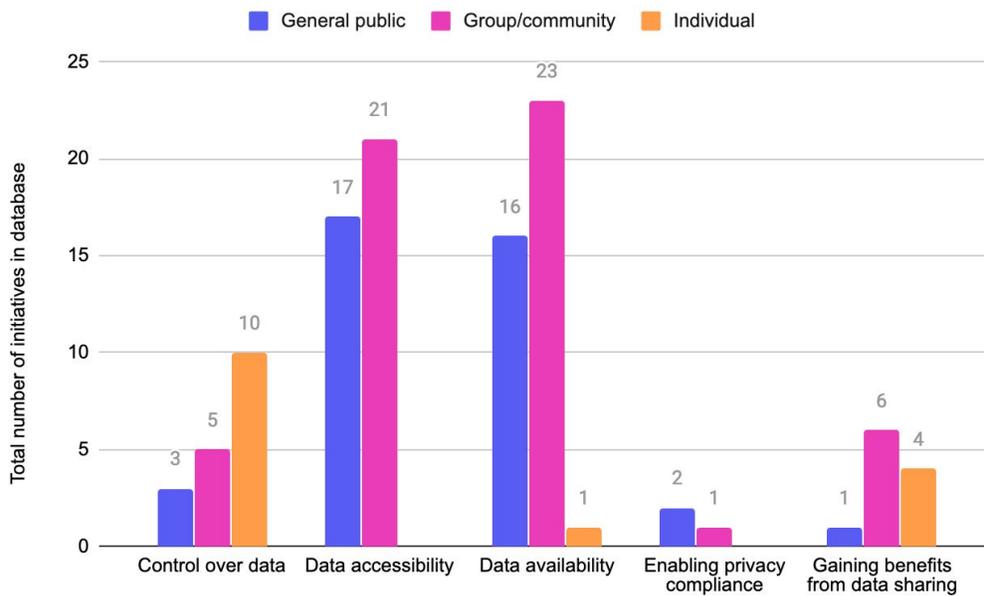
Figure 4: Primary benefit by type of primary beneficiary.

Correlating the stated primary benefit of initiatives in our database with the primary beneficiary type shows clear trends for each group (Figure 4). First, most initiatives that claim to benefit individuals offer more control over their personal data, or promise benefits from the sharing of data (e.g. by monetizing personal data). Second, most initiatives that claim to benefit the general public aim to either make data available or more accessible. Similarly, initiatives that primarily aim to benefit particular groups or communities mostly also aim to make data available or more accessible to them. Only to a lesser extent do they encourage data sharing or promise better means to control their data. The benefit 'Enabling privacy compliance' only had a very small number of initiatives, which makes it difficult to make general statements about what type of beneficiary is usually targeted by it.
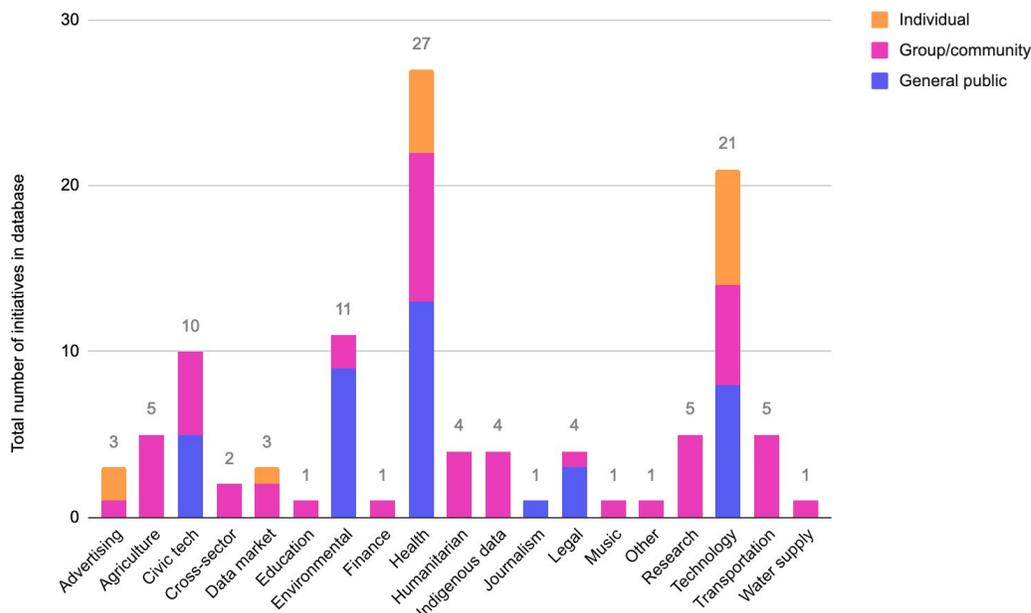
Figure 5: Primary beneficiaries by sector.

Correlating the primary type of beneficiary group with the sectors also shows a couple of interesting details (Figure 5). Of the two most common sectors in our database: most initiatives in the health sector aim to benefit the general public and particular groups or communities, while most initiatives in the technology sector aim to benefit individuals. Initiatives aiming to benefit groups or communities are the most diverse in terms of sectors, indicating that the idea of managing group data to benefit group interests is common across sectors.

Looking more closely into these correlations, we identified five 'types' of initiatives that differ in their arrangements of benefits, beneficiaries, and data sharing:

- **Data donation initiatives:** These act as mediators between data subjects and groups interested in their data. Typically, these are initiatives in the health sector that facilitate data donations by patients for medical research like Salus.coop, a non-profit data cooperative that aims to give users more control of their own health records while facilitating data sharing. Importantly, the data is donated without an expectation of remuneration.

- **Individual control initiatives:** These give individuals more control over their personal data. Mostly, these are technology focused initiatives that protect privacy or help individuals better control how their personal data is shared with various entities. Initiatives that enable individuals to monetize their personal data also fall into this category. Examples are digi.me, a smartphone app that puts

25

data from various services in one place, makes it searchable, offers analytics, and give users control over who can access it; or Aiisma, a 'data marketplace' that allows individuals to be "rewarded for consensually and anonymously sharing data points."

- **Group protection and empowerment initiatives:** These are dedicated to protecting the rights and interests of particular groups or communities, such as worker rights or Indigenous data sovereignty. Importantly, this is not necessarily about data rights, but can be about all kinds of human or collective rights being protected via data in some way. An example is the Driver's Seat Cooperative, a cooperative owned by ride-hail drivers that share data and profit from the revenue made by the cooperative.

- **Public-facing data collection initiatives:** These collect data and make it publicly available. These initiatives are often focused on social issues where data advocacy can drive change such as environmental, government transparency, and humanitarian issues. An example is the SaveEcoBot, which combines air quality data from public, non-governmental, and private institutions as well as individual citizens who donate data from sensors.

- **Data exchange networks:** These pool data held by numerous organizations that agree to create a shared resource. An example is the California Data Collaborative, a network of water professionals that provide members with access to cleaned and standardized data about water use.

The distinction between these five types is not always clear cut in practice. For example, the initiative Amazonia Socioambiental collaborates with several other groups to gather data about the endangered Amazon region and map it to highlight environmental destruction and support Indigenous groups. It could therefore be considered both 'Group protection and empowerment' and 'Public-facing data collection.' Moreover, these different types can overlap and intersect to represent different 'layers' within a project, for example 'Individual control' can play a role in 'Data donation' initiatives. Still, these five types provide a basic orientation that can be further refined using the key distinguishing features described above.

4.3 Regional trends: Is 'alternative data governance' a global phenomenon?

When the seven regional researchers began their work, most raised concerns about whether concepts such as data stewardship and data governance would resonate in

their countries, regions, and languages. English terms such as 'data stewards,' 'data cooperatives,' and 'data trusts' are difficult to translate and were in some cases difficult to explain to local digital rights or open data advocates whom they sought advice from.

As the researchers began sharing observations from conversations they had with actors in relevant sectors and communities, this hunch was validated. Based on desk research and interviews with members of the open data and civic technology communities, privacy and digital rights activists, as well as people associated with IT hubs and technology start up incubators/platforms, they identified a number of key contextual factors and trends across five regions – Eastern Europe, Sub-Saharan Africa, the Middle East and North Africa, Southeast Asia, and Latin America (including Brazil).

In joint conversations of the larger research group, a number of prerequisites for innovation around alternative data governance approaches emerged. These include the maturity of data privacy movements as well as data protection and data privacy laws, but also numerous cultural factors. For example, Indigenious communities might have different understandings of individual vs. collective. The existence (or lack of) local traditions for managing collective resources can also influence how frequent and understandable concepts around 'alternative data governance' are. Moreover, data collection by private companies is not always considered problematic, at least not in comparison with government data collection. These regional trends are explored in more detail below.

Eastern Europe

In Eastern Europe, discussion and practice addressing data governance and data privacy is happening within special interest and activist communities, but not outside of these circles. There appear to be more active open data movements in countries with more advanced data protection legislation. Generally, people express more concern about governments collecting their data than corporations, with the exception in some countries (due to geopolitics and conflicts) of Russian businesses collecting data.

Examples of interesting practices in the region include ecological data collection initiatives, such as SaveEcoBot in Ukraine (mentioned above). [Breathe.Moscow](#) is another example, which relies mostly on data supplied by individual contributors who install environmental sensors. There are impactful anti-corruption initiatives in both [Russia](#) and [Ukraine](#) where activists accumulate asset declarations of public officials and make them available online.

Notably, most of the collected initiatives do not discuss how they govern data. Moreover, they hardly even mention the word "data" on their websites. Within the

technology sector, profits and technological innovations are more of a priority than the agency of data subjects.

Sub-Saharan Africa

Concepts such as data stewardship, data cooperatives, data commons, or data trusts were largely non-existent in the surveyed countries (both francophone and anglophone). The majority of activity and discussion was around data protection legislation, with these efforts being led by national governments in most cases, with other groups playing an active role, such as law firms, civic tech activists, and other civil society organizations.

The majority of initiatives found in this region were public-facing data collection initiatives. However, there is a potential for other alternative data governance based on pre-existing organizational and cooperative structures. An example for what might be called a 'potentially nascent data governance initiative' is the [Kenya Tea Development Agency (KTDA)](), an organization owned by tea farmers which was founded in the 1960s. Among other things, it keeps track of how much tea is produced in what facilities in order to increase efficiency. In 2019, it introduced a 'smart card' to track this digitally. Thus far, data is collected solely for management and efficiency purposes, but as an organization owned by tea farmers, KTDA might adopt more elements of data trusts, data coops, or other forms of alternative data governance in the future.

Middle East and North Africa

In the MENA region, the lack of (strong) privacy and data protection laws would seem to inhibit innovation in data governance. Even when there are laws, they tend to focus on protecting consumer privacy in relation to corporate actors, even while government surveillance is rampant.

The focus of most internet startups is on delivering services for either commercial profit or social gains (education, culture, women's empowerment, etc.), and not on increasing agency and control over data for individuals and communities.

Among digital rights groups, while protecting privacy is a priority, 'data stewardship' is a term that is largely unheard of. Given the dire human rights situation in many of the region's countries, digital rights groups prioritize more urgent threats including government hacking and surveillance, trolls and propaganda, disinformation, threats to free expression, and online gender based violence.

Latin America

In Latin America, the main impulse around data governance initiatives that distribute the benefits of data in more democratic ways are largely coming out of the open data movement. Some of the barriers to more data commons or cooperative centric approaches to data governance could be that, though access to smartphones and the internet has grown in the region, e-commerce is still not as strong as in other regions and digital literacy may still not be widespread. There are few data governance projects related to privacy concerns despite a rich ecosystem of NGOs already working on privacy-related issues.

Established digital rights groups in the region like TEDIC in Paraguay and Fundación Karisma in Colombia conduct privacy campaigns and research, but no real governance products or tools or frameworks have emerged from either.

Most of the initiatives found in this region seek to make public data, and sometimes privately contributed data, more accessible to journalists, civil society, and sometimes citizens. Several of these initiatives include interactive mapping and data visualizations, or tools that make access to government services easier in order to expose corruption, demand better public services, or engage directly with the administration, to name just a few examples.

The initiatives come mostly from established organizations: local governments, national governments, NGOs, and civic tech organizations. There seem to be fewer citizen or community-led initiatives. Uruguay looks like a leader in the region in these efforts, though there are interesting initiatives also in Chile, Argentina, Perú, and Colombia. Key drivers seem to be the social issues outside of the digital economy, for which better access and use of data could be useful, like corruption or environmental degradation.

Brazil

Theoretical discussions around data and governance, especially as in relation to personal data, is still developing in Brazil. A data protection law was approved in 2018, but still has not been implemented, which means that constructs around data as a right, or data ownership and control are far from mainstream. Civil society groups and activist communities that emerged from the broader digital rights discussion are important players in advancing these discussions.

Brazil has a vibrant open data and open government community. At the same time, there's a juridical and cultural tradition around the role of government as the main

steward of public and collective goods. This might be the reason for the high number of initiatives led by the government. From the civil society standpoint, the main activities are open data/government for transparency and to fight corruption. From a government standpoint, the open data/open government initiatives aim for efficiency and innovation. Much of the idea of collective data ownership and control is derived from the open data movement.

Within the startup ecosystem, there is a relevant agritech movement that responds to the need to innovate to compete globally and is incentivized by the government. This movement normally produces proprietary technologies and databases. Ideas around data collectives are driven by universities or research initiatives aiming to develop new technologies and R&D in a closed or proprietary fashion. One example is Brazil's foodtech startup iFood, which builds an AI learning academy and research center (see Jacob Atkins 2019).

The idea of stewardship could be further researched within Indigenous communities that have different epistemologies of the individual vs. the collective. There is much literature and many policies about traditional knowledge in terms of both the management of this knowledge and economic exploitation.

Southeast Asia

The majority of initiatives found in Southeast Asia are established projects managed by non-business entities (e.g. NGOs, independent networks, or volunteer-run). In terms of accountability structures and systems, only a few detailed the mechanisms that hold the initiative accountable. For example, only one project clearly stated on their website that they have a governing board.

The main beneficiaries of these projects is an unspecified public. For example, the Malaysian project on 'COVID-19 Crisis: Atlas of Community Resilience' provides information for users to understand the impact of COVID-19 in the country. In contrast, some projects, such as the Automated Analytics System for Small-Scale Fisheries in Timor Leste, are targeted at specific groups of people such as fisheries officers and researchers.

A number of the projects collect personal data such as an individual's location, credit information, and demographic data. However, it is unclear who has rights to use the data and how. Perhaps this is not surprising as not all the surveyed countries have regulations or frameworks on data privacy and security. Only one project, MCIX, said "user data is and remains the property of users." Another potential reason has to do

with how personal data is understood, and not all the projects involve personal data (e.g. pmhaze.org).

Trends across regions

Based on the observations of regional researchers, it appears that only in Western Europe and North America (primarily in English) is there ample evidence of 'alternative data governance' activity. Moreover, initiatives that are active globally (rather than being tied to a specific region) are most often based in the US and Western Europe (less than 4% of those operating globally have their base in other regions).

Due to the limitations of this study, we must reiterate that this is not a comprehensive scan and that there may be projects that were missed due to our research design.
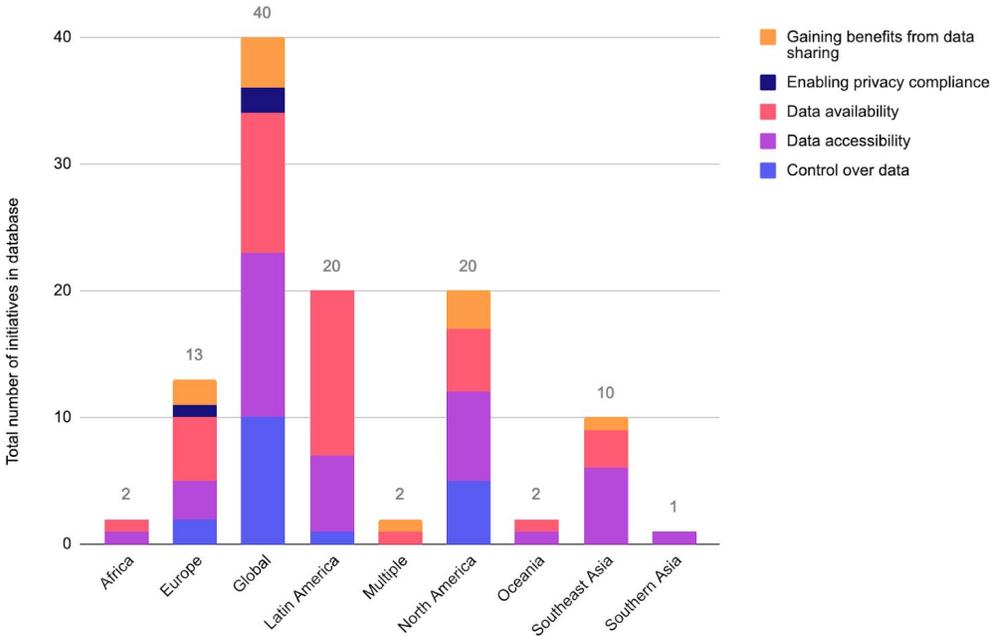

Figure 6: Primary benefit by region where initiatives offer their services.

Looking at the primary benefit of the initiative by region reflects an important trend identified by our regional researchers (Figure 6): In regions with lower levels of privacy and data protections and generally lower levels of data literacy, most of the few initiatives we collected in our database focus on making more data publicly available or accessible. This is at least in part because alternative approaches of data governance are primarily understood in terms of open data by the people interviewed by our regional researchers. Our database did not include any initiatives that claim to give more control over data in Africa or Southeast Asia, and only a few examples were

31

found in Latin America. Taking these findings together, it appears that alternative data governance is most prominently understood as a form of open data, where data is simply shared openly with the public rather than managed in various other ways specific to the interests of particular beneficiaries.

4.4 COVID-19 as a driver for alternative data governance?

At the time of writing, 24 initiatives in our database had an explicit connection to COVID-19. Nine of those were newly invented as a direct response to the pandemic: three are contact tracing apps that provide a privacy-preserving way for citizens to share data with health officials, the rest were mapping projects like the COVID-19 map from Tunisia, a project dedicated to making information and data around COVID-19 more accessible.

Thirteen of the COVID-19 related initiatives apply a previously developed approach to data governance to the COVID-19 situation. Most common are initiatives in the health sector that have developed privacy-preserving data sharing technologies for patients and medical researchers since before the pandemic started. An example is Salus.Coop, a non-profit data cooperative for health research founded in 2017 based in Spain. It developed its own data sharing licence and aims to give users more control of their own health records while facilitating data sharing to "accelerate research innovation in healthcare." Since the COVID-19 outbreak, Salus.Coop (2020) updated its manifesto to invite health authorities to collaborate to "develop licenses for use and technological architectures which encourage people to voluntarily participate in donating their complementary data."

Another example of reapplying earlier ideas to the COVID-19 pandemic is the Commons Prototype by MyData global, a non-profit organization that develops tools and standards to give individuals more control over their data (Iain Henderson 2020) The Commons Prototype promotes and implements standardized tools that give individuals easier controls to share their data with health authorities or others (see above). According to MyData, this project is rooted in ideas dating back as far as 2007 (Iain Henderson 2020) MyData has always been focused on giving individuals more control over their personal data, e.g. by advocating for interoperability and data portability. The blog post presenting the new Commons Prototype calls COVID-19 an "inspiration/driver" that helped to further develop, refine, and implement some of MyData's core principles: "COVID-19 has given us a giant, data-intense use case, the time to study it in detail, and the incentive to move at pace."

Another variation is the Idaho Health Data Exchange, a network of medical institutions that exchange data about patient care (see list of primary benefits above). This initiative responded to the pandemic by allowing broader access to the data of its (paying) members as a way to contribute to the fight against COVID-19 (Idaho Health Data Exchange 2020). Here, an initiative modified its data governance model in response to the pandemic. Finally, some initiatives simply provided a dedicated info page on COVID-19, but otherwise have no direct connection to the pandemic or modification of their data governance approach in response to it.

Thus far, we do not see evidence that the COVID-19 pandemic has been a major driver for the invention of new data governance approaches, but it significantly contributed to innovation that makes use of pre-existing approaches. Around the globe, COVID-19 has sparked discussions about the right balance between privacy and public health that accelerated the development of decentralized contract-tracing technologies in various regions (like [PEPP-PT](#) in Europe). MyData say their Commons Prototype is not about reinvention, but about finding ways to better scale solutions that implement core beliefs and ideas (Iain Henderson 2020). Similar sentiments about the increased relevance of their own core beliefs and the need to better scale them in the pandemic are echoed by several other initiatives in our database.

Apple and Google's joint initiative for an Exposure Notification system in service of privacy-preserving contact tracing (Apple 2020) further illustrates how the pandemic is paving the way for making pre-existing ideas and visions for alternative data governance more scalable. While the initiative raises important concerns around privacy and the power of big tech companies (Michael Veale 2020), it implemented a decentralized approach to the collection and use of data that is also advocated by several alternative data governance initiatives in our database.

4.5 Different understandings of 'alternative data governance'

Our survey responses, the findings by regional researchers, and the contents of other databases all suggest that 'alternative data governance' is often understood in broader terms than in some theoretical literature.

We can roughly distinguish three understandings:

- **Governing data itself:** Arguably the 'proper' definition that forms the basis for theoretical discussions about data governance and was our starting point. Initiatives in this category are primarily about creating data resources as well as managing access and value extraction of data.

- **Affecting governance via data:** In submissions to our survey, as well as among professionals our regional researchers contacted, 'data governance' is often understood as affecting governance via data, for example by creating alternatives to official data and publishing it online (Gray, Lämmerhirt, and Bounegru 2016). The core concern is not about governing access to or sharing data in particular ways.

- **Enabling ways of handling data:** Many submissions to our survey were about enabling certain ways of handling data via technology (tools or data standards), funding, or advocacy. Instead of prescribing particular modes of data governance, these initiatives nevertheless do facilitate particular data governance approaches.

Our findings suggest that even among digital rights savvy professionals worldwide, 'alternative data governance' is mostly understood as 'affecting governance via data.' This is evidenced in the fact that public-facing data collection initiatives to improve or create alternatives to official data were the most common examples of 'alternative data governance' identified in regions outside of North America and Western Europe. Going forward, these findings illustrate the need to 1) make actual alternative data governance approaches more widely known, and 2) clarify how these two common understandings of data governance differ and potentially complement each other.

5. Looking ahead – questions and considerations for future research

5.1 Correlating types and distinguishing features with governance approaches

As noted earlier, many of the initiatives in our database offer limited information publicly around governance decisions or what mechanisms and systems exist to ensure accountability. Therefore, a priority in the next stage of research could be to follow up with specific initiatives to better understand internal governance and decision making. In addition, correlating the distinguishing features and types we identified in our database with the various data governance approaches described in theoretical literature could be helpful in highlighting and understanding the gap between theory and practice. Exploring the relationship between data governance and the different stages along the twelve steps of the Data Value Chain (identify, collect, process, analyze, release, disseminate, connect, incentivize, influence, use, change, and reuse; see Open Data Watch 2018) could also provide a helpful layer of understanding.

5.2 The ecosystem of organizations supporting innovation in data governance

Through this research, we identified dozens of organizations who play a supportive role in the global ecosystem of data governance innovation. As noted previously, identifying ecosystem actors was not the focus of this research and these organizations by no means represent the full ecosystem of support for data governance initiatives. However, insights from a rapid analysis of these organizations could be a helpful starting point for future research.

These organizations play a number of different roles in supporting the ecosystem for innovation, with some playing multiple roles:

- Funding initiatives, such as Nesta in the UK and Knight Foundation.

- Publishing research and thought leadership, such as the Open Data Institute and the Centre for AI and Data Governance in Singapore.

- International conveners, such as the Future of Privacy Forum and Cities for Digital Rights.

- Capacity builders across the ecosystem, such as [Esoko](#) in Africa, [Local Contexts](#) defending data rights of Native, First Nations, Aboriginal, and Indigenous communities, the [Jordan Open Source Association](#), or [Data Union](#).

- Advocates for change in policy and practice, such as [Data for Black Lives](#), [Te Mana Raraunga](#) in New Zealand, [Privacy International](#), and the [Eticas Foundation](#).

- Proponents of data standards and principles for best practice, such as [360 Giving](#) in the UK, the [Global Alliance for Genomics and Health](#), and [Data Commons Blueprint](#) in New Zealand.

- Developers of technological infrastructure that enable alternative approaches to data governance, such as [OpenMined](#).

- Organizations acting in multiple capacities, such as the [Mozilla Foundation](#) and [GovLab](#).

As a next step, understanding a more complete picture of the support ecosystem that exists around innovation in data governance should be a priority. As the field matures, many for-profit consultancies are currently moving into this space (from large consulting groups like Accenture/Deloitte/BCG, to smaller technology studios like Projects by If). It could be interesting to study possible tensions or new dynamics that this creates. Government funding (in addition to government-led initiatives identified through this research) may also be useful to call out specifically as key participants in the ecosystem. For example, governments in Australia, Finland, and the UK, fund a number of individual control initiatives.

It may also be worth investigating whether there is a correlation between regions with fewer supporting entities, including privacy and digital rights groups, and fewer data governance initiatives. In related fields such as digital security, the presence of regional hubs (such as Ukraine's Digital Security Lab) has fostered a growing community of practice.

The role of organizations who develop data standards such as the [Beneficial Ownership Data Standard](#) or the [Open Contracting data standard](#) to help facilitate innovation in data governance is also worth exploring in more detail, as well as interrogating the role of national governments and other intermediaries such as the World Bank in supporting these efforts.

It could be interesting to understand how supporting entities to alternative data governance map onto existing frameworks such as the Deloitte Center for Government Insights' "Five roles in public sector innovation," which highlights five key roles in

innovation ecosystems: Problem solvers, Enablers, Conveners, Motivators, and Integrators (Alan Holden et al. 2017).

5.3 The cultural underpinnings of the data government ecosystem

Much of the theoretical literature is about the legal mechanisms and systems of accountability. However, the findings of our regional researchers illustrate the importance of cultural and social factors that influence the understanding and ecosystem of alternative data governance. Research to investigate epistemological questions of what gets highlighted (e.g. by drawing from feminist data practices) as well as critical examinations of the values and imaginaries that drive founders, decision-makers, and particular communities of alternative data governance initiatives would be useful. Relatedly, the alignment of data governance approaches with political philosophy and how this influences regional differences would be an important addition (e.g. in Europe, individual data initiatives often seem to align with free-market attitudes while data trust/commons often align with leftist positions).

5.4 Other relationships to explore in future research

- The legal/regulatory environment (and whether the innovative data governance approaches create new needs)
- Business models
- Barriers to growth
- Dynamics between big (scalable) and small initiatives (rather than consumer and private vs. public interest tech)

Acknowledgements

References

- Aapti Institute. 2020. "Data Stewardship Explorer." The Data Economy Lab. June 12, 2020. https://thedataeconomylab.com/explorer/.
- Alan Holden, Sonora Braun, Lilian Lee, Abigail Phelps, and Rachel Samuelson. 2017. "Catalyzing Public Sector Innovation." Deloitte Insights. https://www2.deloitte.com/us/en/insights/industry/public-sector/catalyzing-public-sector-innovation-initiatives.html.
- Anouk Ruhaak. 2020. "Data Commons - Projects." Airtable. 2020. https://airtable.com/shrcX38kvQVKpxEQL/tbl5czc9lofWxdxlk.
- Apple. 2020. "Privacy-Preserving Contact Tracing - Apple and Google." April 10, 2020. https://www.apple.com/covid19/contacttracing.
- GovLab. 2020. "Data Collaboratives Explorer." Data Collaboratives. 2020. http://datacollaboratives.org/explorer.html.
- Gray, Jonathan, Danny Lämmerhirt, and Liliana Bounegru. 2016. "Changing What Counts: How Can Citizen-Generated and Civil Society Data Be Used as an Advocacy Tool to Change Official Data Collection?" http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2742871.
- Iain Henderson. 2020. "Guest Blog Post – The MyData Commons Prototype." MyData Global Blog (blog). June 2, 2020. https://mydata.org/2020/06/02/the-mydata-commons-prototype/.
- Idaho Health Data Exchange. 2020. "Ihde Statement on Covid-19." March 17, 2020. https://www.idahohde.org/alerts/ihde-statement-on-covid-19/.
- Jacob Atkins. 2019. "IFood, Brazil's Largest Delivery App, Funds New AI Learning Center." Contxto (blog). April 10, 2019. https://www.contxto.com/en/brazil/ifood-brazils-largest-delivery-app-funds-new-ai-learning-center/.
- Michael Veale. 2020. "Privacy is not the problem with the Apple-Google contact-tracing toolkit." The Guardian. July 1, 2020. https://www.theguardian.com/commentisfree/2020/jul/01/apple-google-contact-tracing-app-tech-giant-digital-rights.
- Mills, Stuart. 2019. "Who Owns the Future? Data Trusts, Data Commons, and the Future of Data Ownership." SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3437936.
- Open Data Watch. 2018. "The Data Value Chain: Moving from Production to Impact." Data2X. https://opendatawatch.com/publications/the-data-value-chain-moving-from-production-to-impact/.

- Verhulst, Stefaan, Andrew Young, and Prianka Srinivasan. 2017. "An Introduction to Data Collaboratives." https://datacollaboratives.org/static/files/data-collaboratives-intro.pdf.