



**People-Centered
Internet**

2024

Data Cooperatives Report

**Launched at UN Science Summit Digital
Governance Series, September 16, 2024**

First Edition By

People-Centered Internet

Preface

In the digital age, data is both a valuable resource and a sensitive asset. As scientific research increasingly depends on vast quantities of data, the need for clear rights and governance around this data becomes critical. Yet, today, the rights to data—whether belonging to individuals, families, communities, or organizations—are not clearly established. This uncertainty can create hesitation for those who might otherwise contribute valuable data to scientific endeavors. Without safeguards, many fear losing control over their data or having it used in ways that do not align with their intentions.

This report on **Data Cooperatives** addresses these challenges head-on, offering a framework for collective data governance that empowers communities while protecting individual rights. One of the core chapters ([link](#)) in this report focuses on the role of **Digital Asset Registries (DARs)**, which we have contributed. DARs provide a mechanism for defining, managing, and securing the rights of data owners in the digital era. By establishing clear ownership and access rights, DARs ensure that data can be shared and used for scientific research without undermining the rights of those who contribute it.

For science to thrive, it is essential that people feel safe contributing their data to research efforts. DARs make this possible by guaranteeing that contributors retain control over their data and that their rights will not be lost in the process. In many ways, this is about building trust—trust that data will be used ethically, and that contributors will be recognized and respected for their participation.

Data Cooperatives, as outlined in this report, offer a collective model where data can be pooled and used for the benefit of all, while still respecting individual and community rights. They enable a more democratic and inclusive approach to data governance, ensuring that the benefits of data-driven science are shared widely. By enabling secure, structured data sharing, these cooperatives can unlock new opportunities for research and innovation across fields ranging from healthcare to environmental science.

As we move forward, the integration of DARs within Data Cooperatives provides a vital foundation for scientific research in the digital age. These systems ensure that data rights are not only protected but also actively managed in a way that fosters collaboration and collective progress. By contributing to this report, we hope to offer a roadmap for a future where data is a shared resource, governed with transparency, and used to advance people's individual well-being and prosperity as well as the common good.

Jascha Stein and Mei Lin Fung

Co-chairs of the People-Centered Internet, co-founded with Vint Cerf

Foreword by Initiator and Curator

This report on data cooperatives is a direct follow-up to a pivotal panel held during the 9th edition of the Science Summit at the 78th United Nations General Assembly (UNGA78). The panel was part of the session on Digital Democracy for Climate Action, which took place at the Harvard Club in New York City on September 18, 2023. Bringing together a diverse group of 40 esteemed speakers, including Carlos Nobre - a Nobel laureate - and representatives from the Amazonian Puyanawa tribe, the event was convened by Rob Eisenberg, Andrea Soccorso, Indrani Pal-Chaudhuri, and myself, Gael Van Weyenbergh.

Our session explored the interplay between civic participation, indigeneity, conscious capitalism, and data cooperatives in addressing climate change. One of our key objectives was to investigate the potential of data cooperatives and digital commons as innovative tools for mobilizing collective action.

Our panel featured a remarkable lineup of contributors: Eugene Braverock, a Native American renowned film actor and advocate for indigenous language preservation; Aungh Chabalala, Director of the Department of Science and Innovation at the South African government, with expertise in Indigenous Knowledge Systems; Morshed Mannan, Research Fellow at the Robert Schuman Centre for Advanced Studies at the European University Institute, specializing in cooperative governance; Kelly Achenbach, Communication Officer at the Max Weber Foundation in Germany, with a focus on citizen science; Hossein Rahnama, a visiting professor at the MIT Media Lab, professor at Ryerson University, founder of a data intelligence company, and recognized as one of the world's top young innovators; Mei Lin Fung, Chair of Sustainability IEEE SSIT, Co-Chair of the People-Centered Internet, and Co-Chair of the Expert Roundtable for Asia, Africa, Europe and the Middle East which provided the basis for the 2023, UN Commission on the Status of Women Report on Digital Innovation; Inder Monga, Director of Scientific Networking and Executive Director of ESnet at the US Department of Energy's Lawrence Berkeley National Labs, known for global collaboration strategies; Gurvinder Ahluwalia, Founder and CEO of Digital Twin Labs and previously CTO of IBM North America, often called the father of live video streaming; Navroop Sadhev, CEO of The Digital Economist, an advocate for a human-centered global economy; and Stephanie Carroll, Director of the Collaboratory for Indigenous Data Governance, who develops research, policy, and practice innovations for Indigenous Data Sovereignty.

The discussion sparked a lively conversation around the potential of data cooperatives, resulting in a surge of interest that led us to invite each panelist to contribute a two-page paper. We also reached out to additional experts to broaden the scope, resulting in this report.

The relevance of data cooperatives in the digital age cannot be overstated. They represent an evolution of cooperative principles, much like social cooperatives and cooperative banks did in the previous century. By addressing the pressing question of who controls data flows,

data cooperatives tackle the governance of digital platforms while also striving to ensure that the benefits of digital transformation are shared more equitably among a broad range of stakeholders. This approach is key to developing digital public infrastructures that are inclusive and fair. Their multistakeholder nature makes them a valuable tool for involving civil society in defining and steering the course of digital transformation.

By bringing various stakeholders together and enabling shared insights through the pooling of data, data cooperatives have the potential to drive a new wave of innovation. For those who have pioneered the internet, this is seen as an opportunity for a renaissance, where the true transformative impact of digital technology is yet to unfold. Beyond their use in climate action, data cooperatives offer an adaptable framework for complex problem-solving across various sectors, such as SME financing, mobility, healthcare, farming, supply chains, and indigenous knowledge systems.

However, data cooperatives are still emerging, facing challenges such as legal uncertainties, regulatory complexities, and operational hurdles. Despite these obstacles, they are propelled forward by several key trends. First, the commercial internet has shown its flaws—particularly in spreading disinformation, fostering digital addiction, and polarizing society—creating a push toward alternative models. Second, the evolution from Web 1.0 to Web 3.0 has shifted the conversation from privacy rights to ownership rights, empowering individuals and organizations to gain more control over their data. Third, institutional support is growing, as evidenced by the Digital Governance Act of 2022, which recognizes data cooperatives as a new type of data intermediary. This is further bolstered by international recognition from initiatives like the Global Digital Compact, which advocates for a human-rights-based digital transformation, along with the G7 and G20's emphasis on the advantages of stakeholder-controlled data.

Looking back at historical precedents, two key examples come to mind. The printing press, initially controlled by a select few, eventually became widespread, transforming the way knowledge was disseminated. Data cooperatives could facilitate a similar transformation in the digital realm. Likewise, the exploration of new frontiers, first driven by exploitation and profit, eventually came under the rule of law—a process we are now witnessing in the digital domain. In this context, data cooperatives represent a significant shift in mindset. They move away from an industrial approach that often overlooks societal impact, toward a holistic model that considers the entire ecosystem of stakeholders. This shift mirrors the evolution of industrial food production into more organic and sustainable models, offering the possibility of a healthier information ecosystem.

Despite the transformative power of the internet, it presents paradoxes and limitations. While it has identified problems, solutions, and resources and connected people who are ready to engage, it has fallen short in facilitating collective action at scale, thus hindering our ability to solve problems collectively. Additionally, while the internet has democratized information, it has also eroded public discourse and weakened the foundations of a shared reality. It has connected everyone yet led to the atomization of relationships and polarization of ideas, resulting in a pervasive sense of isolation and a lack of trust in both individuals and institutions. Data, as the bedrock of the digital economy, remains poorly regulated, leaving users inadequately protected.

Data cooperatives, as a 21st-century solution, offer a viable path forward, allowing us to address the shortcomings of the commercial internet without sidelining the private sector. On the societal front, they provide a flexible framework for complex problem-solving, grounding our digital experiences in ethics, the rule of law, and human rights. This approach better protects users' privacy while promoting a healthier, more cooperative digital ecosystem.

The aim of this report is to contribute to the growing body of literature on data cooperatives, inspire a shift in mindset, and rekindle the original spirit of the internet—not just as a means of communication, but as a global platform for cooperation.

Gael Van Weyenbergh

Foreword by Managing Editor

Data is exceptional: while a small amount of data may have little value, when aggregated, its worth multiplies far beyond the sum of its individual parts. This characteristic makes data a commodity that requires a connecting institutional framework. The contributors to this Data Cooperative Report have provided a compelling framework that reminds me, as a German economist, of a concept that played a key role in the rapid economic recovery of post-war Germany. Germany had to build entirely new, stable structures to counteract excessive concentration of power. The social market economy (thought leaders Walter Eucken, Alfred Müller-Armack, Friedrich August von Hayek and others) provided stability on a large scale by enabling people to act in a self-determined way through an institutional structure based on freedom and protection of property. Today, in the digital age, we face a similar challenge worldwide: the need to build entirely new, stable structures to maximize the potential of our new treasure: data.

Data-collecting companies, as Hossein Rahnama describes, often struggle with internal tensions over customer data. By pooling the capacities of many individuals into member-owned, fiduciary institutions, data cooperatives—introduced into public discourse by Alex Pentland and Thomas Hardjono—offer an innovative counterbalance. These cooperatives, distinct from data commons and data collaboratives (see Banifatemi), are structured to manage the collection, processing, and access to their members' data.

To unlock the full potential of data cooperatives, they need to be embedded in a digital public infrastructure, Sarthak Satapathy explains. Stein et al. emphasize the need to establish a Prosperity Data Network (PDN) that connects data cooperatives for knowledge and data transfer. Fung et al. discuss Digital Asset Registries (DARs) as a means of creating property rights in the digital sphere. Such a networked approach offers flexibility and accommodates regional and cultural differences, thereby fostering acceptance – more effectively than current western data protection in Africa, as Shabaya et al. point out.

In data cooperatives, members retain the power to decide how their data is used, leading to greater empowerment and self-determination (see Rai). This model can liberate various social groups, including youth and Indigenous communities. Eugene Brave Rock, Natalie Tercova, and Izaan Khan further elaborate on how this approach could shape a more equitable digital future for these groups.

Data cooperatives are not only a strategy for navigating the challenges of digitization; they are also a tool for fostering digital transformation in sectors where implementation has been difficult. Less digitized sectors often contain a high concentration of small and medium-sized enterprises (SMEs). Data cooperatives can help these sectors to overcome their coordination challenges. Bühler et al. introduce the German bauform cooperative for construction firms. In addition to this, another hurdle for digitalization lies in highly sensitive data, such as in healthcare. Creus presents salud.coop, a Spanish health data cooperative that empowers individuals to control their health information.

It is clear from this report that we have the necessary ideas to build new structures for the digital age. Kalinina and Seaman highlight the growing number of initiatives. A seed has been sown; now it needs to be watered.

Christine Asjoma

Table of Content

Preface	2
Jascha Stein and Mei Lin Fung	2
Foreword by Initiator and Curator	3
Gael Van Weyenbergh	5
Foreword by Managing Editor	6
Christine Asjoma	7
Table of Content	8
Data Cooperatives	8
Alex S. Pentland and Thomas Hardjono	8
A Scalable Framework for the formation of Data Alliances and Co-operatives	10
Hossein Rahnama	10
Data Cooperatives: Aligning Data Governance with Human Values	13
Mandeep Rai	13
Towards an Equitable Data Future: Data Cooperatives x Digital Public Infrastructure	15
Sarthak Satapathy	15
Improving Global Governance: Data Cooperatives for Global Cooperation	18
Jascha Stein, Mei Lin Fung, Marta Bertolaso, Fulvio Bersanetti, Suruchi Gupta, Ronald Strauß, Charlie Isaacs, Christine Asjoma	18
Data Cooperatives to Empower SMEs in the Construction Sector	28
Michael Bühler, Konrad Nübel, Thorsten Jelinek, David Riechert, Lothar Köhler, Pia Hollenbach	28
Salus.coop, data cooperative for health research.	32
Javier Creus	32
Data Cooperatives: A Case for Collective Participatory Governance	35
Anastasia Kalinina and Alexandra Seaman	35
Legal and Regulatory Frameworks for Data Cooperatives in the EU and beyond	39
Dr. Morshed Mannan	39
Harnessing AI Potentials: from Data cooperatives to AI commons	43
Amir Banifatemi	43
Exploring the Potential of Data Cooperatives in Preserving Indigenous Languages through the Oki Language Project	47
Eugene Brave Rock	47
Implementing Resilience in Scientific Research through Digital Asset Registries (DARs) and Data Cooperatives	49
Mei Lin Fung, Jascha Stein, Sean Masaki Flynn	49

An Understanding of Data Cooperatives When Viewed from the Lens of the 2023 YOUthDIG Messages	54
Izaan Khan	54
Fostering an African Data Commons	56
Marie Shabaya, Simon Sun and Nai Lee Kalema	56
Digital Horizons: Empowering Youth in the World of Access, Literacy, and Opportunities	59
Natálie Terčová	59
List of Publications on Data Cooperatives	61

Why Data Cooperatives

Data Cooperatives

Alex S. Pentland and Thomas Hardjono

MIT Media Lab

During the last decade, all segments of society have become increasingly alarmed by the amount of data, and resulting power, held by a small number of actors. Some famously call data “the new oil,” and they mean records of the behavior of citizens. Why, then, is control of this powerful new resource concentrated in so few hands? During the last 150 years, questions about concentration of power have emerged each time the economy has shifted to a new paradigm.

As the economy was transformed by industrialization and then by consumer banking, citizens felt trapped and exploited by powerful new companies. In order to provide a counterweight to these new powers, citizens joined together to form trade unions and cooperative banking institutions. Eventually, the struggle reached a point where citizens felt that powerful players such as Standard Oil, J. P. Morgan, and a handful of others threatened freedom itself, and the federal government instituted antitrust laws, labor rights, and banking reforms. The citizen organizations were key in helping to balance the economic and social power between large and small players and between employers and workers.

Today, the same sorts of citizen organizations can help us move from the current paradigm of individuals giving up data to large organizations to a system based on collective rights and accountability, with legal standards upheld by a new class of representatives, who act as fiduciaries for their members. There are many examples of community organizations using community data to manage investments for the good of the community. For instance, since 1943, the National Rural Electric Cooperative Association has electrified communities that cover over half the land area of the United States. Similarly, there are over 1,100 community development financial institutions, mainly small banks and credit unions, investing over \$220 billion in community projects, including over 300 focused primarily on economic, social, and political justice. With 100 million people being members of credit unions, the opportunity for community organizations to leverage community-owned data is huge.

Indeed, with advanced computing technologies, it is practical to automatically record and organize all the data that citizens knowingly or unknowingly give to companies and the government and store this data in community organizations’ vaults. In addition, almost all these community organizations already manage their accounts through regional associations that use common software, so widespread deployment of data cooperatives could be surprisingly quick and easy.

The notion of a data cooperative refers to the voluntary collaborative pooling by individuals of their personal data for the benefit of the members of the group or community. The motivation for individuals to get together and pool their data is driven by the need to share common insights across data that would otherwise be siloed or inaccessible. These insights provide the cooperative members as a whole with a better understanding of their current economic, health, and social conditions as compared to the other members of the cooperative in general.

It is technically straightforward to have a third party such as a cooperative hold copies of their members' data in order to help them safeguard their rights, represent them in negotiating how their data is used, alert them to how they are being surveilled, and audit the large companies and government institutions using their members' data. The creation of such data cooperatives also does not require new laws; many community organizations are already chartered to manage members' personal information for them. It does, however, require new regulations and oversight, similar to how the government regulates and provides oversight of financial institutions.

It is critical to note that community organizations that manage members' data must have fiduciary responsibilities to protect the sensitive information that they hold for members, as this is a central element in bringing data rights to the membership. This enables members to improve privacy and transparency regarding data use and empowers members to collectively direct the use of their data to their benefit.

Who will lead this historic and necessary transformation? The answer could well grow out of current-day credit unions, many of which are directly associated with universities, city governments, trade unions, and the like. They are already chartered to represent their members in financial transactions and hold members' data for them.

The ability to balance the world's data economy depends on creating a balance of stakeholders. Today, citizens and workers have no direct representation at the negotiating table, so they lose out. By leveraging cooperative worker and citizen organizations, which in the United States alone have over 100 million members, we can change this situation and create a sustainable digital economy that serves the many, not just the few. The power of 100 million US consumers who control their data would be a force to be reckoned with by all organizations that use citizen data and would be a very decisive way to hold these organizations accountable. The same potential for community organizations to balance today's data monoliths exists in most countries around the world.

What new advantages can communities have if they have the ability to analyze their data? People often think of monetizing personal data, but the reality is that while there is a great deal of value in aggregate data for specific purposes, there is no market mechanism for data exchange, so personal data does not have very much value on an individual basis. Personal data and community data will only become a serious source of revenue when privacy-respecting data exchanges become a major part of the general financial and economic landscape. However, monetization is only a minor part of data's value to a community, especially in today's economic climate. A greater source of value is in improving the living conditions of community members and ensuring the success of future generations. For instance, the COVID-19 pandemic has highlighted major disparities in public health between different communities. Data about community public health is necessary to address these disparities, but today that data is unavailable to communities in all but the most general terms.

A Scalable Framework for the formation of Data Alliances and Co-operatives

Hossein Rahnama

MIT Media Lab

After decades of inconsistent data sharing policies, consumer mistrust, government action, and competition for customers are bringing in a new era. Firms that generate any value from personal data will need to change the way they acquire it, share it, protect it, and profit from it. They should follow three basic rules:

1. consistently cultivate trust with customers, explaining in common-sense terms how their data is being used and what's in it for them;
2. focus on extracting insight, not personal identifiable information; and
3. companies should work together to facilitate the flow of insights, with a common objective of acquiring maximum insight from consented data for the customer's benefit.

For most of its existence, the data economy was structured around a “digital curtain” designed to obscure the industry's practices from lawmakers and the public. Data was considered company property and a proprietary secret, even though the data originated from customers' private behavior. That curtain has since been lifted and a convergence of consumer, government, and market forces are now giving users more control over the data they generate. Instead of serving as a resource that can be freely harvested, countries in every region of the world have begun to treat personal data as an asset owned by individuals and held in trust by firms.

This will be a far better organizing principle for the data economy. Giving individuals more control has the potential to curtail the sector's worst excesses while generating a new wave of customer-driven innovation, as customers begin to express what sort of personalization and opportunity they want their data to enable. Leading firms are already adapting to the new reality as it unfolds. The key to this transition — based upon our research on data and trust, and our experience working on this issue with a wide variety of firms — is for companies to reorganize their data operations around the new fundamental rules of consent, insight, and flow.

Converging Forces that necessitate the formation of trusted data cooperatives

We see three distinct pressures currently driving change in the personal data industry. All three are quickly becoming widespread and intertwined, causing seismic ripples across the sector.

1. Consumer mistrust. The idea of “surveillance capitalism,” which its author Shoshana Zuboff describes as “an economic system built on the secret extraction and

manipulation of human data,” has become common coinage, capturing consumers’ increasing awareness that their data is bought, sold, and used without their consent — and their growing reluctance to put up with it.

2. Government action. Federal lawmakers are moving to curtail the power of big tech. Meanwhile, in 2021 state legislatures proposed or passed at least 27 online privacy bills, regulating data markets and protecting personal digital rights.
3. Market competition. Last year, Apple’s upgrade to its iPhone operating system allowed users to shut down data harvesters’ ability to track them across their many apps. It was a refreshing change, providing customers with power and agency over their data.

This is a remarkable confluence of forces, and they are converging towards a clear endpoint where individuals will soon exercise full control over their personal data. While consumers still seek the conveniences and benefits that flow from their data, they will be the ones to set the terms over what data they share and who they share it with. People want that protection, governments have their backs, and technology firms are already falling in line, with competition over data privacy now impacting financial bottom lines.

Challenges Ahead for Large Firms

For established companies, these changes present a new set of data challenges on top of the ones they already have. Most large firms already suffer from a series of internal tensions over customer data. They typically have a Chief Information Officer whose role is to keep data in: collect it, encrypt it, and secure it from hackers. They also have a Chief Digital Officer whose role is to push data out: mine it, model it, and use it to entice users. Some have also added Chief Data Officers — a notably unstable position due, unsurprisingly, to lack of definition for the job — as well as Chief Information Security Officers and Chief Privacy Officers. All these overlapping roles are embedded in organizations with expansive data collection operations, multiple legacy systems, a complex web of bilateral and multilateral data-sharing agreements and, quite often, an ongoing lack of clarity on how to integrate data into their businesses. Based on our experience, up to 90 percent of current IT budgets are allocated to managing internal complexities, leaving very little funding for data innovation that improves either productivity or the customer experience. The new data economy won’t tolerate this state of affairs for long. If your organization generates any value from personal data, you will need to change the way you acquire it, share it, protect it and profit from it.

The New Rules of Forming Data Cooperatives

Our new rules of the forming decentralized and trustworthy data economy are fairly straightforward, all of them derived from the basic principle that personal data is an asset held by the people who generate it. But each rule entails the breaking of entrenched habits, routines and networks.

Rule 1: Trust over transactions. This first rule is all about consent. Until now, companies

have been gathering as much data as possible on their current and prospective customers' preferences, habits, and identities, transaction by transaction — often without customers understanding what is happening. But with the shift towards customer control, data collected with meaningful consent will soon be the most valuable data of all, because that's the only data companies will be permitted to act upon. Firms need to consistently cultivate trust with customers, explaining in common-sense terms how their data is being used and what's in it for them. Firms can follow the lead of recently-created data cooperatives, which provide users with different options for data sharing and secure each user's consent for the option they are most comfortable with. The more robust and thorough your consent practices are, the more valuable your customer database becomes.

Rule 2: Insight over identity. Firms need to re-think not only how they acquire data from their customers but from each other as well. Currently, companies routinely transfer large amounts of personal identifiable information (PII) through a complex web of data agreements, compromising both privacy and security. But today's technology — particularly federated learning and cooperative trust networks — makes it possible to acquire insight from data without acquiring or transferring the data itself. The co-design of algorithms and data can facilitate the process of insight extraction by structuring each to better meet the needs of the other. As a result, rather than moving data around, the algorithms exchange non-identifying statistics instead.

Rule 3: Flows over silos. This last rule flows from the first two, and doubles as a new organizing principle for internal data teams. Once all your customer data has meaningful consent and you are acquiring insight without transferring data, CIOs and CDOs no longer need to work in silos, with one trying to keep data locked up while the other is trying to break it out. Instead, CIOs and CDOs can work together to facilitate the flow of insights, with a common objective of acquiring maximum insight from consented data for the customer's benefit.

The Data-Sharing Future using Trusted Data Cooperatives/Alliances

New companies are already springing up to provide the structures needed to facilitate these kinds of data-sharing arrangements with data co-operatives. The emergence of data representatives, agents, and custodians make it possible to manage consent at scale, serving as trusted hubs for users' personal data and acting as their user agent in the marketplace. Data cooperatives are becoming common in some parts of the United States. The end of the old personal data economy will not spell the end of its value creation and wealth generation; that wealth will just be distributed better and more equitably, and carry fewer privacy and security risks. People will not hoard their data assets. Instead, they'll invest them in companies that provide them with a return in the form of more and better personalized services. They may even allow those companies to share insights drawn from their data — provided the benefits accrue to them.

Data Cooperatives: Aligning Data Governance with Human Values

Mandeep Rai

In the modern digital climate, the significance of data as a catalyst for innovation, democratic structures, and economic progress cannot be overstated. Nevertheless, concerns about privacy, security, and equitable distribution of benefits have arisen due to the centralised control and exploitation of user data by the world's tech giants. In response to these growing challenges, the concept of data cooperatives has emerged, aiming to explore the ethical dimensions of collaborative data governance and its potential to empower individuals and communities.

Data cooperatives represent collaborative initiatives wherein individuals or organisations voluntarily share, manage, and govern their data for the collective benefit of participants. The pooling of data resources within a cooperative framework enhances bargaining power, provides access to superior services, and allows members to shape data-driven systems in alignment with their values.

Travelling to more than 180 countries in my career, I have witnessed the best of humanity and the virtues that can contribute to our collective wellbeing and evolution. This journey forms the basis of my book, "The Values Compass: What 101 Countries Can Teach Us About Purpose, Life, and Leadership." The book serves as a guide, helping readers to discover their core values, recognise the influences and lessons that shape their worldview, and make deliberate life choices in creating a legacy aligned with their true aspirations.

In addition to personal growth, this level of self-exploration can also influence our views of data privacy and how we participate in the online realm. The creation of data cooperatives serves as an alternative for online users who want to ensure their data contributes to the causes they most value.

In the establishment and operation of data cooperatives, there are many ethical and moral considerations that require attention from the international community, such as:

- **Data Privacy & Consent:** Data cooperatives prioritise principles of privacy and informed consent. Members retain control over the collection, use, and sharing of their data, facilitated by transparent policies and consent mechanisms ensuring awareness and the option to opt-out.
- **Fair Data Practices:** Unlike conventional data gathering, data cooperatives are grounded in principles of fairness and equity. This entails the equitable distribution of benefits derived from data exploitation among members, mitigating existing power imbalances.
- **Democratisation of Decision-Making:** Data cooperatives can provide a platform to aid collective decision-making, ensuring the voices of all participants are heard. Members have the opportunity to actively participate in the creation of governance

structures, engage in data-related policy decisions, and contribute to shaping ethical guidelines governing data usage.

While understanding the ethicality of these initiatives is paramount in their formation, data cooperatives can also bring some key benefits:

- **Empowerment:** Participation in data cooperatives grants individuals control and agency over their data, enabling negotiations for fair terms with data processors, service providers, and other entities seeking access to their data.
- **Knowledge Exchange:** These cooperatives facilitate the sharing of insights and knowledge among members. Collaboration and access to diverse datasets empower participants to gain valuable insights, identify patterns, and discover innovative solutions to common challenges.
- **Community Development:** Data cooperatives can play a pivotal role in community development by unlocking the economic potential of data. Shared data resources inform targeted policies, support local businesses, and facilitate evidence-based decision-making, addressing community needs more accurately.

As the digital landscape evolves, data cooperatives emerge as a promising alternative to existing data ownership models. Anchored in ethics and values, these cooperatives champion privacy, fairness, and democratic decision-making, and provide users authority over their data and its uses. Through collaboration and empowerment, these initiatives have the potential to revolutionise the data economy, transforming individuals and communities into active participants rather than passive subjects. Embracing data cooperatives can reshape power dynamics surrounding data, paving the way for a more inclusive and ethically responsible future on a global scale.

Towards an Equitable Data Future: Data Cooperatives x Digital Public Infrastructure

Sarthak Satapathy

Digital Impact Alliance

Post COVID19 countries, multilateral and bilateral institutions, and civil society organizations have put their heads together on advancing digital public infrastructure (DPI) responsibly towards building a secure global digital economy. The first section of the outcome document from the 2023 G20 New Delhi Leaders' Declaration (G20 2023) defines DPI as -

“Digital public infrastructure, hereinafter referred to as DPI, is described as a set of shared digital systems that should be secure and interoperable, and can be built on open standards and specifications to deliver and provide equitable access to public and/or private services at societal scale and are governed by applicable legal frameworks and enabling rules to drive development, inclusion, innovation, trust, and competition and respect human rights and fundamental freedoms.”

A key piece of this DPI puzzle is sharing and exchanging data across national and international boundaries involving public, private, and civil society organizations in a safe, equitable, and efficient way. A World Bank G2Px report (Qiang et al. 2022), states that countries that could leverage existing databases and ID systems to cross-verify individuals were able to deliver social assistance to 51% of the population, while those that couldn't only reach 16%. This speaks about the critical nature of the data exchange layer to facilitate interactions between other infrastructure layers like ID and payments.

This comes at a time when we are generating zetabytes of data by accessing, sharing, or providing our data to these organizations numerous times a day. To put it in perspective, every minute - Google sees 5.9 Billion searches, 500 hours of videos are uploaded on YouTube, and 104.6K hours are spent on Zoom calls (Domo 2023). In October 2023 alone India's digital payments DPI, Unified Payments Interface clocked 11 Billion transactions.

At the moment all this data is concentrated with a few of these organizations and is hard to access and share even among them. While we see promising examples of sectoral data exchange models, there are not many to draw inspiration from at a national infrastructure level. Another contentious issue in all of this is data governance, with an emphasis on taking a human-centered lens on the same. A note on this by the Digital Impact Alliance, states the lack of focus on finding the right governance models is creating a data divide where it can only be leveraged by a few and extracted from the rest for various purposes (Vora 2023).

As countries increasingly go digital, there are many principles, and strategies that data cooperatives and data exchange systems can cross leverage to benefit individuals through their personal data.

The foundational premise of data cooperatives is unlocking and sharing data safely and securely to derive optimum benefits for all. This is a robust usecase against data being siloed (Girish and Avery 2022). The same inhibits the exchange of data and fails to unlock its true potential and an array of beneficial usecases for society. Data exchange systems can serve all of their stakeholders if measures are taken to make data accessible in a safe, inclusive, and equitable manner keeping people at the center.

Another important value that data cooperatives hold close is the agency of the collective and hence individuals within the collective over their data. This highlights consent as a key factor when it comes to setting up data cooperatives. Similarly, the ability of an individual to have agency over their data and choose whom to share it with is the foundational block of a data exchange network. With these consent networks built into data exchange systems, individuals can approve or reject data requests, revoke access to data, and share data at a granular level. This also results in individuals being able to seamlessly share their data with institutions to access services and goods like loans and

Data cooperatives can be found in groups that have been marginalized or are at the receiving end of the data divide. There are huge gaps in data about many social groups, that make their existence and experiences invisible. As data exchange systems are scaled nationally, these groups are at risk of exclusion as we move towards an increasingly digital economy. Data cooperatives highlight inclusion as an important value as their purpose is to mitigate this risk.

Data cooperatives can hugely benefit from DPI layers like data exchange systems that countries set up both at national and sectoral levels. The infrastructural approach of DPI provides a framework for innovative models like data cooperatives to be tested and scaled. Ideally, open and interoperable data exchange systems will enable multiple data cooperatives to plug in and exchange data across institutions rather than doing the same in a bespoke manner. This will help data cooperatives focus on their business models rather than worrying about

It is critical to build safeguards while thinking of leveraging data exchange systems for data cooperatives. In the current landscape governments are the stewards and builders of DPI initiatives across many countries, and those intent on targeting certain social groups can easily use these systems for surveillance and exclusion.

To avoid such a situation it is important to think about safeguards and accountability at multiple levels (DIAL 2023). Technical safeguards can play a key role in decentralizing data and ensuring that governments don't have access to individual data. Independent oversight institutions can hold the government accountable and public redressal mechanisms. Civil society groups can add to this and help create awareness and provide feedback to the oversight institution and government on implementation challenges on the ground. Policies and regulations like data protection and privacy laws will provide individuals the right to protect and have agency over their own data.

People centered governance models like data cooperatives can pave the way for further innovations that provide individuals benefits and agency on sharing data. Data exchange systems as a DPI layer can help enable this mission and inspire the overarching technical, governance, and service models of doing the same.

References

DIAL - Digital Impact Alliance (2023). What is good Digital Public Infrastructure. March, 2023. <https://dial.global/good-dpi/>

Domo (2023). Data Nerver Sleeps. Received from <https://www.domo.com/data-never-sleeps>

G20 (2023). [G20 New Delhi Leaders' Declaration](#). September 9, 2023.

Girish, S. and Avery, M. (2022). [Data Cooperative: Enabling Meaningful Collective Negotiation of Data Rights for Communities](#), December 1, 2022.

Qiang, C. Z., Rutkowski, M. & Pesme, J. (2022). [The COVID-19 crisis showed the future of G2P payments should be digital. Here's why](#), October 03, 2022.

Vora, P. (2023). [The data divide is real. We urgently need to close it](#). Digital Impact Alliance - Expert Comment. August 9, 2023.

Improving Global Governance: Data Cooperatives for Global Cooperation

Jascha Stein¹, Mei Lin Fung², Marta Bertolaso³, Fulvio Bersanetti⁴, Suruchi Gupta⁵, Ronald Strauß², Charlie Isaacs⁶, Christine Asjoma¹

¹*People Centered Internet*

²*Particip.ai*

³*Università Campus Bio-Medico di Roma*

⁴*Impact Innovation Department, Fondazione Compagnia di San Paolo*

⁵*GIANT Protocol Foundation*

⁶*Salesforce*

Published as Think7 Policy Brief for G7 Italy 2024

The emergence of artificial intelligence (AI) has led to significant changes in all sectors ranging from healthcare to finance. It has fundamentally altered global structures. Its rise has contributed to widening digital, economic and social divides.¹ The fast pace of technological progress has meant that understanding what AI really is and how it can be used effectively is largely confined to those who are already among the technical elite.

The role of artificial intelligence in our lives increases, so does its social and economic impact. As data becomes not only the basis for training artificial intelligence but is shaping by those important parts of our lives, it is becoming an increasingly important social and economic asset. Who controls the data, controls the reality – though controls our lives.

We need to develop governance and safety mechanisms for collecting data that neutrally reflects our decisions and actions. Otherwise in an increasingly evidence-based world the needs and priorities of us as individuals and our communities are not equally weighed against the interests of economic corporations and public institutions (Fung and Stein 2023).

Data responsibility will need to be managed in a collaborative way, so that a wide range of voices and perspectives are able to contribute to the foundational datasets that drive our AI systems.

¹ The divide refers to the socio-economic gap between those who have access to the Internet and related technologies and those who don't, either due to a lack of infrastructure or the necessary knowledge. This gap perpetuates inequalities in economic opportunity and social participation, exacerbating socio-economic disparities.

Current data governance systems are often controlled by a few dominant actors, leading to concerns about privacy, equitable access and fair use of data. They are based on centralized models, with one single authority managing and controlling data. An increasing number of governance experts envision more participative approaches like decentralized and hybrid models that combine central oversight with regional or local autonomy. Other approaches like federated governance for more adaptive governance to regional or local needs and open data governance in public institutions as well as cloud-based systems for managing information across multiple platforms are increasingly utilized.

The concept of data cooperatives emerges as a practical solution to implement these evolving governance aspirations and for bridging the growing digital divide. Data cooperatives are collective organisations in which the control and management of data is decentralised, and the data ownership remains with data originators (Pentland and Hardjono 2021).

People-centred data cooperatives ensure that the value derived from data equip a wider part of society with collective benefits. By giving individuals and their communities control over their data, these cooperatives enable more equal participation in the digital economy and promote innovation and social welfare. A key component that ensures the efficiency of data cooperatives is their integration into a so-called Prosperity Data Networks (PDNs). Via PDNs, the different data cooperatives can learn from each other, exchange their data in a protected manner, sell managed access to or train AI-systems with it (Stein et al. 2023).

As such, data cooperatives significantly advance the Sustainable Development Goals (SDGs) by promoting inclusivity, participation and shared responsibility. They impact SDG 11 (Sustainable cities and communities) by improving urban environments, SDG 10 (Reduced inequalities) through cooperative data access, SDG 17 (Partnerships for the Goals) through collaborative approaches, SDG 4 (Quality education) by improving learning outcomes and information sharing and SDG 8 (Decent work and economic growth) by using all facets of our extensive economic and social knowledge (Stein et al. 2023).

Data cooperatives

Data cooperatives decentralize data ownership and collective data management so that people and their communities can control and use shared data resources. People participating in these cooperatives can influence how their data is used, ensuring that it serves their collective interests and that ethical and security standards are upheld.

Each data cooperative manages its members' data, while the members individually and collectively decide the data oversight and use, for example, governance and who, how, when and why the data may be used. The transparency and integrity of these decisions can be ensured by recording them immutably on a decentralized ledger. AI can help process vast amount of data and automate the decision-making process, ensure that privacy is maintained, and individual and collective goals of the members are met.

Members in data cooperatives have an equal say in how the data is used. Participation is organized using agile methods (from the software industry) and methods for interdisciplinarity. Today marginalized groups are often excluded from traditional

decision-making processes. Data cooperatives can empower these groups by giving them access to data and platforms to voice their concerns and ideas, and to express their perspectives and priorities. This empowerment is essential to ensure that solutions are equitable and consider the needs of all members of the community. Data cooperatives are therefore participative and increase the sense of mutual responsibility and accountability among their members.

Data cooperatives provide a platform, where in a regulated manner, members of a data cooperative can contribute, access and analyse data that can be used to solve individual and collective problems. Such a body of information in the hands of a community of individuals can be an extremely powerful public good for collectively generating knowledge and as a tool for identifying problems, understanding their scale and in turn developing the right focus for action. For example, in a community with an environmental problem, environmental data cooperatives can collect local environmental data, share it with their members and with other data cooperatives through Prosperity Data Networks to look for sources of pollution, solutions developed by other data cooperatives and develop evidence-based plans for solutions. Or, in a context of weakening social capital in communities, sharing relational data (knowledge, skills, interest, etc.) can help rebuild a new community tissue (Bertolaso and Boschetto 2021).

Data cooperatives facilitate the sharing of information both within the community and with external stakeholders. This shared knowledge can include best practices, lessons learned and success stories that can serve as inspiration and guidance for other communities facing similar challenges (Fung et al. 2023). Transparent information sharing builds trust among members and with external bodies, which increases the credibility and influence of the cooperative. Therefore, data cooperatives foster collaboration not only among their members, but also with other cooperatives, organizations and government agencies. By pooling resources and expertise, they can tackle larger, more complex problems more effectively. This synergy maximizes the impact of their efforts, leading to more sustainable and lasting solutions.

Prosperity Data Networks (PDNs)

PDNs exponentially increase the benefits of data cooperatives, creating a broader platform for collaboration, innovation and shared prosperity. They are AI-powered digital data repositories that connect individual data cooperatives into larger networks (Stein et al. 2023). These networks expand efficiency, reach and impact of data cooperatives by facilitating the sharing of data and insights across communities.

PDNs advance the sharing of data among data cooperatives. One of the key benefits of PDNs is that they bring together large amounts of data from multiple sources. By using blockchain, the security and integrity of these data exchanges is ensured. With the help of AI, relevant data across different cooperatives can be analysed to identify trends and insights, enabling seamless data integration and improved collaboration, that can significantly boost innovation within communities and help address important societal challenges like environmental sustainability, infrastructure development, affordable healthcare and education.

Standardized protocols establish interoperability within PDN's and the large network of PDN's. Blockchain technology reinforces interoperability by providing transparent and secure frameworks for data transactions across different systems within the PDN ecosystem. AI enables the automatic translation of data formats between different systems, optimizing exchanges and integration within the PDN ecosystem.

PDNs connecting data cooperatives encourage collaboration. By bringing together people with different backgrounds and expertise, data cooperatives provide a platform for collaborative problem solving. This collaboration not only leads to more resilient outcomes, but also fosters a sense of community among members. As they work towards common goals, trust naturally develops through interactions sharing perspectives and experiences resulting an increasing potential for sustainable resilience and prosperity.

Trust building by implementing data cooperatives and PDNs

Data cooperatives empower people and their communities by giving them control over their most valuable asset in the digital age: their data. While the data cooperative adheres to rigorous regulatory frameworks like GDPR people can track how their data is being used, managed and protected. A dashboard accessible to all cooperative members gives them the opportunity to monitor real-time updates on data usage and data sharing within the cooperative and with external entities.

Participation and engagement are fundamental to building trust within communities. Hence, they are more than just operational aspects of data cooperatives; they are the heartbeat that increases trust between members, especially in the context of data exchange-based collaborations. The participative nature of data cooperatives, where members have a say in decision making processes, fosters a sense of ownership and responsibility. This involvement is critical to building a foundation of trust, as people are more likely to trust systems in which they are actively involved and where their voice is heard and valued.

The participation process in the governance of data cooperatives uses AI algorithms to gather feedback from a diverse range of members, making sure that all voices are heard in the governance process. Additionally, blockchain technology creates an immutable ledger of all data transactions within the cooperative. This includes entries on data access, updates, and sharing, ensuring that all data interactions are traceable and verifiable. With blockchain a transparent audit trail is available for members to verify the integrity of governance processes.

This transparency is critical to allaying fears and suspicions about data misuse – a common concern in the digital age. It also helps set clear expectations and standards for data use. People, as members of a data cooperative, feel more empowered, develop a sense of belonging and commitment and are more likely to participate actively and contribute positively to the goals of their cooperative. Participation strengthens data cooperatives as member-owned and governed institutions able to respond to and manage conflict and build and keep trust.

To manage these collaborations efficiently, implementing smart contracts (Garnett 2023) in

data cooperatives automates various operations, including the enforcement of data sharing agreements and permissions management. Smart contracts automatically execute pre-defined conditions, reducing the need for intermediaries and increasing efficiency. This implementation significantly reduces operational costs and minimises the likelihood of disputes, resulting in a legally enforceable data governance and digital rights management framework.

For data cooperatives as socio-economic systems we propose the creation of community-owned and governed data marketplaces based on the underlying concept of PDNs. Such cooperative marketplaces ensure secure data exchange and economic use under strict guidelines. Sale profits benefit the individual members as well as their community to fund further data collection and support cooperative activities.

Case studies: Global examples of community networks

Recent discussions and policy processes concerning data cooperatives emphasize their potential role in balancing power dynamics in the data economy and enhancing community control over data. Data cooperatives are being recommended for their ability to navigate the complex landscape of data rights and collective governance, providing a framework that could empower individuals and communities against the backdrop of digital and data-driven transformations (Ada Lovelace Institute 2021). Data cooperatives will play a major role in promoting data sovereignty as they act as a tool for communities to retain control over their data and overcome the traditional power imbalance in the data economy (Calzada 2021).

In this evolving context, data cooperatives are being positioned not just as economic entities but as crucial instruments for ethical data governance, highlighting their potential to mediate between individual data creators and large data-utilizing organizations. This approach could pave the way for more equitable distribution of the benefits derived from data, aligning with ongoing international efforts to manage data as a collective resource.

We demonstrate two case studies of organized collaboration. In both case studies, the collaborative model, whether in data management or healthcare, illustrates the potential for greater efficiency, equity and community engagement. The examples showcase valuable insights about how collaborations naturally evolve into data cooperatives. Set up to address various socio-economic challenges, sharing data improved impact and outcomes for regional and global cooperation.

Case study 1: Fondazione Compagnia di San Paolo

[Fondazione Compagnia di San Paolo](#), an Italian philanthropic foundation based in Turin, is an example of how collaboration with data can be aligned with and advance the UN SDGs (Bersanetti et al. 2022). The Foundation has made a strategic shift to focus its efforts on three main areas: Planet (addressing the overarching challenges facing the future of the planet and future generations through effective action that contributes to environmental protection, health system development, research and innovation performance), People (promoting prosperity and social cohesion, fighting poverty and reducing inequality) and

Culture (supporting culture and the arts, creativity and heritage). This focus on the SDGs has been driven by innovative strategies that emphasize data-driven decision-making.

The Foundation's approach highlights the potential that lies in using shared, ethically managed data to make informed decisions in philanthropic endeavours. Using data cooperatives and matching internal and external data from institutional sources, Fondazione Compagnia di San Paolo ensures that its goals are aligned not only with the broader SDGs, but also with the specific needs and priorities of the communities it serves. This alignment enables the Foundation to measure over time and maximize its impact on critical global challenges such as climate change, social inequality and cultural development. Starting from key performance indicators (KPIs), the Foundation has a strong commitment to public participation and citizen engagement, it releases a digital platform and a yearly impact report.

The Foundation's model demonstrates how data collaboration can enable organisations to leverage diverse data sets for comprehensive analysis, leading to more targeted and effective philanthropic interventions. As stated in its Data Strategy 2023-2030, Fondazione Compagnia di San Paolo is setting a precedent for how foundations can use data collaboration to improve their contribution to global sustainability and community well-being.

Case study 2: Federally Qualified Health Centers

[Federally Qualified Health Centers \(FQHCs\)](#) in the United States are an example of how collaborative structures can have a significant impact on communities. FQHCs are community based health care centres who provide primary care services in underserved areas to vulnerable populations. FQHCs operate under a patient-centred governance structure, with a board of directors composed largely of patients who use the health centre's services and reflect the needs and aspirations of the community they live in (Michener et al. 2019).

These centres are designed to reduce barriers to comprehensive health services such as cost, lack of insurance, distance and language. They provide services regardless of a patient's ability to pay, with fees based on a patient's income level. The structure of FQHCs ensures community participation in health care strategy, planning and implementation of decisions and is consistent with the core principles of data cooperatives, which emphasize community involvement and control over resources.

Additionally, the FQHCs have been organized in so called Breakthrough Collaboratives which allow the 8,000 individual health centres to collaborate on specific topics of interest to people in each health centre. The outcome has been an extraordinary progress on achieving outcomes on a wide range of health and healthcare themes by activating multi-sector partnerships (Michener et al. 2019).

The success of FQHCs underscores the benefits of cooperative structures in improving access to essential services. By involving the community in governance, these centres ensure that the services provided meet the needs of community members. The model demonstrates how collaborative structures can lead to more equitable and efficient delivery

of services that contribute to the overall health and well-being of the community.

Recommendations for the role of the G7 in promoting data cooperatives

As a grouping of the world's leading economies, the G7 is in a unique position to take a leadership role in the development and implementation of data cooperation. Our recommendations for the G7 are:

- Practical research on development of data cooperatives: To research and develop practical collaborative digital structures, we recommend the financial and organisational support of digital cross-border sandboxes for data cooperatives.
- Legal framework for data cooperatives: We recommend the establishment of an institution dedicated to the development of standards and regulatory frameworks that facilitate the creation and operation of data cooperatives. These frameworks should ensure seamless integration into the digital economy and compliance with data protection rules.
- Promote international cooperation on trustworthy data sharing: G7 plays a crucial role in encouraging and facilitating alliances among national/local institutions, corporates, universities and non-profit organisations for data cooperatives between countries to establish a global culture of trustworthy data sharing and cooperation.
- Integrate data cooperatives into global policy agendas: We recommend including data cooperatives in global discussions on digital governance and sustainable development and highlight their contribution to the SDGs.
- Establish secure, automated data cooperatives with AI and blockchain: We recommend using AI and blockchain to uphold the integrity of data exchanges and decision-making processes, enhancing collaboration, innovation, and trust among cooperative members within G7.

Pioneering a conscious path towards a sustainable resilient and prosperous future

What was perceived and criticised as science fiction yesterday has become reality today in the age of exponential development of artificial intelligence – faster than most of us expected. Data increasingly becomes not only a passive result of our lives and overall existence, but especially through AI trained based on it, data will actively shape our realities and lives in the future – in ways we can't even imagine yet.

AI inferences are as good as the data AI is trained on. When AI is trained primarily on data aggregated by large tech or powerful states, or when trained on non-representative

discussion forums like Reddit, AI is perpetuating biases that are not in the broad interests of common good. Strong AI, and later artificial general intelligence (AGI), trained on current sources of data will not ensure the diversity-based resilience of humanity. By training our AI systems based on data that reflects us and our perspectives of reality in the most neutral and participative way – in form of data cooperatives – we can systematically balance the power of AI and its socio-economic value to make its applications and outcomes accessible for everyone. We are positioned to mitigate the risks associated with strong AI and AGI when these are trained on data based on a participative governance. Only when AI training data includes data cooperatives can AI become a part of our socio-economic systems – leveraged by us as a tool for humanity and not humanity leveraged as a tool for AI.

Nature demonstrates that viruses typically do not aim to structurally harm their hosts. Similarly, by applying this natural principle, data cooperatives can act as an intermediate layer between humanity and AI, resembling our immune system. This layer can leverage the positive value of artificial intelligence while managing the internal and interconnected balances of our socio economic systems.

Data cooperatives and Prosperity Data Networks are evolving collective governance concepts, designed to be more adaptive, anticipative, responsive and dynamic than existing more static and less participative data governance concepts.

Implementing data cooperatives will reinforce the G7's status as a leader in promoting equitable and sustainable digital economies. This approach will prepare the way for leveraging the advances in AI and digital technologies for the benefit of our regional and global communities.

References

- Ada Lovelace Institute (2021). Exploring legal mechanisms for data stewardship. <https://www.adalovelaceinstitute.org/?p=9185>
- Bersanetti, F., Candela, F., and Mulassano, P. (2022). Doing philanthropy at the time of the Sustainable Development Goals: The case of Fondazione Compagnia di San Paolo. *The Foundation Review* 13(4): 7-17.
- Bertolaso, M., and Boschetto, E. (2021). Managerial sustainability for relational workplaces. 21st International Symposium on Ethics, Business and Society, IESE Business School, University of Navarra.
- Bertolaso, M., Capone, L., and Rodríguez-Lluesma, C., eds. (2022). *Digital Humanism: A Human-Centric Approach to Digital Technologies*. Cham: Palgrave Macmillan
- Calzada, I. (2021). Data co-operatives through data sovereignty. *Smart Cities* 4(3): 1158-1172.
- Fung, M. et al. (2023). People-centered science and digital transformation: A practical proposal for the G7 and G20. Think7 Policy Briefs. <https://think7.org/?p=397892>

Fung, M., and Stein, J. (2023). Addressing the digital divide and mitigating the risk of AI by people-centered, collaborative digital regulation. *The Bridge* 119(3):15-19. https://www.nxtbook.com/nxtbooks/ieee/bridge_issue3_2023/index.php#/p/14

Garnett, A. G. (2023). Smart contracts vs. dApps—how are they different?. *Encyclopedia Britannica*, 30 March. <https://www.britannica.com/money/smart-contracts-vs-dapps>

Michener, J. L. et al., eds. (2019). *Practical Playbook II: Building Multisector Partnerships That Work*, 2nd ed. Oxford: Oxford University Press

Pentland, A., and Hardjono, T. (2021). Building data cooperatives. In Pentland, Alex, Lipton Alexander, and Hardjono, Thomas, eds. *Building the new economy: Data as capital*: 19-34. Cambridge: MIT Press

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster

Stein, J. et al. (2023). [Enhancing global SME financing through prosperity data networks: An integration of Hayek's and Sen's economic insights in the digital age](#). *SME Finance Forum*, 25 September.

Where we Stand

Data Cooperatives to Empower SMEs in the Construction Sector

Michael Bühler¹, Konrad Nübel², Thorsten Jelinek³, David Riechert⁴, Lothar Köhler⁵, Pia Hollenbach¹

¹*Hochschule Konstanz*

²*Technische Universität München*

³*Centre for Digital Governance, Hertie School*

⁴*GemeinWerk Ventures*

⁵*Benefit Unternehmensentwicklung GmbH*

The construction sector, characterized by a large number of small and medium-sized enterprises (SMEs), faces particular challenges in the digital age. Data cooperatives, such as the Bavarian Construction Data Cooperative (Bauform eG), currently being founded, represent a transformative solution by providing these SMEs with a platform to leverage their collective strength in data management, sharing and use. Initiated by GemeinWerk Ventures and supported by the Bavarian Construction Industry Association, Bauform eG provides shared services and construction data through its member companies via a digital collaboration platform. This innovative approach improves collaboration and organization within the construction value chain and aims at governance innovation to enable trustful data exchange between stakeholders and create a pre-competitive space of trust. This model promotes SME productivity and innovation through ecosystem collaboration and represents an important step in the digital transformation of the construction sector.

Model and Benefits

Data cooperatives are organizations that are owned and managed by members. They facilitate the collective development, collection, processing, management, analysis and sharing of data. They allow members, including SMEs and individuals, to retain control over their data while benefiting from shared resources and expertise. This collaborative structure aims to promote data sovereignty and democratize the digital landscape. Bauform eG represents a sector-specific model in which the trustful sharing of data by stakeholders creates a pre-competitive space that fosters collaboration. Such cooperatives aim to:

- overcome the fragmented nature of the construction sector;
- provide a platform for sharing resources and insights;
- enable SMEs to access and leverage big data analytics; and
- foster innovation through shared insights and data-driven decision-making.

For SMEs in the construction sector, data cooperatives offer a plethora of benefits. They provide a platform for sharing insights, resources, and innovative practices, thus driving

productivity and innovation. By participating in data cooperatives, SMEs can overcome resource constraints and gain access to a larger pool of data, which is crucial for informed decision-making and strategic planning. Bauform eG, specifically, improves collaboration and organization within the construction value chain, intensifying industry collaboration and creating a trustworthy environment that drives productivity and innovation through ecosystem collaboration. Data cooperatives play a crucial role in bridging the data divide, empowering traditionally underrepresented SMEs in the digital space by providing them with equitable participation in decision-making processes and access to data-driven innovation.

Challenges and Future Directions

The model of data cooperatives, while promising, encounters several challenges that need to be addressed for their widespread adoption and success in the construction sector. Primary challenges include establishing trust among members, navigating complex data privacy laws, ensuring operational scalability, and creating sustainable economic structures.

1. Data Privacy and Security: The balance of individual privacy within the data cooperative model necessitates the deployment of robust safeguards. In an environment of extensive data sharing, the implementation of rigorous data anonymization protocols and stringent cybersecurity measures is imperative to uphold the sanctity of privacy.
2. Regulatory Compliance and Barriers: The multifaceted nature of regulatory landscapes poses challenges, with existing regulations in many countries not adequately supporting or even hindering the establishment and operation of data cooperatives. This limited regulatory support necessitates a sophisticated approach that accommodates the variations in laws across different jurisdictions.
3. Resource Constraints: Small communities and SMEs often face resource constraints that restrict their ability to develop and implement digital governance structures, open standards, and cooperative models. This challenge is compounded by the digital divide, where unequal access to digital infrastructure, skills, and resources exacerbates existing inequalities.
4. Participatory Governance: As cooperatives scale, the complexity of maintaining effective participatory decision-making processes magnifies, demanding careful strategizing and adept management.
5. User Awareness and Interoperability: Addressing user awareness, ensuring interoperability, and fostering competition are vital for the cooperatives' success.

To overcome these challenges, this synopsis outlines policy recommendations supporting the establishment and operation of data cooperatives for the construction sector. Emphasized needs include supportive legal frameworks, funding mechanisms, and capacity-building initiatives. These efforts aim to create a conducive environment for data cooperatives, fostering data sovereignty and democratizing the digital landscape. In conclusion, despite the hurdles, with the right blend of policy support, technological

framework, and stakeholder collaboration, data cooperatives can become critical players in democratizing data governance and fostering inclusive digital commons.

Conclusion: Revolutionizing Construction through Data Cooperatives

Data cooperatives in the construction sector are more than just a collaborative framework; they represent a paradigm shift towards a sustainable, inclusive, and innovative future. By embracing the data cooperative model, SMEs in construction are not just leveraging data for mutual benefits; they are partaking in a transformative journey towards integration, efficiency, and groundbreaking innovation. This model serves as a beacon of cooperative innovation, harnessing the collective power to drive novel solutions and informed decisions. By pooling resources and knowledge, data cooperative members are not only fueling technological advancements but also nurturing environmentally conscious practices and sustainable growth.

The economic, social, and environmental impacts of these cooperatives are profound. They offer a reusable and non-depletable capital in the modern knowledge economy, with the potential to significantly contribute to GDP. The return on investment in these cooperatives is not just financial; it ensures equitable distribution of economic gains, promoting a sustainable and inclusive growth model. Data cooperatives, particularly in the construction sector, are poised to make a substantial impact, contributing to inclusive and sustainable growth in SMEs and communities alike.

This synopsis aims to resonate with a broad audience, including policymakers, legislators, journalists, and the public, by showcasing the transformative potential of data cooperatives. It's a narrative about empowering SMEs in the construction sector, paving the way for a more connected, efficient, and innovative industry. As we look towards the future, the role of data cooperatives in driving economic, social, and environmental progress cannot be overstated. They stand at the forefront of the digital transformation, offering a path to a more sustainable and prosperous future for the construction sector and beyond. In summary, data cooperatives in the construction sector are not just a means to an end but a catalyst for a broader change, aligning with global goals of digital transformation, economic development, and sustainability.

References

- Bühler, M. M., Nübel, K., Jelinek, T., Riechert, D., Bauer, T., Schmid, T., & Schneider, M. (2023). Data cooperatives as a catalyst for collaboration, data sharing and the digital transformation of the construction sector. *Buildings*, 13(2), 442.
- Bühler, M. M., Calzada, I., Cane, I., Jelinek, T., Kapoor, A., Mannan, M., Mehta, S., Mookerje, V., Nübel, K., Pentland, A., Scholz, T., Siddarth, D., Tait, J., Vaitla, B. & Zhu, J. (2023). Unlocking the power of digital commons: Data cooperatives as a pathway for data sovereign, innovative and equitable digital communities. *Digital*, 3(3), 146-171.

Bühler, M. M.; Calzada, I.; Cane, I.; Jelinek, T.; Kapoor, A.; Mannan, M.; Mehta, S.; Micheli, M.; Mookerjee, V.; Nübel, K.; Pentland, A.; Scholz, T.; Siddarth, D.; Tait, J.; Vaitla, B.; Zhu, J. Data cooperatives as catalysts for collaboration, data sharing, and the (trans)formation of the digital commons. Policy Brief. Task Force 2: Our Common Digital Future: Affordable, Accessible and Inclusive Digital Public Infrastructure. In Think20 (T20) India. 2023.

Salus.coop, data cooperative for health research.

Javier Creus

Salus.coop

In June 2007 a group of 12 individuals created Salus.coop as a citizen-led data cooperative for health research. Salus.coop enables citizens to share their own health data, which is too often locked in systems to which they have limited access and power. Salus allows people to gain more control on the governance of their health data, and empower them to make decisions on its use, by whom, and under what conditions. Salus acts therefore as a vehicle to generate data abundance for researchers, who typically struggle to obtain relevant data for their endeavors. The data shared through Salus encompasses holistic information of people's health, enabling the research arena to take advantage of a combination of clinical, socio-economical and real world lifestyle data.

Since the proof of existence of the data is substituted by a personal key, Salus.coop itself does not need to store any personal data. The data is therefore stored in a decentralized, distributed manner by design. The salus.coop software stack is also offered to other communities so that they can securely share their data on their own terms and for other purposes on the basis of a cost-sharing agreement.

Why Salus.coop?

As Daron Acemoglu states in "Power and Progress", technological advances do not automatically translate into social progress, it is social innovation through new ways of organizing that forces new value redistribution (Acemoglu and Johnson 2023). There is still a gap between open knowledge (to all) and (individual) sovereign identity, and Salus.coop addresses this issue.

In 2021 only 3% of healthcare expenditure was spent on preventive care (Eurostat 2021), even if its effects on general health and economic savings are unquestioned. Accordingly, most research carried out by medical institutions is focused on illness rather than on the conditions for healthy living and wellbeing, and thus most appreciated data is genetic and clinical, rather than socio-economical or lifestyle (e.g. activity, sleep, diet, social relationships). This is coherent with the fact that business models related to illness research are very well defined: organizations that create a successful treatment for a disease are rewarded for it by private or public institutions.

What are the legal and technological building blocks?

From evaluating 8,000 randomly generated sharing scenarios, citizens co-designed the Salus Common Good Data Sharing License. This license has been integrated into the Ethical Commission Review of the Hospital Clinic from Barcelona to rule project data sharing. The license consists of five main conditions:

1. Use is permitted for health-related research
2. by non profit organizations
3. which openly share results
4. while preserving anonymity
5. and the right to stop sharing data at any time.

Additionally, a usable mobile app for citizens and convenient web tools for researchers, built by a combination of zero-knowledge protocols, the use of personal repositories in the InterPlanetary File System (IPFS), and blockchain notation, allows citizens to keep their anonymized data under their own control and document metadata and conditions of the data that they directly share with the research institutions.

What has Salus.coop achieved so far?

Over the years, Salus.coop reengineered a new ecosystem of trust, and has received support from the Mobile World Congress, the Catalan Government, Abacus and Suara Cooperatives, Ideas for Change, and leading research institutions in Catalonia.

Today, Salus.coop enables participants to share their European Health Summary issued by the Government of Catalonia, all the information available at the Google Fit and Apple health apps and compatible devices, and answers to questionnaires. Health institutions are adopting these new data altruism paradigms. For instance, since December 2022, one hundred beta testers are already contributing data to ISGlobal (Barcelona Institute for Global Health) "Cities Health" project which studies how heat, air quality and proximity to green zones affect mental health and cognitive capacities. These advances would not have been possible without the persistence and altruistic dedication of a selected group of individuals, and sustainability is not yet guaranteed.

How does Salus.coop sustain itself?

Salus.coop's business model was supposed to generate income from fees charged to research projects in exchange of recruiting data donors that would use the Salus app to share their data, or by sharing its platform Software as a Service with other communities.

We are learning as we develop and experiment that as we, as citizens, acquire agency over our own data, we start finding out how our own agenda might be misaligned with the actual incentives for the use of our data.

There is no doubt that we are in need of a new breed of social institutions, able to reframe discussion on citizen's data from an individual menace to a collective opportunity and to structure an alternative organization to commercial or public centralization of data in the AI era. We hope that Salus.coop and other similar initiatives can be part of this movement towards the use of data for common good which is now becoming more and more integrated in key directives such as the Data Governance Act.

References

Acemoglu, D. & Johnson, S. (2023). Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity. Public Affairs.

Eurostat (2021). 3% of healthcare expenditure spent on preventive care. January 18, 2021. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210118-1>

Data Cooperatives: A Case for Collective Participatory Governance

Anastasia Kalinina and Alexandra Seaman

reState Foundation

Our policy decisions today have the potential to shape society and democracy for decades, offering an opportunity for genuine flourishing and data governance renewal. Participatory bottom-up governance and cooperatives provide opportunities to shift from the current neoliberal focus on individual data toward more collaborative approaches. Data cooperatives can be ideal test beds for establishing new experimental governance models that are collective and participatory. Advocates of this framework argue for a generative and public value-centric approach, challenging the dominance of individual property in social relationships (Broumas 2020; Marella 2016). This concept also extends beyond traditional commons frameworks, such as the Ostromian tradition, which primarily views commons as social institutions managed by communities for resource management (Coyle et al. 2020; Ostrom 1990; Prainsack 2019).

Collective Governance of Data Cooperatives

Data, often described as a "democratic medium" (Viljoen 2020), can be considered a shared property, created collectively by large groups through interactions, sociality, and labor, rather than by any one individual. For instance, Facebook posts not only belong to the person who publishes them but also to those who like and comment. Similarly, emails involve both the sender and the respondents. Even seemingly personal genomic data become fundamentally shared, as they can be used to predict patterns in others associated with the subject.

The value of data increases through aggregation, containing predictive insights about people in relation to each other. New generation AI-based technologies focus on analyzing aggregated data sets rather than individual data. The aggregation phenomenon is also beneficial for data brokers and other structures in the data-driven economy, which have established lucrative markets based on personal information. Therefore, the most effective approach to address the negative consequences of this economy – including heightened inequality, limited access to opportunities, and the emergence of dominant platform monopolies – is to consider data as a shared resource requiring collective governance solutions.

Participatory Governance of Data Cooperatives

Incorporating mechanisms for participatory data stewardship allows individuals and society to gain a deeper understanding and recognition of the broader public advantages derived from 'donating their data' within a specific context (Ada Lovelace Institute 2021). Rather than feeling like passive consumers of Internet services, users would feel empowered as active producers and participants in the creation of value. Data can be regarded as labor, giving people personal dignity and political responsibility (Posner & Weyl 2018). As AI is

understood more as a "collective intelligence" with data laborers being the sources of value in digital systems, people would treat the useful insights of ChatGPT or other AI-enabled solutions not as advice from robots, but as repositories of human contributions, similar to how they understand Wikipedia or the insights on their Facebook wall.

Involving people in the design, development, and deployment of data governance frameworks encourages interactions with a range of views, perspectives, norms, and lived experiences, helping minimize the risks of groupthink, unconscious biases, and misalignments between intended and actual outcomes. Effective participation enables a better understanding of public perspectives, values, trade-offs, and choices, leading to an increased quality of data production over time. Data designers are thus able to understand the boundaries of data use and management: what people feel is acceptable and unacceptable for the access, use, and sharing of their data in a given context.

Engaging in participatory approaches can establish mechanisms for checks and balances, promoting greater societal and economic fairness, addressing power imbalances, and enhancing public trust in the utilization of data (Ada Lovelace Institute 2021). Importantly, participatory governance contributes to process legitimacy, defined as the 'reservoir of support that allows governments to deliver positive outcomes for people' (Centre for Public Impact 2018). In the case of data, it increases public confidence in the way data is gathered and used, particularly where third parties access data, by ensuring people can oversee and hold accountable decisions about rights to access, share, and control that data. Besides overall public support, such legitimacy gives developers and deployers a 'social license to build' (Ada Lovelace Institute 2021).

Landscape Mapping of Data Cooperatives Applying Collective Participatory Governance

A few groups of data cooperatives lead the way in unlocking new governance of data that is both collective and participatory – from [MIDATA](#) providing a citizen-directed model of data administration through the general assembly consisting of members of the cooperative, to [Open Data Manchester](#) and [Carbon Co-op](#) exploring collective approaches to both organizational governance and governance of data access, to [Salus Coop](#) creating a 'common good data license for health research' through a crowd-design mechanism.

There are many other cases identified of data cooperatives testing and implementing bottom-up governance models, where technological architecture is used to allow individuals to know about and manage any data they contribute. Citizen science initiatives like [Zooniverse](#) enable people to decide how and when data should flow. Technical solutions like [Solid Pods](#) allow people to determine who can access data about them, and digital rights like data portability enable people to move data from one provider to another. Data cooperatives like [POSMO Coop](#) enable people to collectively decide how data is stewarded and contribute it towards a common cause, such as improving city mobility. [Camden Data Charter](#), [Digi.me](#), PolyPoly's [polyPod](#) vault, and [Swash](#) are some other remarkable examples of bold experimentation.

Conclusion and Going Forward

While participatory models encourage people to play a more active part in stewarding data about themselves or that they have a vested interest in, the adoption of cooperative or similar infrastructures is yet to emerge stronger as a feasible model for governing data. We have already come a long way from accepting data monopolization as a legitimate business practice. However, more efforts must be put into studying, initiating, and promoting new approaches to collective governance to arrive at a new political-economic settlement with better incentives for people and less alarming concentration of power in the hands of a few.

The public sector has an essential role to play in establishing spaces that break away from extractive, property-based models of managing data. Data cooperatives need to become mandatory counterparties for businesses looking to use data. They also need to be subject to special regulations. These regulations would ensure that they remain wholly independent of conventional data-exploiting businesses and that they genuinely serve their members' interests.

Policymakers are starting to apply new legislation that seeks to reclaim the power balance and pilot projects to develop alternative data governance models that are more collective and participatory (Millard 2023). Without further adoption of strong policies and regulation to reinforce them, data cooperatives will either fail or remain marginal in an adverse economic and legal environment. On the other hand, the space of grassroots data commons needs to be resourced and sustained to create a safe space for experimentation and the creation of new governance tools. Both will serve to create new collective governance models, which can be the foundation of the co-creation of public value.

Just as we established new institutions and democratic methods for all levels of societal organization, we now need to embed collective data rights in the participatory collective decision-making over data. State-level actors can play a key role, but other actors can engage as well, at different scales and layers of society.

References

Ada Lovelace Institute (2021). Exploring legal mechanisms for data stewardship. Retrieved from <https://www.adalovelaceinstitute.org/report/legal-mechanisms-data-stewardship/>

Ada Lovelace Institute (2021). Participatory Data Stewardship: A Framework for Involving People in the Use of Data. Retrieved from <https://www.adalovelaceinstitute.org/report/participatory-data-stewardship/>

Broumas, A. (2020). Intellectual Commons and the Law: A Normative Theory for Commons-Based Peer Production. University of Westminster Press. Centre for Public Impact. (2018). Finding Legitimacy. Retrieved from <https://www.centreforpublicimpact.org/assets/documents/Finding-a-more-Human-Government.pdf>

Coyle, D., Diepeveen, S., Wdowin, J., Kay, L., & Tennison, J. (2020). The Value of Data: Policy Implications.

Marella, M. R. (2016). The Commons as a Legal Concept. *Law and Critique*, 28(1), 61–86. doi:10.1007/S10978-016-9193-0

Millard, J. (2023). Impact of digital transformation on public governance - New forms of policy-making and the provision of innovative, people-centric and inclusive public services. In M. Manzoni & S. Schade (Eds.), Publications Office of the European Union, Luxembourg. doi:10.2760/204686, JRC133975

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.

Posner, E. A., & Weyl, E. G. (2018). *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton University Press, Princeton.

Prainsack, B. (2019). *The Governance of Digital Health: Reimagining the Future of Health Technologies*. Cambridge, UK: Cambridge University Press.

Viljoen, S. (2020). A Relational Theory of Data Governance. *SSRN Electronic Journal*. doi:10.2139/ssrn.3727562.

Legal and Regulatory Frameworks for Data Cooperatives in the EU and beyond

Dr. Morshed Mannan

Robert Schuman Centre for Advanced Studies, European University Institute

Data cooperatives are an organizational structure and intermediary for data governance that has received increased attention in recent years. As explained in a policy brief for T20 India 2023 that I co-authored (Bühler et al. 2023), “a data cooperative is a member-owned organisation that democratically manages and shares data while allowing members to retain control over their data”. Compared to other types of data intermediaries, data cooperatives have collective transformative potential as they can promote sovereign data exchange, enable digital entrepreneurship, protect digital rights, and contribute to community well-being.

From the growing prominence of Big Tech over the past 15 years to the more recent surge in the use of artificial intelligence systems (Shearer 2024), we see local communities and MSMEs struggling to self-determine how their data is used and, consequently, shape their own digital future. Data cooperatives can be seen as a preferable alternative as they enable the local control, access, and use of personal and community data. Data access and sharing can generate considerable value (Reimsbach-Kounatze 2020), with data cooperatives potentially making market access less expensive. As member-led organizations, they strive to compensate members fairly and appropriately for their contributions. They facilitate the use of both individual and collective data rights (where the latter exists).

In recent years, several cooperatives and collectives have emerged that act as such an alternative. However, not all of them use the cooperative legal entity form. [The Drivers' Seat Cooperative](#) enabled gig economy workers in the ride-hailing sector to collect, pool, and analyse the data collected on a smartphone while undertaking work for ride-hailing platforms, such as Uber and Lyft. The pooled data allowed insights to be fed back to members so that they could optimise their income. The cooperative also sold data and insights to city agencies to enable better policy decisions, with profits from sales redistributed among members. This cooperative has recently been wound down with the application being transferred to the Workers' Algorithm Observatory at Princeton University. In Mexico, the COBI NGO worked with fishers' organizations to co-develop the [Pescadata](#) platform in a bottom-up manner, that enabled the sharing of data among fishers and the state. The application facilitates the sharing of catch data, creates an online marketplace for fish and other items relating to fishing, and enables the sharing of community solutions to problems experienced by fishers across the country and the wider region.

Data cooperatives are also appearing in the agricultural sector, sometimes with the assistance of traditional, longstanding cooperatives. The [Paliyakal Service Cooperative Bank](#) was founded in Kerala in 1943, before India's independence, and despite experiencing financial difficulties over the course of its long lifespan, it has supported a number of pioneering initiatives. One such initiative that it has collaborated with is the [Kerala Food Platform](#), which is currently being trialled with 60-70 households in Ezhikkara panchayat.

This platform, which is being developed by the Government of Kerala, uses transaction data to integrate producers, consumers, and other actors within the agricultural supply chain to deliver higher-quality organic produce, increase farmers' incomes, and improve the privacy and data control of concerned stakeholders. As these fledgling examples show, such cooperatives, collectives, and adjacent organizations can improve resource allocation, create jobs, and foster community development, thereby strengthening social cohesion. Cooperatives are expected to adhere to the International Cooperative Alliance's seven internationally recognised cooperative principles (coop 2024), which include education, training and information as well as concern for community. Consequently, these entities are expected to concentrate resources on improving the skills and capacity of members, including digital skills and the ability to use insights generated from data. This can, in turn, support community well-being by improving local economies and sustainability. Pescadata, for instance, can help with national statistics management and help local fishers meet mandatory sustainable fishing requirements.

Data governance is, unsurprisingly, becoming an issue for all cooperatives. Consequently, all cooperatives may become data cooperatives or supporters of data cooperatives in some sense. Indeed, recent publications like Trebor Scholz's *Own This!* and the EU JRC's report on Data Intermediaries document a wide range of collectives that enable participatory forms of data governance and control (Micheli et al. 2023).

While the appearance of data cooperatives is encouraging, the regulatory landscape for data cooperatives is uneven and often uncondusive. The European Union is the first major jurisdiction to recognize data cooperatives through its new [Data Governance Act](#). This supranational legislation recognizes a particular conception of data cooperative, which focuses on the negotiation of data rights, rather than data pooling as such. The Data Governance Act explicitly includes within its remit services provided by data cooperatives, as one of the types of data intermediation services. Data cooperatives are expected to help user-members make informed choices and to negotiate terms and conditions with data users before individual consent is given.

As legal entities, cooperatives come in many forms. Depending on the jurisdiction, they can be for-profit or non-profit/not-for-profit. The Data Governance Act, however, does not contemplate the latter forms of data cooperative, with charitable activities being broadly bracketed within the idea of 'data altruism'. The Data Governance Act creates the basis for a new organization—the 'Data Altruism Organisation'—which can collect and process personal and non-personal data to make this processed data available for general interest purposes. In a sense, the Data Altruism Organisation encroaches on some of the activities that data cooperatives could perform if they were more broadly conceived.

The Data Governance Act also sits uneasily with other legislation. It explicitly seeks to enhance the agency and control of individuals over their data, but the rights under the EU's General Data Protection Regulation (GDPR) continue to be personal and non-delegable beyond certain limited exceptions. It is clear that even with these regulations, there is still a need to legally encourage more collective approaches to data governance and data pooling (Mannan et al. 2022).

In other parts of the world, legal frameworks for data cooperatives are even more elusive. As mentioned in Bühler et al. (2023), there is thus an urgent need to develop support legal

frameworks for data cooperatives that are co-designed with budding data cooperatives. Partnerships with local governments and larger, traditional cooperatives could also open access to funding and other technical resources. If a conducive environment for data cooperatives are created, I believe that they can help humanity achieve SDGs 8 (“decent work and economic growth”), 9 (“industry, innovation and infrastructure”), and 11 (“sustainable cities and communities”) in particular.

References

Bühler, M. M.; Calzada, I.; Cane, I.; Jelinek, T.; Kapoor, A.; Mannan, M.; Mehta, S.; Micheli, M.; Mookerjee, V.; Nübel, K.; Pentland, A.; Scholz, T.; Siddarth, D.; Tait, J.; Vaitla, B.; Zhu, J. Data cooperatives as catalysts for collaboration, data sharing, and the (trans)formation of the digital commons. Policy Brief. Task Force 2: Our Common Digital Future: Affordable, Accessible and Inclusive Digital Public Infrastructure. In Think20 (T20) India. 2023.

coop (2024). Cooperative identity, values and principles. <https://ica.coop/en/cooperatives/cooperative-identity>

Mannan, M., Wong, J. & Bietti, E. (2022). [Data Cooperatives in Europe: A preliminary Investigation](#). Network Industry Quarterly 24(3).

Micheli, M., Farrell, E. Carballa Smichowski, B., Posada Sanchez, M., Signorelli, S. Vespe, M. (2023). [Mapping the landscape of data intermediaries](#). Publications Office of the European Union, Luxembourg, 2023.

Reimsbach-Kounatze, C. (2020). [Enhancing access to and sharing of data: Striking the balance between openness and control over data](#). In. Data Access, Consumer Interests and Public Welfare.

Scholz, T. R. (2023): [Own This!: How Platform Cooperatives Help Workers Build a Democratic Internet](#). Verso Books.

Shearer, E. (2024). [Take Back the Future: Democratising Power over AI](#). Common-wealth.org. September 11, 2024.

Where to Go

Harnessing AI Potentials: from Data cooperatives to AI commons

Amir Banifatemi

AI For Good, K5 Ventures, Resilient Futures Consortium, Responsible AI WG at the Global Partnership on AI

In an era where artificial intelligence (AI) is rapidly reshaping our world, it is becoming evident and imperative to integrate responsible AI practices with inclusive data management models. Future policies should advocate for the transformative potential of this integration, focusing on the synergy between data stewardship efforts and Digital and AI Commons.

The Evolving Landscape of Data Management and its Role in AI and Digital Commons

The last two decades have brought forth diverse data management models, each shaping our interaction with digital resources and their benefits. Data Cooperatives offer a member-centric approach, Data Commons ensure open access for all, and Data Collaboratives foster partnerships for societal good.

Data Cooperatives: A Member-Centric Approach

Data Cooperatives emphasize member ownership and democratic governance. Here, individuals or members collectively manage their data, prioritizing privacy, consent, and shared benefits. This model stands in contrast to traditional data governance structures, empowering individuals with greater control over their data.

Data Commons: Open Access for All

Data Commons represent shared digital spaces where resources like software, educational materials, and scientific data are accessible to all. This model fosters a culture of open access, collaboration, and communal improvement, extending the benefits of digital resources to a broader community.

Data Collaboratives: Partnerships for Societal Good

Data Collaboratives involve partnerships between various entities, including private companies, government agencies, and non-profits. These collaborations leverage combined data and expertise to address societal challenges, offering a model for responsible and impactful data use.

	Data Cooperatives	Data Commons	Data Collaboratives
Ownership and Governance	Member-owned and democratically governed organizations. Members have a say in data use and management.	Shared digital spaces governed by principles of open access, transparency, and community benefit.	Partnerships among entities like private companies, government agencies, and nonprofits. Governance varies based on agreements.
Purpose	Empower members through collective ownership and control of data. Provide benefits like financial rewards or improved services.	Democratize data access, making it freely available for use and analysis by anyone. Often used in academic and research contexts.	Solve societal challenges by leveraging combined data and expertise. Focus on research, policy-making, or social issues
Data Sharing and Use	Data managed and shared based on collective member decisions. Emphasis on privacy and consent.	Data made accessible with minimal restrictions to promote collaboration and innovation. Emphasis on open-source principles and community engagement.	Data sharing guided by collaborative goals, targeting specific projects. Focus on ethical and responsible data use
Membership	Composed of individual members who contribute and control their own data.	Open to the public or specific communities, with no formal membership structure.	Involves a range of stakeholders, including organizations, institutions, and sometimes individuals.

Digital and AI Commons: Expanding the Data Ecosystem towards public benefits

Digital Commons encompasses a broader range of digital resources, promoting open access and collaboration. The concept serves as a foundation for knowledge sharing and innovation across various sectors. Digital Commons serves as communal hubs, supporting a wide array of applications from academic research to grassroots projects, to knowledge sharing, training, and more, all underpinned by the principle of open and equitable access.

A subset of Digital Commons, AI Commons was originally envisioned as open platforms for collaborative AI development, is now evolving into a framework promoting equitable public access to AI resources including data, models, code, and computing power, aiming to develop and deliver AI technologies as accessible utilities that benefit society as a whole.

This expanded focus aligns AI Commons with the public interest, governed by ethics and accountability measures that enable inclusive innovation, backed by policy and regulation that encourages open and equitable participation. Rather than AI being an isolated or strictly academic endeavor, this vision positions AI Commons as a shared resource accessible to all, with technology leaders and governments contributing to a robust infrastructure for developing and deploying AI to foster the common good.

	Digital Commons	AI Commons
Focus and Scope	Encompasses a broader range of digital resources including software, educational materials, scientific data, etc.	Specifically focuses on AI resources (datasets, algorithms, tools) and fundamental services. Aims to serve the broader public interest.
Governance and Ethical Framework	Values ethical and responsible use of resources, but with diverse governance structures focusing on open sharing.	Operates under a robust ethical framework. Stringent governance for data privacy, intellectual property rights, and ethical AI standards.
Infrastructure and Accessibility	Diverse infrastructure supporting various digital resources. Emphasis on sharing and collaborative improvement.	Requires sophisticated infrastructure for widespread AI resource sharing. User-friendly platforms for universal access.

Public Utility Mode	Digital Commons is not typically characterized as a public utility, but rather as a shared space for collaboration.	Functions as a public utility, emphasizing universal and equitable access to AI technology. Aims to serve public interest like traditional utilities.
	Digital Commons	AI Commons
Subset and Specialization	Encompasses a wide array of digital resources, including but not limited to AI.	Considered a specialized subset within the broader Digital Commons, with a specific focus on AI resources.
Complementary Roles	Promotes open access and collaboration across a variety of digital resources, not specifically service-oriented.	Focuses on AI as a fundamental service for public benefit, going beyond open sharing to universal service provision.

The integration of Data Cooperatives with Digital and AI Commons is crucial. Data Cooperatives provide ethically sourced, diverse data, essential for developing inclusive AI technologies within AI Commons. This collaboration along with appropriate policy support ensures ethical AI development, leveraging the shared, open-access nature of Digital Commons. Such an integrated approach promises a future where data and AI serve not just economic interests but the greater good of society, democratizing technology access and fostering responsible innovation.

Exploring the Potential of Data Cooperatives in Preserving Indigenous Languages through the Oki Language Project

Eugene Brave Rock

Language is the foundation of cultural identity, a vessel carrying the stories, knowledge, and worldviews of a people. The preservation of Indigenous languages, like those featured in the [Oki Language Project](#), is crucial for maintaining the cultural heritage and identity of Indigenous communities. The Oki Language Project aims to document and share the diverse greetings and stories of Indigenous Elders across Turtle Island (North America). One innovative approach to safeguarding these languages and stories could be through the establishment of a data cooperative, a model that aligns well with the values of collective stewardship and community empowerment.

What is a Data Cooperative?

A data cooperative is a collective organization where members pool their data for mutual benefit. In the context of Indigenous language preservation, this "data" refers to more tangible elements such as recordings, stories, and linguistic information related to the languages featured in the Oki Language Project. Unlike traditional data ownership models, where data is often controlled by large corporations, a data cooperative allows the community to maintain control and ownership over their data. Members decide how the data is used, ensuring that it aligns with their values and serves their interests.

In the context of the Oki Language Project, a data cooperative could involve Indigenous communities collectively managing and sharing the recordings, stories, and linguistic data related to their languages. This model would empower communities to preserve their languages on their terms while facilitating broader access to this invaluable cultural heritage.

Potential Benefits of a Data Cooperative for the Oki Language Project:

1. **Community Ownership and Control:** A data cooperative ensures that Indigenous communities maintain ownership of their linguistic data. This control is essential for safeguarding the cultural integrity of the languages featured in the Oki Language Project and ensuring that their use aligns with the community's values.
2. **Sustainable Preservation:** By pooling resources and data, communities can create a more sustainable model for preserving the languages documented by the Oki Language Project. The cooperative structure allows for shared responsibility in maintaining and updating the language database, ensuring it remains relevant and accessible over time.

3. Enhanced Collaboration: A data cooperative fosters collaboration among different Indigenous groups, researchers, and linguists. By working together, these stakeholders can share knowledge, develop new tools, and create educational resources that support the revitalization of Indigenous languages.
4. Ethical Data Use and [CARE Principles](#): With a data cooperative, the community decides how their linguistic data is used, preventing exploitation and ensuring that the data is used in ways that benefit the community. This approach is closely aligned with the CARE Principles for Indigenous Data Governance, which emphasize Collective Benefit, Authority to Control, Responsibility, and Ethics (Carroll et al. 2020). These principles ensure that data use respects Indigenous rights and interests, fostering innovation and self-determination. By adhering to the CARE principles, the cooperative would safeguard the cultural significance of the languages featured in the Oki Language Project.
5. Educational Outreach: The cooperative model can facilitate educational initiatives, making the languages documented in the Oki Language Project more accessible to younger generations and non-Indigenous allies. By providing a structured repository of language resources, the cooperative can support language learning and cultural education efforts.

Global Indigenous Data Alliance: A Key Stakeholder

To further strengthen the governance and ethical use of the data cooperative, the involvement of organizations like the [Global Indigenous Data Alliance \(GIDA\)](#) can be instrumental. GIDA advocates for Indigenous data sovereignty and the application of the CARE Principles in data governance. Their partnership could ensure that the data cooperative aligns with global best practices in Indigenous data management and governance, providing an additional layer of support and legitimacy.

Conclusion

The preservation of the languages documented by the Oki Language Project is not just about maintaining words and phrases; it is about safeguarding a way of life, a worldview, and the identity of a people. The Oki Language Project's efforts to document and share Elders' stories are invaluable, and the establishment of a data cooperative could further amplify these efforts. By leveraging the power of collective ownership, ethical stewardship through the CARE principles, and the support of organizations like GIDA, a data cooperative can play a crucial role in preserving Indigenous languages for generations to come, ensuring that their stories and wisdom continue to enrich our world.

References

Carroll, S., Garba, I., Figueroa-Rodriguez, O., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sra, R., Walker, J., Anderson, J. & Hudson, M. (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal*, 19: XX, 1-12.

Implementing Resilience in Scientific Research through Digital Asset Registries (DARs) and Data Cooperatives

Mei Lin Fung¹, Jascha Stein¹, Sean Masaki Flynn²

¹People Centered Internet

²Scripps College

The international recognition of Digital Asset Registries (DARs) and Data Cooperatives can significantly advance scientific research by providing secure frameworks for managing data ownership and ensuring that contributions to research are properly recognized. DARs, in combination with Data Cooperatives, enable the equitable and transparent management of the vast amounts of digital data generated in the modern era. Together, these structures can mitigate risks related to data mismanagement, while ensuring that individuals, communities, and organizations retain ownership of their data and feel confident contributing to the scientific enterprise.

Drawing on the historical development of land rights and the analogy of property title systems, we advocate for Digital Asset Registries as the digital equivalent of land title registries. Data Cooperatives, in this context, act as collective organizations that manage and govern the pooled data of their members, ensuring equitable access to the benefits of digital assets for all stakeholders. By working in tandem, DARs and Data Cooperatives foster a secure and inclusive environment for data sharing, thereby unlocking the potential for scientific progress.

Introduction

In the digital age, data is the foundation of scientific research and innovation. However, the rights to data—whether they belong to individuals, communities, or organizations—remain ill-defined and under-protected. This lack of clarity can result in exploitation or exclusion, especially for marginalized communities, and may inhibit the equitable distribution of scientific benefits.

DARs and Data Cooperatives offer a solution by formalizing data ownership and governance structures. DARs establish clear legal and technical frameworks for managing digital assets, while Data Cooperatives provide a mechanism for communities to collectively manage, share, and benefit from their data. Just as land title registries secured property rights in the physical world, DARs and Data Cooperatives together ensure that digital data is controlled and utilized fairly in the digital world.

Without these protections, data contributors risk losing control of their data, much like communities in history who lost their land due to unclear property rights. This is a lesson drawn from the historical experience of Native Americans, who were forced off their land due

to a lack of formal land rights. In today's digital era, we face a similar challenge: if we do not establish strong systems for data rights, entire communities could lose control over the data they generate, and more importantly, their art, culture, traditions - critical to social science.

AI, Data Cooperatives, and Linguistic Diversity

One of the most pressing challenges in the current AI landscape is that most AI systems are trained on data from only 4-5 major languages, despite the fact that almost 400 languages are spoken by more than 200,000 people. This limited training data risks creating a digital divide in AI, where the voices, art, traditions and experiences of many communities are not represented.

Data Cooperatives can play a critical role in addressing this issue by enabling communities to pool their digital resources in a way that retains local control while contributing to global research. By establishing DARs alongside Data Cooperatives, we can ensure that data from all languages and cultures is safeguarded and properly reflected in AI development. This would prevent the erasure of linguistic and cultural diversity in the digital age and ensure that AI systems reflect the experiences of the people who generate the data, and those who create art and experiences and transmit culture and traditions.

Without the implementation of DARs and Data Cooperatives, we risk repeating historical injustices, such as the [Trail of Tears](#), where land was taken from Native American communities without their consent (Perdue and Green 2004). At that time, many Native Americans did not know their consent was needed because they did not believe land could be owned in the way European settlers understood ownership. Land, for them, was a shared resource, integral to their way of life and identity, not something to be parceled and sold. This fundamental difference in perspective meant they were slow to respond, and eventually unable to assert their legal rights effectively, leading to the loss of their ancestral lands.

We face a similar situation today in the digital realm, where it is extremely difficult for individuals and communities to assert their digital rights or manage the data they generate. Much like the Native Americans' view of land, many people today are unaware that the digital data they generate—whether through online activity, personal devices, or community platforms—has value and rights that can and must be managed, and can be 'owned' or controlled. Even more concerning, many languages of the world don't even have a definition for data or digital rights. Without having conceptual framework for the digital dimension, these communities may unknowingly surrender control over their digital assets to larger, more powerful entities - in effect Digital Colonization will take effect just as the Dutch and British East India Companies and Spain created colonies in Asia, Africa and Latin America 300-400 years ago.

If we do not take steps now to implement DARs and Data Cooperatives, we risk repeating the same pattern of disenfranchisement, where people and communities are stripped of control over their data, just as they were stripped of their land. By establishing clear systems for managing digital rights, we can ensure that the data generated by all communities is protected and that the voices and contributions of these communities are recognized in the digital age

By securing data rights now, we can avoid a modern-day equivalent of the Trail of Tears, where data is taken without recognition of the rights of the people cultures and traditions. We can ensure that the global data ecosystem is inclusive and representative of all.

Digital Assets and Data Cooperatives: A Collaborative Approach

Digital assets, which include datasets, research outputs, and intellectual property, are central to the future of science. However, much like land in the past, these digital assets often lack clear ownership structures. Digital Asset Registries provide a formal mechanism for assigning and protecting ownership of these assets, but this effort must be complemented by Data Cooperatives, which allow communities to manage their data collectively.

Data Cooperatives offer a democratic approach to data governance, ensuring that communities retain control over their data and that the benefits of data-sharing are distributed equitably. By pooling data within a cooperative, members can enhance their collective bargaining power, ensuring that their data is used ethically and that their rights are respected. Digital Asset Registries, in turn, ensure that data ownership is formalized and that contributors have legal recourse if their rights are infringed.

The combination of DARs and Data Cooperatives creates a robust system for data governance that fosters both individual and collective rights, allowing scientific research to thrive in a fair and inclusive manner.

Lessons from Land Rights in the Netherlands

The development of DARs and Data Cooperatives can be likened to the historical evolution of land rights in the Netherlands. The Dutch system of Waterschappen (Water Utility Boards) provides an instructive model for managing both private property and collective resources. The Waterschappen were created to manage dikes and water systems that protected privately owned land from flooding, balancing individual property rights with collective security.

Just as the Dutch established local control over land and water, communities today need to establish local control over their digital assets. Data Cooperatives, like Waterschappen, are collectively governed bodies that manage data for the benefit of their members. Digital Asset Registries act as the formal system for recording and protecting ownership, ensuring that both individual and collective rights are upheld.

Together, DARs and Data Cooperatives provide the infrastructure needed to protect digital assets and unlock their value for science and society.

Digital Common Law for Data Governance

We propose a Digital Common Law system to guide the development of DARs and Data Cooperatives. Modeled on English Common Law, this system would allow local courts and

governing bodies to establish precedents in digital asset governance. As digital technologies and scientific practices evolve, this approach would enable policies and regulations to remain dynamic and responsive to the needs of communities and contributors.

By building a global framework for digital rights that respects local governance, DARs and Data Cooperatives can ensure that data is managed ethically and inclusively, allowing all communities to benefit from scientific progress.

Conclusion

The integration of DARs and Data Cooperatives is essential for the future of scientific research and the equitable management of digital rights. Together, these frameworks protect the rights of data contributors, ensuring that data can be shared and used for the common good while respecting individual and community oversight and rights.

By establishing these structures now, we can safeguard the linguistic and cultural diversity of the world in the age of AI, preventing a repeat of the historical injustices that occurred when land rights were ignored. DARs and Data Cooperatives offer a path toward a more inclusive, equitable, and innovative future for science, ensuring that the benefits of the digital age are shared by all.

References

Stein, Jascha, Mei Lin Fung, Marta Bertolaso, Fulvio Bersanetti, Suruchi Gupta, Ronald Strauß, Charlie Isaacs and Christine Asjoma (2024): [Improving Global Governance: Data Cooperatives for Global Cooperation](#). Think 7 Policy Brief for Japan G7.

Fung, M. (2023). Make AI Work for Everyone <https://www.edelman.com/trust/edelman-trust-institute/Publication/2023/AI-for-everyone>

Fung, M., Tobing, D. H., Bertolaso, M. & Potluri, V. A. (2023). [People-Centered Science and Digital Transformation: A Practical Proposal for the G7 and G20](#). Think 7 Policy Brief for Japan G7.

Fung, M, and Stein, J. (2023). [Addressing the Digital Divide and Mitigating the Risk of AI by People-Centered, Collaborative Digital Regulation](#). IEEE - HKN Bridge Magazine.

Gambacorta, L., Huang, Y., Li, Z., Qiu, H. & Chen, S. (2023). [Data versus Collateral](#), Review of Finance 27(2).

Perdue, T. and Green M. (2004). The Cherokee Removal: A Brief History with Documents. Bedford/st Martins; 2nd edition

Singh, T., Fung, M., Murray, E., Lacavaro, A., Gamser, M., Omwenga, B., Bitange, N., & Kharas, H. (2022). [Global Public-Private digital utilities for MSME recovery and transition](#). Think7 Policy Brief 2022.

Stein, J., Fung, M., Ndemo, B. & Flynn, S. (2023). [Enhancing Global MSME Financing through Prosperity Data Networks: An Integration of Hayek's and Sen's Economic Insights in the Digital Age](#). Global SME Finance Forum.

Park, K-S. (2023). [Property and Sovereignty in America: A History of Title Registries & Jurisdictional Power](#). Georgetown Law – The Scholarly Commons

An Understanding of Data Cooperatives When Viewed from the Lens of the 2023 YOUthDIG Messages

Izaan Khan

Internet Society Youth Standing Group

The European Dialogue on Internet Governance (“EuroDIG”) is an annual, pan-European, multi-stakeholder initiative focused on driving the discussion of pressing Internet governance issues, and forms part of the broader Internet governance ecosystem through its upward inputs to the global United Nations Internet Governance Forum (“UN IGF”). Established in 2008 as the first such regional initiative, since 2017 the EuroDIG programme has also included a youth programme, the Youth Dialogue on Internet Governance (“YOUthDIG”), which facilitated the introduction to - and subsequent involvement of - individuals aged 18-30 in the ecosystem.

The inclusion of youth voices within the broader Internet governance debate has been widely recognised as being crucial for upholding the true spirit of multi-stakeholderism and for understanding the demographic’s key concerns for the future Internet that they inherit as a result of the policies of today. Within the European context, this occurs through the YOUthDIG Messages process, where the participants of the programme come together to identify their areas of interest or concern, and develop policy suggestions that are presented to an audience composed of legislators, technologists, civil society organizations and the community at large.

In 2023, the YOUthDIG messages encompassed four key areas: Artificial Intelligence, the Digital Divide, Data Governance, and Evidence-based Cooperative Internet Governance. The unifying factor in these topics is ensuring that individuals are not betrayed by the promise of the Internet and new technologies, and that stakeholders currently in positions of power or influence should do more to increase the individual’s agency rather than further marginalizing their presence and involvement in discussions and deployment of technologies.

Further, there is an explicit understanding of the importance of the role that data (particularly personal data) plays in relation to these areas, which lends itself conveniently to an analysis of how data cooperatives can offer a potential solution to the broader issues raised. The urgency for this type of analysis cannot be understated, especially from a youth perspective, as there have been a number of high-profile examples where personal data of young individuals was not appropriately handled in line with existing rules. One needs to look no further than Meta’s €405m fine for publicizing teenager contact information as they were freely allowed to create business accounts, or TikTok’s €345m fine for similarly making content posted by children public by default, both instances occurring only in the last two years.

Data cooperatives are an attractive proposition to counter these issues, as they are deliberately designed to act in the fiduciary interest of their members, and act as stewards in relation to the data that they collect that is generated from the online activity of their members. As such, in relation to the issue of data mishandling, the bargaining power of the cooperative can be leveraged to engage with companies on fairer terms, and to only provide as much information as the members of the cooperative are comfortable with. The institutional design of such cooperatives also ensures that individuals are actively involved in the decision-making processes over their data which naturally leads to more privacy and agency, rather than simply accepting the terms provided to them as in the status quo. This is very much in line with the YOUTHDIG messages on bridging the digital divide, recognising it as more than just an issue of access to new technologies, but also as a social phenomenon. As quoted: “[We] Urge all relevant stakeholders to collaboratively prioritize the inclusion of marginalized voices in the decision-making process with the aim to create an inclusive digital landscape in Europe.”

In order to make data cooperatives a sustainable alternative, it is recognised that there need to be technical and institutional frameworks and mechanisms to support them. This was also captured within the YOUTHDIG messages, as a large number of them are calls to action aimed at governments and policymakers to create a data environment that is conducive to such innovation, while also providing more education and capacity-building to smaller stakeholders, and which is ultimately grounded on human rights. Data cooperatives need more attention for research and experimentation to understand how they can fit in and add value within the digital ecosystem, and a first step would be to create regulatory and technological sandboxes where a number of interested parties can test their ideas with a view to developing common standards for how data cooperatives can pull, process and share data in an interoperable way whilst in line with the potentially varied needs of their members. Next, legislators can consult with these stakeholders to understand how the power of these cooperatives can be effectively utilized so as to level the playing field between them and larger institutions via legislation, codes of conduct, or standard commercial contracts which can be used as the basis for engagement between cooperatives and their counterparts. Lastly, ensuring that this exchange of data does not commoditize the individual from whom it originates, and ensuring that human rights are continually underlined throughout will be critical to prevent privacy from being a luxury good. Indeed, as another YOUTHDIG message emphasizes: “All stakeholders involved in regulation or development of the Internet should conduct thorough impact assessments to identify the effects of their activities on the Internet and the user experience in order to protect security and privacy and avoid fragmentation.” With these points in mind, we can aim to make data cooperatives a new foundation for a fairer Internet.

Fostering an African Data Commons

Marie Shabaya, Simon Sun and Nai Lee Kalema

A version of this paper was presented at the African Society for International Law Conference in October 2023 held at the African Union Headquarters in Addis Ababa, Ethiopia.

The proliferation of Artificial Intelligence (AI) has propelled the global data economy to new heights, yet this remarkable progress has also precipitated significant disparities within the digital realm. At the heart of this debate is the issue of data governance as the symbiotic relationship between data and AI is indisputable. Ample data fuels AI algorithms, underpinning their predictive capacities and emphasising data's indispensable role in driving these advancements. Although the African continent holds immense potential for leveraging AI, the progress of data governance in the region has been gradual. In the African context, the primary obstacle to data sharing stems from the power asymmetries ingrained in the colonial system, commonly referred to as “data colonialism.” Currently, Western tech giants such as Amazon, Google, and Facebook wield considerable control over Africa's digital infrastructure. Fundamentally, the challenge arises from the existing data governance framework which is rooted in the philosophy of “privacy,” an inherently Western concept that presents numerous complexities for African citizens in comprehending its associated rights. Our innovation derives from two philosophical and theoretical frameworks to aid in constructing a data governance infrastructure for the African continent, for posterity.

First: The pan-African philosophy of Ubuntu which frames a logic inherent in a number of sub-Saharan countries and its myriad Bantu sub-cultures. A modern axiom comes from South Africa where various Zulu language sub-groups relate it as ubuntu ngumuntu ngabantu (I am because of who we all are). Essentially, these ideas underpin societal relationships and civic duty with the foundational understanding that humanity is, at its very core, interconnected. This ethos has numerous implications for application within AI, and its regulation, particularly as new tools are being developed for use within African jurisdictions. For the purposes of this argument, Ubuntu frames a value system (Mhlambi 2019) which can be used as a lens through which to assess practical frameworks for the use of common pool resources, as pioneered by Elinor Ostrom.

Ostrom provided four fundamental contributions to the economic governance literature: Firstly, substituting the term “rivalry of consumption” with “subtractability of use” (Ostrom 2010). Secondly, suggesting a conceptual shift that involves considering subtractability of use and excludability as variables that can range from low to high, rather than categorising them as binary attributes of either present or absent. Thirdly, introducing a fourth category of goods—common-pool resources. These goods share the attribute of subtractability with private goods and the challenge of exclusion with public goods. Noteworthy examples of common-pool resources encompass forests, water systems, fisheries, and the global atmosphere. In 2009, Ostrom became the first woman to receive the Nobel Memorial Prize in Economic Sciences for this body of work.

		Level of Subtractability	
		High	Low
Difficulty of exclusion	High	Common-pool resources (e.g. street trees)	Public goods (e.g. public parks with trees)
	Low	Private good (e.g. private trees)	Toll good (E.g. private parks with trees)

Source: Adapted from Ostrom, 2005, p.24

Our proposal relies on the core assumption that data can be (and should be) treated as a Common-Pool Resource (CPR). We contend that treating data as a CPR would establish a healthy data ecosystem for Africa. Currently, most data centres in Africa are run by foreign service providers located overseas (Raji 2022). For example, a significant portion of Nigerian government agencies, approximately 70 percent, currently opt to store their data on servers located outside of the country. This choice is often motivated by several factors, including cost considerations, the reliability of overseas data hosting services, and the capacity to accommodate large volumes of data storage. These factors continue to influence the preference of many African firms and government entities to maintain their data abroad.

In this manner, private investment has the potential to expand the reach and the breadth of practical and potential applications of open data (Lindman and Kuk 2015). As explained by Juho Lindman and George Kuk, the extension of reach by private actors can manifest through several means:

1. commercial companies may opt to donate datasets and resources for ongoing maintenance,
2. private individuals could contribute their own datasets, or
3. novel data complementarities could be established. The goal is to formulate effective incentives for various types of contributions.

It is worth noting that, while the private sector and academia have maintained a relatively passive role within the African open data ecosystem, primarily focused on internal data generation and research-driven objectives, a notable shift is taking place (Mutuku and Tinto 2019). In recent times, both the private sector and academic institutions are displaying a growing interest in actively contributing to the sustainable development agenda by becoming integral partners in the expansion of the open data ecosystem. These endeavours underscore the evolving role of academia as a proactive player in the African open data landscape. Similarly, the private sector, too, is recognizing its potential to actively participate in fostering the growth of the open data ecosystem as part of its commitment to sustainable development objectives. As the private sector increasingly embraces this role, there are emerging instances showcasing the valuable contributions it can make to the open data landscape.

References

Lindman, J. and Kuk, G. (2015). [From Open Access to Open Data Markets: Increasing the Subtractability of Open Data](#) 48th Hawaii International Conference on System Sciences, Kauai, HI, USA

Mhlambi, S. (2019). [Ethical Implications of AI and Ubuntu as an Intervention](#). Berkman Klein Center. August 29, 2019.

Mutuku, L. & Tinto, T. (2019). Open Data Around the World - Sub-Saharan Africa In: T. Davies, S. Walker, M. Rubinstein, & F. Perini (eds.), *The State of Open Data: Histories and Horizons*.

Ostrom, E. (2010). [Beyond Markets and States: Polycentric Governance of Complex Economic Systems](#). *American Economic Review* 100(3), June 2010, 641 - 72.

Raji, R. (2022). [Where is Africa in the cloud?](#) NTU-SBF Centre for African Studies (CAS), September 15, 2022.

Digital Horizons: Empowering Youth in the World of Access, Literacy, and Opportunities

Natálie Terčová

Internet Governance Forum (IGF) Czechia

Access to entertainment. Access to communication. Access to employment.

These are just a glimpse of the myriad offerings facilitated by information and communication technologies (ICTs). Infiltrating our daily lives, these tools operate so seamlessly that their true reach often eludes our awareness. The ability to navigate this digital landscape has evolved into a fundamental skill set, even a necessity, essential for keeping pace with our ever-changing, dynamic world. This aptitude is encapsulated in the concept of digital literacy.

While the term 'digital natives' is often ascribed to the younger generation born after 2000, presuming innate digital savviness, such assumptions warrant reconsideration. The belief that all children grow up in an environment saturated with ICTs is naive, accounting for disparities not only across countries but also within households of varying socio-economic statuses. Access to a range of technologies significantly influences an individual's familiarity with basic principles and skills. Hence, it is imperative not to underestimate the necessity of imparting digital skills to the younger generation. Digital skills encompass a diverse array of competencies, spanning technical proficiency, effective information navigation, adept communication and interaction in digital environments, and proficiency in creating and producing digital content. These competencies exhibit uneven distribution across individuals, communities, and nations. Their prevalence and development are intricately shaped by factors such as individual aptitude, aforementioned socio-economic conditions, educational opportunities, and regional or national contexts. A deficiency in digital skills or neglecting their enhancement significantly curtails opportunities for personal and professional growth, as well as active societal engagement (d'Haenens et al. 2023). To ensure digital literacy among young people, experts advocate a collaborative effort involving the formal education system, the private sector, and the research community. A systemic approach to digital literacy development, addressing both connectivity and quality education, is deemed crucial in overcoming the digital divide on its two levels—(1) access to digital technologies, and (2) gaps in usage and skills.

However, the inequalities go further. The third level of the digital divide delves into the differential outcomes resulting from ICTs use. Despite possessing the same set of digital skills and access to technology, individuals derive varied experiences from online interactions. Generalizations about the harms or benefits of technology are thus deemed inappropriate. Recognizing the diverse content encountered online, differing motives for seeking it out, and variable responses to potential risky materials emphasizes the need for nuanced understanding. Education and open dialogue through active mediation, particularly with young people and children, are identified as the key starting points for addressing such digital literacy challenges. On the contrary, the restrictive approach, marked by bans and online constraints, has proven ineffective, resulting in a dearth of youth's digital skills

development (Smahel et al. 2023). Building on this, we should recognize that escaping all online risks is an unrealistic goal. Understanding the limitations in shielding young people from every potential threat, our collective emphasis shifts. Instead of pursuing an unattainable escape, we should redirect our focus to nurturing their collective capacity to employ digital skills for self-protection and secure navigation in the online realm. This, coupled with fostering the development of effective coping strategies, would equip young individuals to navigate challenges adeptly while concurrently mitigating online risks. Such approach would empower them not only to safeguard themselves but also to harness the myriad online opportunities. In pursuit of these objectives, data cooperatives can play a pivotal role by actively advocating for policies that champion digital literacy, taking into account barriers related to all three levels of digital divides. Through collaborative efforts, they may contribute to effectively communicating the latest findings, dispelling myths, and addressing associated needs to foster innovation that is both impactful and inclusive.

The online world, a vast and enchanting yet unknown and dark realm, increasingly blurs the boundaries between itself and the offline world. Rather than turning a blind eye, it is prudent to prepare and harness the tremendous opportunities it offers. Such as the access to entertainment. Access to communication. Access to employment.

References

- d'Haenens, L., Joris, W., & Bossens, E. (2023). Synthesis of ySKILLS results. KU Leuven: ySKILLS.
- Smahel, D., Mascheroni, G., Livingstone, S., Helsper, E. J., van Deursen, A., Tercova, N., Stoilova, M., Georgiou, M., Machackova, H., & Alho, K. (2023). Theoretical Integration of ySKILLS: Towards a New Model of Digital Literacy. KU Leuven: ySKILLS.

List of Publications on Data Cooperatives

Aapti Institute, & Data 2X (2023, May). *Data Cooperatives: A Real-World Roadmap for Social Impact*. Youtube.

Aguirre, A., Clomax, A., DiAntonio, N., & Parikh, S. (2021). *Should We Certify Platform Cooperatives?*, White Paper created as part of The New School's Platform Cooperativism Consortium and Harvard University's Berkman Klein Center for Internet & Society Research Sprint.

NITI Aayog (2020). [Data Empowerment And Protection Architecture - A Secure Consent-Based Data Sharing Framework To Accelerate Financial Inclusion](#). August 2020, NITI Aayog.

Banterle, F. (2018). *Data Ownership in the Data Economy: A European Dilemma*. EU Internet Law in the digital era, Springer.

Bayamlioglu, E. (2021). *Data cooperative: A new intermediary on the horizon*. KU Leuven, Center for IT and IP Law.

Berkman Klein Center (2022). [What we learned from our research sprint on cooperative data governance](#). Berkman Klein Center for Internet & Society.

Bietti, E., Etxeberria, A., Mannan, M., & Wong, J. (2021). [Data cooperatives in Europe: A legal and empirical investigation](#) (The New School's Platform Cooperativism Consortium) [White paper].

Blasimme, A., Vayena, E., & Hafen, E. (2018). Democratizing health research through data cooperatives. *Philosophy & Technology*, 31, 473–479.

Bühler, M. M., Nübel, K., Jelinek, T., Riechert, D., Bauer, T., Schmid, T., & Schneider, M. (2023). Data cooperatives as a catalyst for collaboration, data sharing and the digital transformation of the construction sector. *Buildings*, 13(2), 442.

Bühler, M. M., Calzada, I., Cane, I., Jelinek, T., Kapoor, A., Mannan, M., Mehta, S., Mookerjee, V., Nübel, K., Pentland, A., Scholz, T., Siddarth, D., Tait, J., Vaitla, B. & Zhu, J. (2023). [Unlocking the Power of Digital Commons: Data Cooperatives as a Pathway for Data Sovereign, Innovative and Equitable Digital Communities](#). *Digital 3*(3), 146-171.

Bühler, M. M.; Calzada, I.; Cane, I.; Jelinek, T.; Kapoor, A.; Mannan, M.; Mehta, S.; Micheli, M.; Mookerjee, V.; Nübel, K.; Pentland, A.; Scholz, T.; Siddarth, D.; Tait, J.; Vaitla, B.; Zhu, J. Data cooperatives as catalysts for collaboration, data sharing, and the (trans)formation of the digital commons. Policy Brief. Task Force 2: Our Common Digital Future: Affordable, Accessible and Inclusive Digital Public Infrastructure. In Think20 (T20) India. 2023. Calzada, I. (2021). Data co-operatives through data sovereignty. *Smart Cities 4*(3): 1158-1172.

Carovano, G., & Finck, M. (2023). Regulating data intermediaries: The impact of the Data Governance Act on the EU's data economy. *Computer Law & Security Review*, 50, 105830.

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal*, 19.

Data Governance Act - Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724. (2022, May 30). Official Journal of the European Union.

Fung, M. L., Herdiawan Tobing, D., Bertolaso, M., & Aditya Potluri, V. (2023). People-Centered Science and Digital Transformation: A Practical Proposal for the G7 and G20. In Think7, G7 Japan 2023.

Fung, M. L., Stein, J., Bertolaso, M., Bersanetti, F., Gupta, S., Strauß, R., Isaacs, C., & Asjoma, C. (2024). Improving Global Governance: Data Cooperatives for Global Cooperation. In Think 7, G7 Italia 2024, Policy Brief.

Fink, A. (2024): [Data Cooperative](#). *Internet Policy Review* 13 (2).

Hafen, E., Kossmann, D., & Brand, A. (2014). Health data cooperatives—Citizen empowerment. *Methods of Information in Medicine*, 53(02), 82–86.

Hardjono, T. and Pentland, A. (2019): [Data Cooperatives: Towards a Foundation for Decentralized Personal Data Management](#). May 21, 2019.

Mannan, M., Bietti, E., Etxeberria, A., & Wong, J. (2019). Data cooperatives in Europe: A legal and empirical investigation. *Computer Law & Security Review*, 35(5), 105297.

Mannan, M., Wong, J. & Bietti, E. (2022). [Data Cooperatives in Europe: A preliminary Investigation](#). *Network Industry Quarterly* 24(3).

Micheli, M., Farrell, E., Carballa-Smichowski, B., Posada-Sánchez, M., Signorelli, S., & Vespe, M. (2023). [Mapping the landscape of data intermediaries: Emerging models for more inclusive data governance](#). [Report]. Publications Office of the European Union.

Miller, K. (2021). [Radical proposal: Data cooperatives could give us more power over our data](#). *Stanford University Human-Centered Artificial Intelligence*. October 20, 2021.

Mulgan, G., & Straub, V. (2019). [The new ecosystem of trust: How data trusts, collaboratives and cooperatives can help govern data for the maximum of public benefit](#) [Report]. Nesta.

Pentland, A., Hardjono, T., Penn, J., Coclough, C., Ducharme, B., & Mandel, L. (2019). [Data cooperatives: Digital empowerment of citizens and workers](#) [White paper]. MIT Connection Science.

Pentland, A., and Hardjono, T. (2021). Building data cooperatives. In Pentland, Alex, Lipton Alexander, and Hardjono, Thomas, eds. *Building the new economy: Data as capital*: 19-34. Cambridge: MIT Press.

Platform Cooperativism Consortium (2022). Data Cooperatives – Presenting Results from the PCC/BKC Research Sprint | #TheNewCommonSense [Video]. YouTube

Salau, A., Dantu, R., Morozov, K., Upadhyay, K., & Badruddoja, S. (2022). Towards a threat model and security analysis for data cooperatives: Proceedings of the 19th International Conference on Security and Cryptography, 707–713.

Salau, A., Dantu, R., & Upadhyay, K. (2021). Data cooperatives for neighborhood watch. 1–9.

Scholz, T., & Calzada, I. (2021). Data cooperatives for pandemic times. Public Seminar.

Scholz, T. R. (2023): [Own This!: How Platform Cooperatives Help Workers Build a Democratic Internet](#). Verso Books.

U. N. Executive Office of the Secretary-General. (2023). A global digital compact — An open, free and secure digital future for all. In UN-EOSG Policy Briefs and Papers.

Stein, J., Fung, M. L., Bitange, N., & Flynn, S. (2023). [Enhancing Global SME Financing through Prosperity Data Networks: An Integration of Hayek’s and Sen’s Economic Insights in the Digital Age](#). SME Finance Forum.

Tanwar, A. S., Evangelatos, N., Venne, J., Ogilvie, L. A., Satyamoorthy, K., & Brand, A. (2021). Global open health data cooperatives cloud in an era of COVID-19 and planetary health. *OMICS: A Journal of Integrative Biology*, 25(3), 169–175.

UNStats. (2023). (TA1.02) [Data Cooperatives: A Real-World Roadmap for Social Impact](#) [Video]. YouTube.