# Socializing Data Value

## Reflections on the State of Play

Synthesis of themes from a roundtable
organized by IT for Change
with support from Friedrich-Ebert-Stiftung, India office

July 2021

# Credits

# Socializing Data Value

## Reflections on the State of Play

Synthesis of themes from a roundtable organized by IT for Change
with support from Friedrich-Ebert-Stiftung, India office

July 2021

*This page left intentionally blank*

# Table of Contents

# **Roundtable Participants**

- Aaron Martin, Postdoctoral Researcher, Tilburg Law School
- Aditya Singh Chawla, Doctoral Researcher, Centre for Technomoral Futures, University of Edinburgh
- Amay Korjan, Research Associate, IT for Change
- Amshuman Dasarathy, Research Assistant, IT for Change
- Anita Gurumurthy, Executive Director, IT for Change
- Anuradha Ganapathy, Senior Program Manager, IT for Change
- Anurag Shanker, Program Manager, Friedrich-Ebert-Stiftung India Office
- Arindrajit Basu, Research Lead, Centre for Internet and Society, India
- Astha Kapoor, Co-founder, Aapti Institute
- Barbara Prainsack, Professor, Department of Political Science, University of Vienna
- Bruno Carballa Smichowski, Researcher, Joint Research Centre of European Commission – Unit B6 – Digital Economy
- Burcu Kilic, Director, Digital Rights Program & Research Director, Access to Medicines, Public Citizen
- Cecilia Rikap, Tenure researcher at CONICET, Argentina's National Research Council; Visiting professor of the IDHES, Université Paris Saclay
- Chee Yoke Ling, Executive Director, Third World Network
- Freyja van den Boom, PhD candidate Bournemouth University and Research Fellow, Weizenbaum Institute
- Gurumurthy Kasinathan, Director, IT for Change
- Ingrid Schneider, Professor, Department of Political Science, University of Hamburg
- Kristina Irion, Associate Professor, Central European University; and Marie Curie Fellow at Institute of Information Law, University of Amsterdam
- Mandvi Kulshreshtha, Program Adviser, Economy of Tomorrow Project, Friedrich-Ebert-Stiftung (FES), India office
- Marina Micheli, Researcher, Project DigiTranScope, Centre for Advanced Studies of the Joint Research Centre (Ispra, Italy)
- Nadezhda Purtova, Associate Professor, Tilburg Law School, Tilburg University
- Nandini Chami, Deputy Director, IT for Change
- Parminder Jeet Singh, Executive Director, IT for Change
- Paul-Olivier Dehaye, Founder, PersonalData.IO
- Raymond Onuoha, Research Fellow, Lagos Business School, Nigeria; and Doctoral Student, Nelson Mandela School of Governance, University of Cape town, South Africa
- Robert (Bob) Fay, Managing Director – Digital Economy, Centre for International Governance Innovation (CIGI)
- Salomé Viljoen, Affiliate, Berkman Klein Centre for Internet and Society, Harvard University
- Sarah Ganter, Friedrich-Ebert-Stiftung, Berlin

- Shamel Azmeh, Lecturer, Global Development Institute (GDI), University of Manchester
- Siddharth de Souza, Postdoctoral Researcher, Tilburg Law School
- Simonetta Vezzoso, Adjunct Professor, University of Trento
- Sneha Bhagwat, Communications Assistant, IT for Change
- Stacco Troncoso, P2P Commons Advocate and Organizer, Founder of DisCO and Gureillia Translation, Project Lead at Commons Transition
- Stefaan Verhulst, Co-Founder and Chief Research and Development Officer, The Governance Laboratory at New York University
- Tanvi Kanchan, Communications Assistant, IT for Change
- Trebor Scholz, Founding Director, Institute for the Cooperative Digital Economy, The New School
- Vinay Narayan, Research Assistant, IT for Change
- Yogesh K S, Senior Project Associate – Technology, IT for Change

*For more details about the participants, [visit this page](visit this page).*

# List of Acronyms

API: Application Programming Interface

CPR: Common-Pool Resource

DC: Data Cooperatives

DSP: Data Sharing Pools

DSB: Digital Stability Board

EDPS: European Data Protection Supervisor

EU: European Union

FAANG: Facebook, Amazon, Apple, Netflix, Google

GDPR: General Data Protection Regulation

HMB: Harm Mitigation Bodies

IP: Intellectual Property

NYU: New York University

PDT: Public Data Trusts

PIMS: Personal Information Management Systems

TRIPS: Trade-Related Aspects of Intellectual Property Rights

WTO: World Trade Organization

# Introduction

The past year has seen an increasing buzz around the need to address prevailing deficits in the governance of the digital economy. With data becoming central to market power, reining in the often monopolistic control of platform firms over society's data has assumed urgency. Current regulatory efforts like competition law reform, data protection laws and taxing digital services may succeed in tempering the concentration of digital power to some extent. However, they do not address the core problem of who exercises control over the immense socio-economic value of data. 70% of the new value created in the global economy over the next decade is likely to be generated from data capital-intensive platform businesses.[1] Yet, the platform model as we know it has only spawned an exponential increase in inter-firm and capital-labor inequality.[2] It is built on the grand premise that data is no one's property, a free resource that is available to all. However, the truth is that if newcomer firms are unable to effectively compete with dominant players such as FAANG, it is because the latter's data ecosystems are enclosed. Rules about the governance of data ecosystems are, therefore, crucial to reappropriate the public and social value of data for economic development and societal wellbeing. This is important not just for national digital development and economic security, but also for political sovereignty and socio-cultural independence and diversity.

Recognizing the systemic nature of data's resource structure is a precursor to such institutional rule-making. A bold institutional design based on this shared starting point can provide the basis for collective claims over data, and a just and equitable future for all. There is, thus, a critical need for theoretically and normatively grounded thinking around the issue of data access and claims, as well as the building blocks of a robust data governance framework.

It is towards this goal that IT for Change and Friedrich-Ebert-Stiftung, India office hosted a virtual roundtable with a small group of scholars, scholar-activists and practitioners on May 11 and 12, 2021. Titled 'Socializing Data Value: How Can Data Governance Meet the Challenge?', the roundtable explored the contestations characterizing the current economy and its data (dis)order, existing and emerging governance responses, policy visions and alternative practices that seek to socialize data value, as well as the key elements of an institutional governance framework for data that can democratize value creation and distribution.

**Rules about the governance of data ecosystems are crucial to reappropriate the public and social value of data for economic development and societal wellbeing.**

---

1    World Economic Forum, Shaping the Future of Digital Economy and New Value Creation.

2    UNCTAD, Digital Economy Report 2019.

**There is a critical need for theoretically and normatively grounded thinking around the issue of data access and claims, as well as the building blocks of a robust data governance framework.**

In this report, we have gathered some of the highlights from the roundtable's various discussions, and organized them thematically to provide a snapshot of the event and its key debates.

The content presented here has been put together from both the written inputs that participants submitted beforehand as well as the presentations and conversations that occurred during the roundtable. The report also draws from a wider body of literature, including contributions made by participants in their published works.

This report is intended to be a selective curation of important ideas and arguments, and by no means is a comprehensive representation of views of participants. Nonetheless, we hope it can provide a flavor of the rich proceedings and the critical questions and possible answers that were discussed during the two days about democratizing data value.

# 1. ASYMMETRIES OF THE DIGITAL ECONOMY

The status quo
## ADVANTAGES
### BIG TECH
at the expense of small businesses,
consumers and nation-states from
the Global South

Current individualist approaches to
data governance remain limited in
scope and cannot address
## DATA'S
### ECONOMIC VALUE

The data economy is mired in
## ASYMMETRIES
### OF POWER

The absence of de jure ownership
frameworks for data allows
de facto domination by a
### HANDFUL OF
## DIGITAL FIRMS

Power imbalances are
exacerbated by an unjust,
## INTERNATIONAL
### IP REGIME

- The digital economy is heavily skewed in favor of certain firms who control data value chains.

- Asymmetries of power – between consumers and Big Tech companies, smaller economic actors and large tech firms, nation-states and transnational tech corporations, economies of the South and those of the North – characterize the digital economy.

- Big Tech corporations exploit and cement an unjust macroeconomic governance regime in trade and IP, cornering value from data-driven intelligence rents.

- The absence of a de jure governance regime establishing ownership rights in data resources has prevented the emergence of fair data markets.

- Individualist approaches to data governance, such as the European Union's General Data Protection Regulation (GDPR), are based on a simplistic yardstick of privacy and anonymization, failing to address the de facto privatization and enclosure of data's economic value.

Set out below are some of the key ideas that emerged on the deeply unequal relationships that constitute the current state of the data economy, reflecting an urgency for governance mechanisms that redistribute the benefits of data value chains.

## Intelligence advantage and market power

Both **Barbara Prainsack** and **Ingrid Schneider** referred to a particular type of power asymmetry – that between people on the one hand and powerful corporate data collectors on the other.[3] Schneider noted that consumer data is used by platforms to create digital profiles, and huge user numbers help to create market power by providing the data for innumerable digital profiles.[4] This gives platforms a much greater bargaining power than individuals, further deepening an already existing asymmetry. This point was supported by **Kristina Irion's** assertion in her input paper, that data production and extraction take place inside walled gardens of online platforms so that data accumulates and concentrates in the hands of a few firms, which then commodify and monetize it, while excluding everyone else from its potential benefits. Schneider also pointed to the asymmetry at the global level,

> Data production and extraction take place inside walled gardens of online platforms so that data accumulates and concentrates in the hands of a few firms.

---

3    Barbara Prainsack, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

4    Ingrid Schneider, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

with most countries feeling like they are sandwiched between the data-extraction operations of libertarian US firms and the totalitarian Chinese transnational giants.

**Arindrajit Basu** noted that the global digital economy today is both structured around and dependent on entrenched power asymmetries, further augmented by the continued assertion of control by entities wielding power. **Paul-Olivier Dehaye** elaborated on this, directing attention to how actors that are able to pool large and diverse data sets seize market power, putting data subjects and smaller competitors within the data economy at a huge disadvantage. Combining the volume of data collected (which has a bearing on the quality of the data) and the variety of data collected (which adds to its exclusivity), they consolidate their position.

## An Unjust Macroeconomic Regime and Big Tech Hegemony

The untenable power imbalance in the digital economy has empowered Big Tech firms and countries in the developed world to write the rules of the global market in a manner that can ensure a cementing of existing power structures.

As Basu put it, "Legal, political and social structures amplify and enable these asymmetries, as entities in power – nations of the developed world, large multinational corporations, and global governance bodies – make rules that all the others in the global ecosystem must abide by. Trade agreements, taxation clauses and municipal laws are all brokered by and scripted for the benefit of a limited set of actors that hold the keys to the global digital economy."[5]

**Chee Yoke Ling** too drew attention to the unilateral imposition of digital rules through inter-governmental negotiations in e-commerce. These rules, she contended, are pushed under the guise of "building a strong and robust digital economy", but in reality serve to provide the technology giants with more rights. This push entails active advocacy at the World Trade Organization (WTO), where e-commerce negotiations are ongoing among a group of countries,[6] and at the regional/bilateral levels, whereby trade agreements and economic partnerships are building a new set of rules opposed to the rights of indigenous peoples and local communities. These rules will set the legal norms for decades to come, and it is imperative that we push back on these efforts now.[7]

The untenable power imbalance in the digital economy has empowered Big Tech firms and countries in the developed world to write the rules of the global market in a manner that can ensure a cementing of existing power structures.

---

5    Arindrajit Basu, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.
6    There is no consensus, as required by WTO rules, for negotiations to take place on a new subject matter – in this case, e-commerce.
7    Chee Yoke Ling, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

The modern intellectual property (IP) regime, bolstered by the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) at the WTO, has also played its part in allowing Big Tech firms to build up their market power. **Cecilia Rikap** detailed how the tech giants, mostly coming from the US (and some, from China), have monopolized critical data sources that are then used to extract data-driven intellectual rents. The result is a process whereby value is cornered by tech giants, even as the whole world produces data.[8]

Highlighting the global crisis over access to Covid-19 vaccines, Yoke Ling drove home the point of how virus sequence data is accessed freely from open databases and then the products developed therefrom are locked up in proprietary claims through the legal tools of patents and trade secrets. She observed how, in a statement to the WTO dated March 22, 2021,[9] several organizations and individuals had noted the damning consequences of intellectual property laws that prevent researchers in developing countries from using the most advanced methodologies for developing treatments. Despite support from various nations, experts and activists, the proposal for a TRIPS waiver issued by India and South Africa on Covid-19 vaccines and other medical products[10] was blocked by a number of countries from the Global North, including the EU and the US. While the US has reversed its stance, though limited to vaccines, the EU remains recalcitrant. However, the momentum for the waiver has grown with 63 co-sponsors and another 40 countries in support (as of June 30, 2021).[11]

## Inadequacies in Current Approaches to Data Governance

Exponential growth in the generation, collection, use and sharing of data has seen numerous calls for governance regimes to be put in place for the data economy. While most nations are working to come up with specific models that will help address asymmetries in data access and control, current models and their underlying imaginaries have only furthered the data power of Big Tech.

> Exponential growth in the generation, collection, use and sharing of data has seen numerous calls for governance regimes to be put in place for the data economy.

8    Cecilia Rikap, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

9    Statement on Copyright and Proposal of a Waiver from Certain Provisions of the Trade-Related Aspects of Intellectual Property Rights (TRIPS) Agreement for the Prevention, Containment and Treatment of COVID-19, Mar 2021.

10   Anoo Bhuyan, "Rich nations may use patents to cut off easy access to Covid-19 vaccines in developing countries", Scroll.in Nov 2020.

11   D. Ravi Kanth, Developing countries remain upbeat on TRIPS waiver negotiations, Third World Network, Jun 2021.

As **Bruno Carballa Smichowski**[12] explained, "Data governance models rely on a data collector (e.g., a platform) retaining exclusive control over the data it collects, typically through draconian clickwrap general conditions of use, particularly when the data collection involves individuals. After the data is collected, given that there is no such thing as de jure data ownership, the data collector 'owns' it de facto, although there are legal ways to protect third parties from accessing the data (yet not the data itself) through copyright over the database and/or the software that allows access to it."[13]

There have been efforts, at national and inter-governmental levels, to address the specific issues of the data economy, but these do not really measure up. Rikap pointed out that data privacy laws, although aimed at limiting tech giants' power, further contribute to knowledge privatization by fostering individual property over data.[14] Further, Irion noted that the "current governance frameworks of data trade have failed to produce transparent and functioning data markets. We only have an ad-hoc and incomplete picture of the trade and flows in personal data, for instance, and therefore, it is impossible to ensure that individual rights are not breached in the course of, or as a result of, such transactions. Quite to the contrary, there are indications of irregular and shady data markets while regular practices of data sharing and trade are underdeveloped."[15]

**Part of the inadequacy in current approaches to data governance stems from their being circumscribed within a selective framing of the problem.**

Moreover, part of the inadequacy in current approaches stems from their being circumscribed within a selective framing of the problem. That is, they begin with individual data subjects and try to remedy the situation by formulating rights that such subjects ought to have over their data. Such 'individualist' approaches comprise the mainstay of data governance today, and can be exemplified by the EU General Data Protection Regulation (GDPR), which grants citizens quasi-ownership rights in their personal data. This includes the right to determine if, and on what terms, their personal data enters the data market through a notice-and-consent regime, within the boundaries for the data market.

---

12  Bruno Carballa Smichowski participated in the roundtable in his capacity as an economist, and the views he expressed were his own. They do not represent the Joint Research Centre or the European Commission.

13  Bruno Carballa Smichowski, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

14  Cecilia Rikap, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

15  Kristina Irion, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

As **Anita Gurumuthy and Nandini Chami** argue in their paper proposing a "distributive integrity" framework for the data economy,[16] individualist approaches suffer from a number of shortcomings. To begin with, apart from a couple of special cases, the GDPR does not allow for citizens to make claims on data processors for a share in the economic value generated from their personal data, once such data is anonymized and aggregated into a data pool. Such data is treated implicitly and automatically as the private property of data processors.[17] Another shortcoming of the individualist approach is that it is based on an erroneous assumption that privacy risks are limited only to personal data processing. Data-based profiling may occur with the mixing of machine-observed data and personal information, even without lapses in anonymization. Such individualistic approaches, adopting a simplistic, technical stance that privileges privacy and data security, also disregard the differential impacts on different economic actors, doing little to effectively restore market fairness in the data economy or tackle the market power of existing Big Tech corporations.

A shortcoming of the individualist approach is that it is based on an erroneous assumption that privacy risks are limited only to personal data processing.

---

16    Anita Gurumurthy and Nandini Chami, Governing the Abstract Object of Data: Towards a Distributive Integrity Framework, IT for Change, 2021.

17    Ibid.

# 2. DATA STEWARDSHIP AND THE LIMITS OF DECENTRALIZED CONTROL

# DATA STEWARDSHIP AND THE LIMITS OF
## DECENTRALIZED CONTROL

Fledgling, decentralized **data stewardship initiatives** are starting to emerge

Data sharing arrangements are **vulnerable to being co-opted**

Data trusts show promise, but questions remain on their **sustainability and capability** to deliver payoffs for social equity
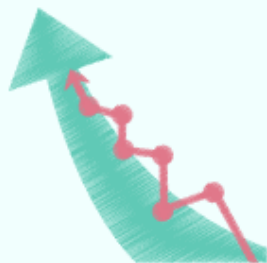
Given current distributions of power and data, **data cooperatives struggle to scale**

Such decentralized initiatives are only viable with the **support of strong public legislation**

- The umbrella term 'data stewardship' refers to fledging initiatives – data sharing pools, data cooperatives, data trusts etc. – for collaborative governance of data resources and cooperativist models of data value generation, working within an imperfect mainstream data economy.

- Regulated by private contracts, data pooling arrangements involving public and private actors often fail to break data silos controlled by Big Tech firms and can even undermine public interest.

- Given the differences in bargaining power among the various actors in the data economy, data cooperatives could typically end up as elitist enclaves, or fail to scale up.

- Public data trusts do have the potential to promote data sovereignty and create public data infrastructure, but their sustainability depends on how payoffs for social equity can be ensured through appropriate institutional frameworks.

- To unlock the generative potential of data and intelligence, and socialize data value, (meso-level) models may not succeed without (macro-level) public law on data governance.

**Discussions in the roundtable analyzed both the promise that data stewardship initiatives hold as well as their limitations.**

Acknowledging the governance deficit in the data economy, the roundtable also mapped emerging models and arrangements to communitize/decentralize data value, affirming the collective and social dimensions of data relations. These initiatives have come to be known in the literature through the umbrella term 'data stewardship', and include "any institutional arrangement where a group of people come together to pool their data and put in place a collective governance process for determining who has access to this data, under what conditions, and to whose benefit".[18] Discussions in the roundtable analyzed both the promise that such initiatives hold as well as their limitations.

## Data Sharing Pools

Data sharing pools (DSP) are horizontal partnerships between different data-holders to facilitate combining datasets for augmenting data-analytics and broadening the scope of intelligence that can be generated. There are a number of such arrangements in place within the digital economy today.[19]

---

18   Jared Robert Keller, Pauline L'Hénaff, and Jeni Tennison, Applying New Models of Data Stewardship to Health and Care Data, Open Data Institute, 2020.

19   "An example of DSP is the Connected Citizens Program, a collaboration between Waze, a community-based traffic and transport app, E.S.R.I., a global commercial software company, and municipal governments. As part of the pool, municipal governments share real-time construction and road closure data through the E.S.R.I. platform, and in exchange Waze shares its community-collected real-time traffic data". Marina Micheli et al., Emerging Models of Data Governance in the Age of Datafication, Sept 2020.

**Marina Micheli** presented some of her research on emerging data governance models at the roundtable, and noted a number of salient, critical points with respect to the DSP model. Firstly, while in principle such data sharing agreements may occur among any entities, given the state of affairs, they currently take place only among large private companies or between such companies and branches of the public sector. As a result, data subjects themselves have almost no control over what data is pooled together, how analysis is undertaken and how it is used. Thus, the idea of having such decentralized, horizontal agreements is appealing when the actors involved are entities such as data cooperatives or trusts, but as things stand, this model seems to simply facilitate the growth of Big Tech's data silos.

The fact that such agreements are regulated only by private contracts is also concerning. As Micheli points out in her paper[20] on emerging governance models, and as was reiterated in the roundtable by **Ingrid Schneider, Parminder Jeet Singh and Freyja Van Den Boom** during the discussions, contracts are fair when they are negotiated between equals. Yet, when one side has leverage – as is the case with big data behemoths – they can dictate terms and institute one-sided agreements that fortify their power within the domain.

Nandini Chami elaborated why pooling data may not bring bargaining power. The data commons and data stewards are, in the end, forced to trade with the monopolist. Avoiding this is not easy either, since it is very difficult to replicate the carrier layer of data, which network monopolists control completely.

Paul-Olivier Dehaye also pointed to how less-than-ideal laws or rules for mandated data sharing can, paradoxically, end up undermining the public interest when the entity that has collected the data is able to dictate the terms of negotiations. To illustrate this, he offered the example of the data-sharing partnership between the not-for-profit project SharedStreets and Uber, which was built to provide aggregate data to cities on where electric scooters were parked and whether they were parked at the right place, etc. He pointed out that, at the level of creating the contractual relationship with municipalities who wanted to benefit from this project, Uber was able to play different cities against one another, and gain the best terms for itself. Thus, it is important to ensure that legislative interventions are adequately robust, and cannot be taken advantage of by Big Tech to further their own market interests.

> While in principle such data sharing agreements may occur among any entities, given the state of affairs, they currently take place only among large private companies or between such companies and branches of the public sector.

---

20    Marina Micheli et al., Emerging Models of Data Governance in the Age of Datafication, Sept 2020.

**Personal Information Management Systems (PIMS)**

A special case of data sharing arrangements, Personal Information Management Systems (PIMS) received attention at the roundtable as a potential tool to wrest back control from the dominant data economy and its extractivism. Presented as an alternative to current 'centralized' data processing approaches, whereby user data is (rather opaquely) collected and processed by organizations, PIMS intend to provide users with technical mechanisms for aggregating and managing their own data, determining when and with whom their data is shared, and the computation that may occur over that data.[21]

The European Data Protection Supervisor (EDPS) has noted that PIMS allow for a human-centric approach to personal data and to new business models, protecting against unlawful tracking and profiling techniques that aim at circumventing key data protection principles.[22] Specifically, PIMS would enable individuals to track back who has had access to their digital behavior, as a basic feature of a common concept of PIMS involves providing access control and an access trail.

Paul-Olivier Dehaye led the roundtable to a discussion on PIMS, drawing from his experience with MyData Global, a non-profit that seeks to empower individuals by improving their right to self-determination regarding their personal data.[23] He noted that within MyData, there were contestations around how to ensure that the entity responsible for managing an individual's consent is distinct from the one that holds and/or processes an individual's data. The popular solution to this, he said, seemed to favor democratic oversight over the technical requirements that are enforced in building the MyData system.

In response to this, Ingrid Schneider highlighted that the MyData model itself relied, at a certain level, on an implicit trustful relationship with the state, which is not true for a large part of the world. Given that states request access to data, this was a point of tension and concern. Additionally, from a social sciences standpoint, the term 'MyData' could be critiqued for not representing 'OurData'. Given that data must be seen in terms of social relationships (an abiding thread that ran through the roundtable), a system like MyData might not prove very effective. Ingrid also added that for some sensitive data, PIMS should definitely be financed by the state, by public entities or by non-profit organizations to prevent commercial interests from taking on such intermediary services.[24]

---

21  Heleen Janssen et al., Personal Information Management Systems: A User-Centric Privacy Utopia?, Dec 2020.

22  European Data Protection Supervisor, TechDispatch #3/2020- Personal Information Management Systems, Jan 2021.

23  About, MyData.

24  The EDPS has also stated that, if adopted, the EU Commission's Data Governance Act would provide a framework for PIMS in the EU. This would include making available the technical or other necessary means to enable such services.

# Data Cooperatives

Data cooperatives (DCs) are one form of 'data stewardship'. Deriving from the traditional cooperative movement, DCs involve a group pooling together their data for various uses under a democratic governance model. "Participants of DCs share data while retaining control over it, having a say on how it is managed and put to value."[25] This makes data subjects a key stakeholder and brings them into the governance framework. By "establishing a relationship of trust with the cooperative that manages data on their behalf, they preserve democratic control over their data and might demand an equitable share in the benefits produced."[26]

While certainly promising, experiments with DCs are still at an incipient stage and it is hard to evaluate their effectiveness. Even at this stage, however, a number of participants raised concerns, including the fact that collective modalities could become "a device for large data monopolies to externalise their regulatory burden, reducing administrative costs and reputational risks in the process of data collection and processing".[27]

Barbara Prainsack cautioned against this model, saying that they run the risk of not genuinely amounting to a socialization of data value, and instead "foster[ing] exclusive or even elitist forms of in-group solidarity."[28] In her 2020 paper, mentioned above, Micheli makes a different point, noting that "this model struggles to compete and scale up against Big Tech that are advantaged by their monopolistic position, their critical mass of users, and greater financial resources."[29]

Discussions at the roundtable also engaged with the question of how cooperative approaches to communitize data value need public law approaches. Singh observed that for DCs to be effective, enabling public legislation needs to exist – for instance, provisions in the EU's draft Digital Markets Act that mandate private companies to share data with such data cooperatives.

**Trebor Scholz** pointed out that when dealing with these models, we should not simply throw around the word 'cooperativism' without precision, and instead engage with the rich tradition behind it that spans almost 200 years. For example,

**Deriving from the traditional cooperative movement, data cooperatives involve a group pooling together their data for various uses under a democratic governance model.**

---

25    Marina Micheli et al., Emerging Models of Data Governance in the Age of Datafication, Sept 2020.

26    Ibid.

27    Anita Gurumurthy and Nandini Chami, Governing the Abstract Object of Data: Towards a Distributive Integrity Framework, IT for Change, 2021.

28    Barbara Prainsack, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

29    Marina Micheli et al., Emerging Models of Data Governance in the Age of Datafication, Sept 2020.

data cooperatives – as currently conceived – are a form of 'consumer cooperative' – a model which has a number of strong critiques within the cooperative movement itself. It is surprising that there isn't more recourse to this literature in thinking through these problems. Moreover, Scholz called for much more participatory research that involves communities in this work, instead of legislating models from the top down. He emphasized the need to work with actual cases and communities and see how these data are used and what kind of data sovereignty they want, what kind of technological sovereignty they want, and then see what makes sense for them.

## Public Data Trusts

Another version of data stewardship is the public data trust (PDT), wherein public, local or civic administrations serve as the intermediary that aggregates and uses its citizens' data. It does so by acting both as the steward that provides third-parties and private actors access to this data under certain conditions, and using such data for the creation/commissioning of projects and services deemed to be in the public interest. In short, "a key goal of PDTs is to integrate data from multiple sources to inform policy-making, promote innovation and address societal challenges, while adopting a responsible approach to the use of personal data."[30]

**Another version of data stewardship is the public data trust, wherein public, local or civic administrations serve as the intermediary that aggregates and uses its citizens' data.**

One of the main examples of the PDT model in action – which was cited favourably by a number of participants in the roundtable – is the pilot project under such a framework being run in Barcelona as part of the EU-funded DECODE initiative.[31] Presenting itself as a 'Civic Data Trust', the pilot in Barcelona attempts to create a publicly-funded data infrastructure that also regulates the conditions in which data from the city and its inhabitants is collected and used by third-parties. Some of its key measures include:

- introduction of 'data sovereignty' clauses in all public service contracts that impose a mandatory obligation on any supplier to the Barcelona municipality to share associated data in machine-readable format and using open APIs and open standards to guarantee interoperability across the data pool;

- smart contracts and cryptographic tools that enable citizens themselves to directly contribute data to the city data commons in privacy-compliant ways and with full autonomy over the terms and conditions of data sharing;

---

30   Ibid.

31   DECODE project.

- opening up the data commons to local companies, cooperatives and social sector organizations that create public value through data-based innovations.[32]

This initiative, as well as the PDT model in general, resonates well with the project of socializing data value and seemed to go in the direction that many of the speakers believed represented a genuine alternative to the status quo.

**Stefaan G. Verhulst** raised concerns about the financial sustainability of data trusts, noting that most models that he had seen simply were not financially viable in the long run and this was a serious problem given the large investments needed in setting up such public infrastructure.[33] Micheli also spoke briefly about the importance of creating sustainable models for local administrations, and in her input to the roundtable, underscored the need for the necessary technical, legal and operational frameworks and infrastructures to ensure social benefits and outcomes. Noting that "once access to data is achieved, the question of how data is used to serve the public interest remains", she reflected how "it remains to be seen to what extent the use of such data assets have an impact, and of what kind, to the people from whom the data comes. This is especially relevant when the analysed personal data belongs to vulnerable or less privileged groups."[34]

Discussions about collectivizing data management and commonsifying data value pointed to how stewardship models for data, in their current design, have limitations. The redistribution of data value at a whole-of-economy scale may not be achieved with practices of pooling and sharing that do not make a dent on distributing the generative potential of data and intelligence, and, hence, call for a paradigmatic change in data governance.

One of the main examples of the PDT model in action is the pilot project under such a framework being run in Barcelona as part of the EU-funded DECODE initiative.

32    Anita Gurumurthy and Nandini Chami, Governing the Abstract Object of Data: Towards a Distributive Integrity Framework, IT for Change, 2021.

33    Stefaan G. Verhulst, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

34    Marina Micheli, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

# 3. NEED FOR CONCEPTUAL REFINEMENT

An important facet of the data governance challenge is **epistemological**

Theoretical directions need to evolve more precisely to address **data's uniqueness as a resource**

# CONCEPTUAL FRAMEWORKS
# FOR DATA GOVERNANCE MUST BE SHARPENED FOR POLICY-BUILDING

**Interplay between** macro- and meso-level governance frameworks needs attention

**Normative foundations** for policy must be based on tenets of community autonomy and protection from harms

- An important facet of the data governance challenge is epistemological. The unique nature of data as a resource has to be comprehended accurately in order to effectively address its governance deficits. We need to separate the physical and syntactic aspects of data, and comprehend digital intelligence as distinct from digital traces.

- Theoretical directions need to evolve more precisely to address data's uniqueness as a resource. The normative assertion of data as commons needs appropriate principles – without slippage into those deriving from common property resource frameworks applicable to natural resources.

- Normative positions that privilege data's social relational ontology could evolve from constitutional tenets of community autonomy. The individual and societal data harms at the heart of the extractivist data economy also present a strong normative case.

- In grappling with data value, legal and economic analyses need to be complemented by a social science perspective that puts power analysis at the centre. Similarly, the interplay between macro-level laws and meso-level frameworks and practices need to be studied more closely.

**A recurring thread in the roundtable discussions related to the need for a refinement in theory and narrative-building.**

One of the chief obstacles in designing effective regulatory frameworks and institutions for the digital economy, and for 'data value' in particular, is epistemological in nature. The breadth of technological transformation across sectors, the rapidity with which novel economic forms are evolving and the unique character of data as a resource have created a situation where we lack the conceptual and rhetorical tools to render these dynamics fully transparent. Moreover, as with any contested terrain where there are strong interests at play, the representation of issues in mainstream debates is often skewed by ideological currents that constrain the framing of the problem.

A recurring thread in the roundtable discussions related to the need for a refinement in theory and narrative-building – both in order to better understand the data economy and develop more effective frameworks for regulating and socializing data value. It was also felt that alternative framings are needed to dispel dominant ideological assumptions and myths perpetuated by mainstream discourse that tend to distort popular perceptions of what the key issues really are when we think about data and activities around it.

The following are some of the high points from the roundtable pertaining to this topic.

# Capturing Data's Uniqueness

At the heart of the theoretical quandaries that plague discussions around data governance are the many unanswered and not fully clarified questions about the nature of data itself. What exactly does it constitute? In what sense is it a resource? What lends such value to it? How can we conceive of general claims about its economic properties, or the forms of knowledge and power that it can bestow?

Many roundtable participants made valuable contributions to tackling these questions in their inputs and discussions. Anita Gurumurthy and Nandini Chami unpacked 'data' into its constituent layers – the physical (network-data architecture), the syntactic (machine-readable data), and the semantic (encoded information). Using these distinctions, they articulated a clear goal that alternative data governance frameworks ought to strive towards – the creation of institutional checks that prevent the possessor of the syntactic or physical layers of data from claiming exclusive rights over its semantic layer.

In his 2019 paper, Parminder Jeet Singh characterizes the core competency of Uber and Amazon in being the 'brain' of the vast transportation and commerce ecology that they manage and control.[35] This 'brain status' requires the analysis of huge volumes of data, but it is achieved once that analysis yields actionable intelligence. Thus, "the real resource at the core of the digital economy, and its new relationships, is digital intelligence. This intelligence is built from data but not reducible to it. Data is something inherent in the concerned social relationships, left as digital traces over platforms from where it is collected and processed by digital companies".[36] The separation of data from data-enabled intelligence may be seen as another dimension about the digital phenomenon that becomes vital in thinking about what exactly needs regulation, and how best to go about it.

Apart from this, a recurring and particularly salient point that came out of the roundtable discussions was the need for a richer conception of the social and relational character of data, and the imperative to take this as a starting point in developing new governance frameworks.

> **How can we conceive of general claims about data's economic properties, or the forms of knowledge and power that it can bestow?**

35   Parminder Jeet Singh, Data and Digital Intelligence Commons (Making a Case for their Community Ownership, IT for Change, 2019.

36   Ibid.

Barbara Prainsack stressed that in order to bolster collective forms of responsibility, oversight and ownership of data requires nothing less than a "a new way of thinking of data subjects and of data that is underpinned by a relational ontology."[37]

**Nadya Purtova** sounded a note of caution on the theoretical work on 'data commons' frameworks, identifying an equivocation regarding how the concept is employed in much of the literature. She noted a conflation between thinking of data as a **common-pool resource (CPR)** in the Ostromian tradition,[38] and simply thinking of data 'being held in common'. In her input paper, Purtova submits that the latter framing "does not imply that data has certain inherent characteristics. Instead, it is more of a normative claim: data should be held in common, as opposed to, for instance, by a few tech giants."[39]

While the value of the CPR framework is its status as a well-developed research program that has formulated a series of design principles for the governance of common-pool resources, these have been tailored for resources with very specific properties. Namely, those that are rivalrous, excludable, and generally associated with the notion of the 'tragedy of the commons'. This paradigm may not be directly applicable in the case of data. Conversely, the normative claim that data ought to be commonly owned is useful as a tool for advocacy and can be employed as a way to shift mainstream narratives and perceptions, but it does not come readily equipped with the resources to build a novel governance framework.

Purtova pointed out that a number of authors and articles often put forward the normative proposition, making a claim towards the need for data to be commonly owned, but then reach for the CPR design principles as a working model for how such a data commons could operate. Such eclecticism, she claimed, was both misleading and counterproductive. We must be rigorous in our scholarship in order to justify either trajectory. The CPR framework does have resonances with our predicament and it can be utilized, but it requires a better articulation of what it is in 'data' that is the common-pool resource we are dealing with and what aspect of it can be broken down into something that is genuinely rivalrous, and would fit the CPR frame. Similarly, building a strong normative basis for data to be held in common is also a viable path worth pursuing, but it needs explicit debate regarding how and on what principles such normative claims may be made most effectively.

> Building a strong normative basis for data to be held in common needs explicit debate regarding how and on what principles such normative claims may be made most effectively.

37    Barbara Prainsack, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

38    See for e.g, 'Governing the Commons', 1990; 'Crossing the Great Divide: Co-production, Synergy and Development', 1996; and 'A Framework for Analyzing the Knowledge Commons', 2003.

39    Nadezhda Purtova, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

# Laying the Normative Foundations

Noting how it may be more intuitive for normative claims to be made on behalf of individuals and towards their personal data, Purtova submitted that this is harder when it comes to other forms of machine-generated or non-personal data. Yet, these may be the key resources whose value we are looking to socialize. A number of participants at the roundtable touched on how this conundrum may be addressed, articulating possible normative foundations to realize data's collective claim.

Arindrajit Basu, for instance, emphasised how "the battle today is about the loss of control over an individual or group's own data — which is a summation of their lived experiences. This loss of control is what data governance mechanisms should look to check by using governance tools that shaped previous struggles of resistance."[40] Given this, 'autonomy' ought to be seen as "a central societal/constitutional tenet to empower and safeguard the interests of individuals and communities".[41] Basu noted that autonomy has been implicated in instituting a rights-based approach for the protection of individual data, and while this is not sufficient precisely because it lacks a collective dimension, it could be expanded to think of the autonomy of communities as well, grounding a normative claim for community rights over data. That said, Basu also noted how finding an adequate definition of 'communities' may be challenging.

Such a perspective finds resonance with some of IT for Change's recent work[42] that attempts to ground the normative basis for collective claims on data to the economic rights enshrined in various national constitutions. As the research points out, in a number of constitutional texts, these rights implicitly frame the nation as a community, and enjoin the state to ensure an equitable allocation and use of the nation's resources in a manner that fortifies the common good. If one can make the case for 'data' as one of these resources, it is possible to find a legal ground on which to build policies for collective rights to data.

Coming at it from a different angle, **Salomé Viljoen** argued that reconceiving data that places emphasis along its 'relational' dimension also carried the seed for a strong normative position, since such a reconceptualized account "offers

> If one can make the case for 'data' as a national resource that fortifies the common good, it is possible to find a legal ground on which to build policies for collective rights to data.

40    Arindrajit Basu, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

41    Ibid.

42    Anushka Mittal, Exploring the Constitutional Tenability of Data-Sharing Policies, IT for Change, 2020.

an alternative (and it argues, more precise) normative argument for what makes datafication – the transformation of information about people into a commodity – wrongful. What makes datafication wrong is not (only) that it erodes the capacity for subject self-formation, but also that it materializes unjust social relations: data relations that enact or amplify social inequality. This egalitarian normative account indexes many of the most pressing forms of social informational harm that animate criticism of data extraction, yet fall outside typical accounts of informational harm."[43]

## Some Methodological Heuristics

Marina Micheli and Kristina Irion provided important inputs on research methodologies to study data governance models. While their analytical frames are developed with respect to the European context, their insights may be seen as relevant across the board.

**Participants also provided important inputs on research methodologies to study data governance models.**

Micheli delved into the theoretical scaffolding of her work, delineating what she called a **'social-science informed perspective on data governance'** that could be used to enrich and complement the legal and economic analyses that currently proliferate. This approach attempts to study how data governance is made through the everyday practices, attitudes, perspectives and imaginaries of social actors and understands data governance as "the power relations between all the actors affected by, or having an effect on, the way data is accessed, controlled, shared, and used; the various socio-technical arrangements set in place to generate value from data, and how such value is distributed between actors." This framework is employed in the analyses of emerging data governance models that Micheli's 2020 paper explores,[44] key threads of which have already been discussed in this report.

Irion presented the key thesis from her 2020 paper[45] that evaluating the efficacy of data governance regimes requires us to pay greater attention to the interplay between macro and meso-level frameworks. She argued that over the last decade, apart from the evolution of various types of national data governance legislation (the macro level), there has also been "considerable variation in technical, legal and normative frameworks that govern the production, extraction and exploitation of data. Different firms, industries, national governments and municipalities, and a diverse group of techno-legal driven communities have come up with their own data

43    Salomé Viljoen, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

44    Marina Micheli et al., Emerging Models of Data Governance in the Age of Datafication, Sept 2020.

45    Kristina Irion et al., Competing Logics of Data Ordering Regimes: Law, Technology, or Governance, Special Issue of Internet Policy Review on Protecting "European Values" Inside Data Flows.

governance practices, frameworks, technologies, such as data sharing agreements, data trusts and cooperatives, or distributed ledgers and personal data stores"[46] (the meso-level). Her central claim was that the efficacy of macro-level legislative regimes depended largely on whether the right kinds of institutions and mechanisms were at work at the meso-level. Yet, we have not studied these phenomena enough to know what the optimal policy response is. Thus, she ended with a call for more sustained research on how to create the most harmonious dynamic between these two levels.

**We need more sustained research on how to create the most harmonious dynamic between macro and meso-level frameworks.**

---

46    Kristina Irion, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

# 4. NEW FRAMEWORKS AND APPROACHES

New policy frameworks to democratize and socialize data value

Diverse stewardship models for collective data governance tailored to specific contexts

# TOWARDS GOVERNING
# DATA AS A RESOURCE

Informed policy support for developing country negotiators in multilateral forums

Global norm-building to stop Big Tech companies from being the default norm-setters for the data economy

- New governance frameworks have been proposed in emerging scholarship to democratize and socialize data value, including:

    ◊ Solidarity-based Data Governance: Bringing benefit-sharing and collective responsibility for harm-mitigation to the fore.

    ◊ Semi-commons data regime: Sketching a distribution of rights over data that challenges the de facto ownership claims of data-collectors.

    ◊ Data re-use maximization: Finding ways to expand the use of data-sets to more public-minded and collective purposes.

    ◊ Democratic Data Governance: Re-thinking the nature and purpose of governance, with data's relationality as a point of departure.

    ◊ Peer-to-peer production: Bringing the communitarian dynamism of peer networks to the data economy.

- No single alternative approach may be adequate in challenging the intellectual monopoly capitalism of the data epoch, owing to the contextual diversity in data collection and processing. Context-specific, collectivist approaches are the need of the hour.

- Global norm building is urgently needed to stop Big Tech companies from continuing to be the default norm-setters for the data economy.

    ◊ While this may not be easy through the global, multilateral route, the UN still remains an important and relevant space that needs to be reclaimed, especially from the standpoint of Global South nations and marginalized populations.

    ◊ A multilateral regulatory body – Digital Stability Board – with the membership of participating nation-states who set clear mandates to be taken up by working groups could be one way forward.

    ◊ Regional agreements may also be useful, as also a legal pluralism based on bottom-up norm building processes.

- There is also an urgency to ensure that critical scholarship and a systemic and forward-looking analysis of the data economy and its governance can reach and inform policymakers in the Global South.

One of the most important threads running through the roundtable's various sessions was an attempt to think through how a project like socializing data value can actually be realized. To this end, participants put forth sketches of novel frameworks and building blocks for reimagining data governance. Participants also deliberated on the institutional arrangements – the appropriate level (local vs. national vs. global) at which transformation was most viable as a fulcrum for change.

In this final section, we present the new frameworks and ideas that were proposed, as well as some of the highlights from the discussions and debates around how best to take this crucial project forward.

## Models for Socializing Data Value

Given the extensive stocktaking and critiques featured in the roundtable on both individual-centric approaches as well as emerging models for novel data intermediary institutions, the obvious question to be grappled with was about addressing the impasse in the economic governance of data: what were the steps needed to overcome the deficiencies in current approaches and proposals, and what would alternatives that recognized the nature of the problem and were radical enough to address them look like?

In a number of sessions, such new frameworks were articulated, and the following represents an overview of some of the key ideas put forward.

**What are the steps needed to overcome the deficiencies in current approaches and proposals, and what would alternatives that recognized the nature of the problem and were radical enough to address them look like?**

## Solidarity Based Data-Governance

Drawing from her work on health data over recent years with Alena Buyx and her wider research group, **Barbara Prainsack** presented the outline of her 'Solidarity Based Data-Governance' program that attempts to provide an alternative paradigm rooted in the relational character of both data and human relationships, and committed to both re-purpurposing the data economy to working for the collective good, and upholding normative values of justice and solidarity in the way it operates.

The program itself consists of three main pillars:

1. Facilitating the use of data in the public interest.

2. Introduction of Harm Mitigation Bodies (HMBs) that would provide unbureaucratic, low-threshold support for people who have plausibly been harmed by data use but cannot access legal remedies.

3. Strengthening the mechanisms of benefit-sharing to ensure that some of the profits that emerge from commercial data use come back to the public domain that has enabled data use via public infrastructures and the data work of patients and other citizens.[47]

---

47    Barbara Prainsack, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

## A Semi-Commons Governance Regime for Data

**Anita Gurumurthy and Nandini Chami** presented what they called a 'Semi-Commons' framework. This approach explicitly "demarcates the boundaries between common property and private property in data resources ownership",[48] and does so "through a differentiated, rights-based resource ownership regime."

Essentially, this means that data-holders within a semi-commons regime are only allowed to "have the right to non-exclusive access over the base layer of data they have collected, without exclusive possession rights." As a corollary to this, such a regime grants everyone "the right to seek data in the datasets collected, aggregated and controlled by for-profit legal entities, altruistic organisations and public agencies through an entitlement of accessibility."[49]

Of course, the right to seek data is envisioned to be graded, and have accompanying conditionalities that go along with the type of entity asking for data (for-profit, public agency, data altruistic organization, individual data subject), the type of data (personal/non-personal, raw/aggregated), and the intended purpose for which it is being sought.

The authors cautioned that in order for a genuine transformation, such a semi-commons distributed ownership regime would have to be supplemented with a series of organizing principles to regulate the data economy as such. This would have to include a traceability obligation on data businesses and altruistic data organizations to disclose their sources of data collection to an appropriate authority, investments in public data infrastructure to unlock the public and social value of data; protection against re-identification and data sharing and safeguards against profiling risks, and a new structural separation in data value chains between the data layer the cloud computing layer, intelligence layer and consumer-facing intelligence services layer.

---

48    Anita Gurumurthy and Nandini Chami, Written input to Roundtable on Socializing Data Value, IT for Change, May 2021.

49    Anita Gurumurthy and Nandini Chami, Governing the Abstract Object of Data: Towards a Distributive Integrity Framework, IT for Change, 2021.

## Towards an Effective Framework for the Re-use of Data

**Stefaan G. Verhulst** presented some of the work he has been conducting with the GovLab, an action-oriented think tank that he co-founded, located within the Engineering School of New York University (NYU). One of the central tenets of this work is the 're-use of data', or the opening up of previously siloed data-sets to be employed by actors other than the original data collector, and in the service of problems other than those they were originally harvested to solve. According to him, "Better and more transparent re-use of data is arguably the single most important measure we can take to unleash the full possibilities of data."[50]

In order to forward this agenda, and make such practices of re-use central to the data economy, Verhulst highlighted four important steps he believed were necessary to pursue:

1. Develop new participatory methodologies to identify and measure the value of data;

2. Develop enabling ecosystems and collaborative frameworks to move from extraction to co-creation of value;

3. Innovate with new data collaborations and re-use conditions;

4. Identify and nurture data stewards.

Ultimately, according to Verhulst, "Re-using data is a vital step toward generating social value in data. Yet we are only beginning to understand the tradeoffs involved in re-using as well as the institutional frameworks and structures that can encourage it. The four points outlined above represent a start, but we need a rigorous assessment into what's already being done, and more experimentation to push the frontiers of what's possible."[51]

---

50   Stefaan G. Verhulst, Written input to Roundtable on Socialising Data Value, IT for Change, May 2021.

51   Ibid.

## Democratic Data Governance

Building on discussions on the vital need to foreground the relational character of data, **Salomé Viljoen** presented her work on the project of what she calls 'Democratic Data Governance', outlined in her 2020 article, "Data as Property?".[52]

She prefaced her own framework by noting how current paradigms of data governance are rooted in particular notions of data captured in two approaches: 1. The 'propertarian', which conceives of data as a kind of asset that is subject to individual ownership and which grants legal rights that will allow it to be monetized; and 2. the 'dignitarian', which "conceives of data as an expression (or extension) of individual selfhood", and which "take a further step beyond asserting rights to data-as-property, and resist data's commodification altogether, drawing on a framework of civil and human rights to advocate for increased protections."

According to Viljoen, both of these approaches isolate important dimensions of the problem, but they fall short in being unable to overcome an individual-centric account. Hence, she offers the scheme of "Democratic Data Governance", which "views data not as an expression of an inner self subject to private ordering and the individual will, but as a collective resource subject to democratic ordering." In this perspective, "information about one individual is useful (or harmful) precisely because it can be used to infer features about – and thus make decisions affecting – others. Data production places individuals in population-based relations with one another; the social effects that result cannot be adequately reduced to individualistic concerns nor can they be addressed via individual-centric institutions". Thus, what is called for is that "rather than proposing individual rights of payment or exit, data governance should be envisioned as a project of collective democratic obligation that seeks to secure those of representation instead."[53]

---

52   Salome Viljoen, Data as Property?, Phenomenal World, Oct 2020.

53   Ibid.

## Peer-to-Peer Production

Drawing on the work of Kostakis and Bauwens,[54] **Raymond Onuoha** suggested a return to and development of the radical potential of 'peer-to-peer production' as a model for data governance. One of the hallmark experiments spawned by the internet, peer-to-peer production is a form of creative activity that relies on self-organizing communities of individuals who can break down tasks in a modular fashion and engage in collective, network-centric production. Free and Open Source Software and Systems, such as Linux, are a result of peer production, and amenable to being situated in non-profit organizations, with Wikipedia and Mozilla as prime examples. Moreover, peer production can also be used as a methodology for the development of new ventures, such as the Wikispeed project,[55] which aims to design a new, affordable and environmentally-friendly car using peer-based collaboration.

Onuoha suggested that the model of peer-to-peer production could be applied to data governance. If instituted, he notes, it would require a commons-based reciprocity license that would permit any user to benefit commercially from the data commons insofar as they contributed to the co-created (consumer+producer) data value chain. However, taking from Elionor Ostrom's work, he warned that such a system cannot be self-regulated privately due to perceived free-rider arguments, but must involve access and withdrawal rights to the data commons, especially for commercial appropriation. Onuoha highlighted that operationally, such a data governance system, at a large scale, is still largely underdeveloped and lacks a conceptual framework.[56]

---

54    Vasilis Kostakis and Michel Bauwens, Network Society and Future Scenarios for a Collaborative Economy (2014, Palgrave Macmillan).

55    Car, Wikispeed.

56    Raymond Onuoha, Written input to Roundtable on Socialising Data Value, IT for Change, May 2021.

# Nuancing Policy Interventions

Given the multitude of ideas in circulation, and the many potential directions to pursue, participants engaged with the question of what the most appropriate way to streamline efforts to produce more effective policy results was, along with the kind of advocacy initiatives that needed urgent attention, and the pitfalls to be avoided. While these questions are too broad to have been settled, there were important points raised for informing and shaping policy interventions, going forward.

The central thesis of Bruno Carballa Smichowski's presentation, for instance, was that "no one-size-fits-all alternative data governance model can respond at once to the many issues the data economy poses".[57] He pointed out how the different types of data being collected and the contexts in which they were being employed were so diverse, it was very difficult to fashion a single model that was optimally applicable to all. Given its sensitivity and strong links to individual personhood, health data, for example, is most suited to a model that involves the provision of legal property rights to individuals. On the other hand, forms of aggregated or machine-generated data where individual rights are harder to claim could be tailored to a more collaborative approach such as a version of data stewardship. In his presentation, Paul-Olivier Dehaye echoed this view, arguing that "there is no single answer for the governance question in collective-based models. He recommended that approaches be tested, and that such a trial-and-error approach was the only way to match governance mechanisms and collective forms to the structure of data sets and data-use purposes most naturally suited to them.

Notably, a similar call to promote an enabling environment for the proliferation of local initiatives and a trial-and-error approach for determining the best, context-specific governance models was endorsed also by Trebor Scholz and Stefaan G. Verhulst.

On a different note, Cecilia Rikap emphasized that as much as we push for novel models to be adopted and work on tailoring them to perfection, it is imperative that we confront and dismantle the data-enclosures and forms of value capture that are becoming ever more entrenched in the current data economy.

Analyzing the contours of what she called "intellectual monopoly capitalism", she detailed how the core of Big Tech's business models revolved not only around the

**What is the most appropriate way to streamline efforts to produce more effective policy results?**

57    Bruno Carballa Smichowski, Written input to Roundtable on Socialising Data Value, IT for Change, May 2021.

accumulation of intangible assets (data, intellectual property), but also the gradual usurpation of networks of research and innovation, and their key infrastructures. Through one-sided partnerships with public universities, the monetization of open-source projects and strategic mergers and acquisitions, America and China's tech behemoths are currently monopolizing the frontiers of innovation within the data economy.

She pointed out that for alternative experiments in the use of data to be economically sustainable, one had to break down these monopolies – and that tactically, this had to be at the front and center of any advocacy agenda that was aiming at socializing data value.

## Global Norm-Building for Data Governance

**What kinds of multilateral arrangements could support productive and inclusive norm-building for data at a global level?**

**Aaron Martin and Siddharth De Souza**, representing Tilburg University's Global Data Justice Project, described the work the project was currently engaged in, particularly on researching "the lived experience of data technologies in high- and low-income countries world-wide, seeking to understand people's basic needs with regard to these technologies" and relating these findings "to current governance and rights frameworks in order to understand whether they match with people's subjective needs".

Both participants also delved into the project's more recent work on peremptory norms as a potential principle that data governance advocates could borrow from international law. The idea being to institute certain global 'red lines' that prohibit behavior, no matter the circumstances. Martin and De Souza went on to suggest that thinking about what may intuitively be considered off-limits, or alternatively seen as indispensable, and therefore the duty of particular actors in the digital sphere, could act as possible starting points to establish international norms.

This led to an important discussion dealing with the current geo-political situation on data governance, and what kinds of multilateral arrangements could support productive and inclusive norm-building for data at a global level.

Arindrajit Basu claimed that it would be very difficult to build alternate multilateral institutions for data governance or even to subvert current ones, given their stickiness for over 70 years. However, he felt it was possible to challenge the dominant narratives being propagated in these institutions and the process for framing norms within them, and this was a route that was worth pursuing.

Alternatively, and especially for countries in the Global South, he believed it might be strategically beneficial to create robust agreements at the level of regional blocs, if a genuinely inclusive global norm-making process was not forthcoming.

Martin highlighted the urgency of this discussion, noting that it often seems like we are force-fitting a particular sort of discussion on institutions that come with a certain baggage, legacy and pragmatism. And yet, while we spend more time debating this issue, Big Tech is already developing its own norms, and is well ahead in terms of shaping the debate. Given the perceived inadequacy of most existing multilateral forums for this process then, Martin asked whether it was not worth considering whether "we need a new international body to govern data globally? And if so, what would it look like? What would it draw inspiration from? How can we improve on previous attempts at global/transnational regulation?"

A concrete proposal for such an institution has been put forward in recent years by another of the roundtable's participants, **Robert Fay**. Making an analogy with the multilateral regulatory body set up to regulate global finance in the wake of the recession in the late 2000's, Fay proposes the establishment of a Digital Stability Board (DSB). Overseen by a plenary body that would consist of officials from all the member countries, the DSB would "work with standard-setting bodies, governments and policy makers, regulators, civil society and the platforms themselves via a set of working groups with clear mandates that would report back to the plenary."[58] Broadly, the remit of such an institution would cover the following objectives:

1.  Coordinate the development of standards, regulations and policies across the many realms that platforms touch (including areas such as privacy, ethics, data quality and portability, algorithmic accountability, social media content, competition policy, and electoral integrity).

2.  Monitor development, advise on best practices and consider regulatory and policy actions needed to address vulnerabilities in a timely manner.

3.  Assess vulnerabilities arising from these technologies, including their impact on civil society and the regulatory and policy actions needed to address them on a timely basis.

4.  Ensure that this work feeds into other multilateral organizations, especially the WTO.[59]

The non-aligned tech movement is moving beyond regional or international forums, and is building solidarity among organizations at a peer level.

58   Robert Fay, Digital Platforms Require Global Governance Framework, CIGI.

59   Ibid.

**It is crucial that we think about ways in which generated critical scholarship actually finds a way to inform the decisions of the Global South's nations' leaders.**

De Souza put forward the possibility of "a global legal pluralism" – different institutions from where certain norms can emerge. He also proposed the idea of whether such norm-building can be a bottom-up process – one where concepts and issues are taken from actual social movements, civil society actors and others on the ground, and where some consensus is built before it is pushed to a particular governing body. In this regard, he gave the example of the non-aligned tech movement, which was moving beyond regional or international forums, and was building solidarity among organizations at a peer level. He called for a greater exploration of these possibilities, and to examine if we can begin from a plurality of institutions and see if something can percolate upwards.

Chee Yoke Ling made a case for the United Nations as an appropriate global institution for the global governance of data – arguing how it is the one forum conceived to serve this kind of purpose, and where the multitude of voices across nations and civil society can be channelled. As she put it – "the UN, as the norm setting forum, allows for a seat at the table, not only for nation-states, but also civil society, citizens and marginalized populations. It has the potential to create a space of genuine accountability, and we should fight to secure that space." She also acknowledged that the UN was indeed a contested space, where Big Tech was also hoping to push forward and legitimize their agenda; so there certainly lay many challenges ahead. Perhaps one of the most significant, she noted, was the extent to which leaders and politicians from the Global South were duped into supporting policies that were clearly against their national interest. This has a lot to do with the complexity and obfuscation around issues in the data economy, and so, it is crucial that we think about ways in which the critical scholarship that is generated on these subjects actually finds a way to inform the decisions of these nations' leaders.

# Concluding Thoughts

The multiple shades of dystopic futures increasingly evident in our datafied present make thinking about data governance a key imperative for scholarship and practice. The roundtable was an effort in this direction.

The hallmark of the digital paradigm is 'governance by data'. Informational power and knowledge hierarchies have always cemented structures of exploitation. What is new about data-based control is the invincible technical prowess of platform behemoths to engineer economies and their constituent parts through a post-market takeover; the raw power to play God, determining desires and gratifications, human capabilities and destinies, without regard for the rules of the marketplace.

Mapping this institutional terrain, contemplating the conceptual gaps and engaging with frontier ideas, the roundtable brought together cutting-edge analysis and theoretical directions for imagining how data value can be socialized through appropriate governance frameworks.

The discussions traversed a wide spectrum – from problem statements of an unequal global data order to the systemic frames needed for a normatively-grounded, institutional change. The event also saw an emphasis on an integrated approach to ethico-political and economic considerations in contemplating the data economy, underscoring the significance of economic democracy as key to individual and collective autonomy.

The debates reflected the abiding need to grapple with the geopolitics of the data order and narratives of global data justice, using interdisciplinarity to inform policy and practice. Indeed, the nature of capitalism may remain the same in its essence. But the task of bridging social paradigms through involved debate can help us figure out exactly how law and policy can intervene towards equality and justice. The roundtable and this report, we believe, are a contribution to joining the dots, re-politicizing data value chains.

As the policy challenges of data governance intensify, the task of building a commons-based framework for data gathers urgency. A vibrant platform ecosystem with a rightful space for all market actors calls for policy measures to ensure distributive integrity in the data economy. The current discourse on data governance obscures the due attention needed to unlock the non-personal data sets held by private actors in which – as noted in the roundtable – individual claims cannot exist.

**The debates reflected the abiding need to grapple with the geopolitics of the data order and narratives of global data justice, using interdisciplinarity to inform policy and practice.**

How the idea of the 'community' or the 'collective' can be centered in the data rights debate – to check harms, and equally, to shape social and public outcomes from data value – and what, therefore, needs to be done towards breaking the data enclosures of tech giants, would be key. This devil in the policy-detail to galvanize a thriving data commons emerges as a vital epistemological nucleus for scholarship and practice.