A Primer on Data and Economic Justice

November 2022

GPAI | THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE

This report was developed by Experts and Specialists involved in the Global Partnership on Artificial Intelligence's project on Data Justice. The report reflects the personal opinions of the GPAI Experts and Specialists involved and does not necessarily reflect the views of the Experts' organisations, GPAI, or GPAI Members. GPAI is a separate entity from the OECD and accordingly, the opinions expressed and arguments employed therein do not reflect the views of the OECD or its Members.

Acknowledgements

This report was developed in the context of the Data Justice project, with the steering of the project Co-Leads and the guidance of the Project Advisory Group, supported by the GPAI Data Governance Working Group. The GPAI Data Governance Working Group agreed to declassify this report and make it publicly available.

Co-Leads:
Alison Gilwald*, Research ICT Africa/University of Cape Town
Dewey Murdick*, Georgetown University

The report was written by: **Parminder Jeet Singh**‡, IT for Change; **Anita Gurumurthy**‡, IT for Change; with the involvement of the following GPAI Expert: **Alison Gillwald***, Research ICT Africa; **Dewey Murdick***, Georgetown University.

GPAI would also like to thank the following individuals for their support in preparing the complementary Primer on Data and Social Justice: **Kelle Howson**‡, Research ICT Africa, **Alison Gillwald***, Research ICT Africa. GPAI would like to acknowledge the research assistance provided: **Carina Singh**‡, IT for Change. GPAI is also grateful to the following individuals for reviewing the work: **Paul Dalby***, Australian Institute for Machine Learning; **Kim McGrail***, University of British Columbia; **Jaco du Toit****, UNESCO. Finally, GPAI would like to acknowledge the efforts of colleagues at the International Centre of Expertise in Montréal on Artificial Intelligence (CEIMIA). We are particularly grateful for the support of **Arnaud Quenneville-Langis** and **Edward Teather** from CEIMIA, and for the dedication of the Working Group Co-Chairs **Jeni Tennison*** and **Maja Bogataj Jančič***.

* Expert of GPAI's Data Governance Working Group
** Observer at GPAI's Data Governance Working Group
† Invited Specialist
‡ Contracted parties by the CEIMIA

Citation

GPAI 2022. Data Justice: A Primer on Data and Economic Justice, Report, November 2022, Global Partnership on AI.

Contents

This primer on data and economic justice is a complementary primer to the first data justice primer in which it is located in the wider social justice context. A review of the literature on data justice together with three guides for policy makers, developers and users/communities developed by The Alan Turing Institute (ATI) for Global Partnership on Artificial Intelligence Data Governance Working Group provides further background and depth to the emerging concept of data justice and its application to practice.[1]

Abstract

Economic justice is about ensuring the material conditions that enable or are the basis of other kinds of justices. In a digital society, data and digital intelligence are the key resources, and therefore economic justice has to focus on how these resources and their control are distributed. Based on some existing work, this primer proposes four rights regarding data that can form the bedrock of data related economic justice; (1) right to benefit from one's data, and avoid harm; (2) right to access and port one's data; (3) right to appropriate representation in data; and (4) right to govern data and data based systems. These rights would be both individual and collective, depending on the nature of data. Practical steps for ensuring data related economic justice will center on digital platforms, which are the sites of most digital economic injustice – be it in relation to users/consumers; or economic actors like traders, farmers, SMEs, small service providers and workers; or across countries. The counter institution that can address these injustices and distribute data power more evenly is data infrastructures. Indeed, a number of legal/policy and programmatic initiatives already exist, mostly in the EU and India, that seek to employ the institution of data infrastructures to redistribute data power, and thus ensure economic justice. What is needed is to bring theoretical work on data justice closer to the ongoing economic legal/policy efforts, which in default have largely been patchy and inconsistent, and therefore unlikely to be effective. In the last section, the primer lays out a few areas for future theoretical research and evolution of data governance for economic justice.

[1] GPAI (2022). Advancing Data Justice Research and Practice: An Integrated Literature Review, Interim Report, March 2022, Global Partnership on AI. https://gpai.ai/projects/data-governance/advancing-data-justice-research-and-practice-literature-review.pdf

[2] GPAI. (2022). Advancing Data Justice Research and Practice: An Interim Guide for Policymakers, Interim Report, March 2022, Global Partnership on AI. https://gpai.ai/projects/data-governance/data-justice-in-practice-guide-for-policymakers.pdf

[3] GPAI. (2022). Advancing Data Justice Research and Practice: An Interim Guide for Developers, Interim Report, March 2022, Global Partnership on AI. https://gpai.ai/projects/data-governance/data-justice-in-practice-guide-for-developers.pdf

[4] PAI. (2022). Advancing Data Justice Research and Practice: An Interim Guide for Impacted Communities, Interim Report, March 2022, Global Partnership on AI. https://gpai.ai/projects/data-governance/data-justice-in-practice-guide-for-impacted-communities.pdf.

1. Economic justice in the digital age

Economic justice has been defined as "a set of moral principles for building economic institutions, the ultimate goal of which is to create an opportunity for each person to create a sufficient material foundation upon which to have a dignified, productive, and creative life beyond economics". Economic policies tend to initially focus on economic growth. The concept of 'economic justice' adds to these policies the distributive aspect, which is about how 'the pie' – whether growing or static – is to be divided fairly and equitably among all. Digital economic policies still mostly focus on economic growth, although lately distributional issues have begun to surface – such as, entitlements of gig workers, unfair practices of e-commerce platforms vis-a-vis traders dependent on them, and digital colonization across the globe.

Institutions for economic justice have existed across time. Tax-based redistribution and welfare have been common to all eras. Then there is development of public services and infrastructures; employment protections and trade unions; economic regulation like competition law and consumer protection; and interventions in support of economic actors like farmers, small and medium-sized enterprises (SME), artisans, etc. Essential to all these is the imperative to maintain a strong national economy, requiring macro-economic measures in terms of fiscal, monetary, industrial, and trade policies.

Any era's conceptions of economic justice considerably focus on the key resources and factors of production of the times, and how they are controlled and distributed. In the feudal-agrarian age the central focus was land, in the industrial society the main attention shifted to ownership of industrial means and capacity, which later grew to include financial capital and intellectual property. In the digital age, data is recognised as the key resource, and AI or 'digital intelligence'[2] the key factor of production[3]. Formulations about economic justice in a digital age therefore center on (though are not exclusively about) data and AI, in their diverse economic and social roles.

There do not, however, yet exist comprehensive concepts and frameworks for what moral and legal principles should underpin economic justice in a digital society – including in relation to its key elements of data and AI. Neither is there any degree of holistic understanding about what institutions and practical mechanisms are needed to ensure economic justice in the digital society, and the role in that of data and AI governance.

Some early conceptual work is emerging on social and economic justice in relation to data, and to a lesser extent, AI. But largely it remains poorly connected, if at all, to the more practical matters of how economic justice will actually be ensured, and who will ensure it. This brings us to the institutional side of economic justice. Since the digital phenomenon moves fast and causes deep impacts, governments have come up with various kinds of ad hoc policy and legal responses. A lot of these relate to governing platforms, including with respect to data with them. Such efforts have not had much theoretical support, and are based on spontaneous ideas of fairness in a given economic situation (these could be in the aid of traders, start-ups, workers, traditional media, etc., all of which struggle against the power of global digital platforms). This primer surveys some such efforts, seeking to connect them to the emerging conceptual thinking and

[2] UNCTAD employs the term 'digital intelligence' as more socially situated than the technical 'AI', for instance in its biannual Digital Economy Reports.
[3] Purdy, M., & Daugherty, P. How AI boosts industry profits and innovation. Accenture Research. https://www.accenture.com/fr-fr/_acnmedia/36dc7f76eab444cab6a7f44017cc3997.pdf

principles in the formative areas of economic justice and data justice in a digital society. This exercise will also be firmly placed in traditional concepts and institutions of economic justice from the pre-digital era, underlining the need for normative and structural continuity.

Digital society can be defined as one where sector-wide intelligent systems organize and run (and practically govern) various activities and actors in a considerably autonomous manner.[4] We are currently in the early phases of such a systemically intelligent, digital socio-economic architecture. But the general directions towards it are sufficiently clear, and the society and its policy-makers need to be planning for it.

Digital platforms are the sites of vast data mining, collection, and storage, and its conversion into digital intelligence. This intelligence is then employed to re-organise and orchestrate sector wide activities and actors (think Uber and Amazon) to realize huge efficiencies, apart from developing new products and services. As digital partnerships are consolidated upstream along the digital value chain, as with cloud computing providers, and downstream, with traditional players in different sectors (Apple/Google with hospital chains in health, Microsoft with Bayer in agriculture), the two dimensional notion of platforms is shifting towards more complex and multi-dimensional entities called digital ecosystems. Controlled by a very few global corporate behemoths, these digital ecosystems will fundamentally reconfigure the respective sectors, and hold almost all the economic power in those, other than that which is reined in by effective countervailing measures.

In early stages of their development, platforms rely most on network power – the power to decide who connects with whom, and on what terms. But soon, as digital footprints of actors and activities on the platform grow, data/AI power becomes even more key for platforms. Real time granular intelligence about everyone and everything, and their interrelations, enables platforms and digital ecosystems to reorganize and run things in every sector in a manner that creates huge new efficiencies. This large digital intelligence surplus is entirely in the hands of the platform corporations to employ and distribute. A fairer and more just distribution of this intelligence dividend is at the core of economic justice in the digital age. At its foundation, this requires more democratized access to and control over the resources of data and AI. Data in itself is not so useful; what is of actual benefit is the digital intelligence that data contributes. But whoever controls data would have considerable influence over the resultant AI and its uses. It is therefore the ‘data as the key resource’ element of the emerging digital society that currently attracts the greatest attention in terms of economic fairness, justice and equity. This primer focuses on data related economic justice.

It may be worthwhile to lay out the meaning and interrelation of connected but distinct concepts of economic fairness, justice, rights and institutions. Fairness is what is judged as reasonably appropriate in a given situation vis-a-vis the actors involved. It is our general sense of what is fair that leads to the stronger conceptions of what is one’s rightful due, in a more holistic and abiding sense, which constitutes economic justice. This would normally go beyond material entitlements to also include some elements of voice and participation. Strong generalized conceptions of economic justice can get accepted as prior foundational principles on which economic systems should be built or transformed, and may be recognised as such in legal frameworks and constitutions. These become or provide economic rights. Rights however do not by themselves deliver economic justice. They require appropriate institutions to do so, that can actually

[4] Singh, P.J. (2020). Breaking up Big Tech: Separation of its Data, Cloud and Intelligence Layers. Data Governance Network. https://itforchange.net/sites/default/files/add/Regulating_data__cloud_and_intelligence_-_Paper_9-21.pdf

deliver tangible outcomes. These institutions may be in the form of suitable laws, regulation, policies, and implementation mechanisms. It is important to place any thinking and concepts of economic justice in such a larger milieu. This primer attempts this by focusing on rights based conceptions of data related economic justice, that can be expressed in appropriate institutions resulting in just and equitable economic outcomes.

2. Data and economic justice: A theoretical framework

Based on surveying other efforts in the area, Linnet Taylor has provided a conception of data justice that consists of three pillars: visibility (including the right to remain invisible), engagement with technology (which includes the right to share in benefits from data, and autonomy in technology choices), and non-discrimination.[5] A later work by Solano, Martin, de Souza and Taylor, for the Secretariat of the European Parliament, lists four benchmarks for just data governance; preserving and strengthening public infrastructure and public goods; inclusiveness; contestability and accountability; and global responsibility.[6] The public infrastructure and public good part can be considered as relating to access to data and its benefits; inclusiveness to appropriate visibility in data; contestability and accountability to participation in data governance (other parts of the paper by Solano et al speak about collective will in decision-making, and distributed and devolved oversight); and, global responsibility to the need for an outward orientation of any community's data governance apart from the expected focus on its own interests.

We would try to adopt these ideas of data justice to the narrower field of economic justice, also relating them to some emerging legal, policy, and programmatic developments. Most conceptions of data justice are for individual data subjects, but an increasing recognition is found also of the collective data subject, a group or a community. The collective data subject is especially salient to the field of economic justice because the resource nature of data is much more aggregate/collective than individual.[7] The framework we propose for data-related economic justice is a rights-based one, centering on the rights of data subjects, both individual and collective. It consists of four kinds of rights:

1. The right to benefit from one's data, and avoid economic harm;

2. The right to access one's data, including for third parties of choice;

3. The right to appropriate representation in data, including to invisibility and remaining absent, and;

4. The right to participate in governance of data, and of the relevant economic systems based on data.[8]

The right to benefit from one's data, and avoid economic harm (which could be relative), may appear obvious, but significant tension can exist between such distributive norms and utilitarian logic, as relating to effective working of a digital economy. Data is (mostly) neither labor nor it involves a creative effort by

---

[5] Linnet T. (2017). What is data justice? The case for connecting digital rights and freedoms globally. Big Data & Society, 4(2). https://journals.sagepub.com/doi/full/10.1177/2053951717736335

[6] Solano, J.Z., Martin, A., de Souza, S., & Taylor, L. (2022). Governing data and artificial intelligence for all. Panel for the Future of Science and Technology. https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729533/EPRS_STU(2022)729533_EN.pdf

[7] Most AI-enabled systems become more and more useful based on recognising patterns in aggregate data, which is why aggregate data is valuable. This utility does require corresponding individual data when applied to a particular individual context. But the more effective the system is based on access to aggregate data the less individual data it may need for its application. This is why aggregate data is so valuable, or at least its immense value should not be lost sight of in what is currently an almost exclusive focus on individual data, even in economic regulation

[8] We did not use the 'pillar' of non-discrimination from Taylor's framework, because (1) it is a sub-element of right to benefit appropriately, and avoid harm, and (2) AI does nothing if not discriminate – the issue is whether discrimination is good or bad or rather fair or unfair. This gets configured or decided independently based on other factors which can be subsumed under the right to 'rightful benefit and avoidance of harm'.

the data subject, which are the general bases for economic reward. Data's main value is in the intelligence that it contributes about its data subject, individual or collective. It is not difficult to see the moral-philosophical point, and establish it as a principle, that the data subject concerned should rightfully be in control of intelligence about themselves, and therefore of data about themselves.

Comparable precedents do exist of rights and entitlement to resources that may neither be the product of one's labor nor creativity, especially collective rights and entitlements. Some examples of this are, collective rights to royalty for mineral resources in a territory, and to benefits arising from resources connected to local flora and fauna (including from data related to them).[9] Collective rights exist for natural resources because they are considered as 'belonging to' the community most closely associated with them. Data arising from a group or community can similarly be considered as its 'social resource' (apart from an individual one, at a different level). In fact, unlike natural resources, this social resource of data would not exist but for the social unit concerned. Social data – the single most valuable data – is commonly produced and therefore should be commonly owned, with a collective (and individual) right to share benefits from it. This area does require much more theoretical work, especially when involving rival uses by, and entitlements of, different data actors.

The same arguments make the case for an individual and collective right to access one's data, or 'data about oneself'. This is to ensure autonomy about what uses an individual or a community may want to put its data to. However, in general, most data is not a kind of resource that may by itself be very useful to individuals or even small community actors.[10] In its digital economy context, data is generally a system-level or systemic resource that works at scale, in combination with other resources and factors of production. A right to access one's data is mostly fruitful only in terms of being able to port data to another relatively large system-level actor of choice. It is for this reason that most legal initiatives in this area call for data holders to facilitate transfer of data to third parties of choice of rights holder. This is called the right to data portability. Also, the leverage in porting data to such large-scale systems, with data's main value being in its aggregation, is much more collective than individual.

The right to appropriate representation in data has both a negative and a positive element. A right to not provide data or to withdraw one's data may be relatively obvious. But a positive right to be represented in data, and appropriately so, is more complex. Such a right can be imagined for public systems, as a part of the right to fair and equal treatment by the state. This right is relevant to both individuals and communities. The scope of a positive right to representation in data may also extend to regulated privately-provided utility services, such as water supply or electricity, possibly as a part of their licensing terms. This is due to inherent considerations in such services of welfare and equity. For similar reasons, highly regulated sectors like health, education, and agriculture could also move in this direction, as a part of the state's social obligations. For regular commercial services, this positive right may involve lighter burdens of fairness as per consumer rights, and appropriate transparency and disclosure norms. If a product/service purports to target certain sections then sufficient representativity of relevant data and data subjects may need to be ensured, or at least due disclosures made.

---

[9] The United Nations. (2010). Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity https://treaties.un.org/pages/ViewDetails.aspx?src=IND&mtdsg_no=XXVII-8-b&chapter=27&clang=_en

[10] Data is a very broad term used in many ways and senses, including its historical meaning in statistics. When we make general points about 'data' in this paper we mean the kind and quantities of data that are most significant for the working of the digital economy.

As discussed, a digital society is characterized by large-scale data-based intelligent systems that begin to autonomously run and 'govern' different sectors (in the sense of acting considerably without human intermediation). These digital ecosystems are increasingly monopolistic, and thus not optional. In any case, once a sufficient data lock-in is established, the cost of switching can be prohibitive. As Lawrence Lessig put it, in a digital environment, '(software) code is law'.[11] It means that what gets coded in dominant technology systems acts like it has the force of law, as technology mediates much of our socio-economic life. With the maturing of a digital society, regular law may get relegated to a relatively lower status of considerable helplessness in front of the code and architecture of global monopoly digital ecosystems. If indeed code is law, it stands to reason that code must be representative and fair like law. Such is the quasi-legal force of society's dominant digital systems that the constitutional principle of 'equality before law' may effectively translate into 'equality before society's digital systems'. Any such equality first of all requires appropriate and equitable representation in data.

The same reason, of being 'governed' by sector-wide and society-wide autonomous intelligent systems, also provides the basis of seeking a right for data subjects to participate in data governance. At the first level, data subjects should be able to participate in governing how data is collected, processed and used. But this may not be enough. Data from various groups and communities is what fuels sector-spanning digital ecosystems, enabling them to become not just the main production system of any sector but also take on significant quasi-regulatory roles. This is why the data-contributing communities should have a right to participate in the governance of relevant data-based digital systems. After all, any digital ecosystem's main asset and power is the intelligence derived from the data contributed by the community concerned. As data defines and underpins new socio-economic systems, participatory governance by data contributing communities of data and data-based systems adds a very important new dimension to participatory democratic governance of our societies.

The presented rights-based framework is anchored on the data subject, that whom the data is about. Data subjects – individual and collective – are certainly the most important actors here (although this is not any kind of a natural law but a moral choice of the society). But there also exist other actors who could, and should, have some data related rights. A few recent legislative initiatives focus on those who generate data, or the data generator[12] – the principal party responsible for coming into existence of data (while admittedly more than one party is generally involved). It could be data from equipment, other artifacts, or resources, that are owned or operated by a person or entity. A farmer, for instance, may be given rights to data from their farm, an SME owner to data emanating from their machines, and a trader to data about goods they sell on an e-commerce platform. Another important actor is the 'data worker', who through their employment or work helps generate data that could pertain to others (persons, things or processes) – like a mechanic about a machine, nurses for patients, a teacher about a classroom, and a welfare worker related to social security. An Uber driver, for instance, can be placed in all three categories; they could be data subjects, data generators (as owners of objects generating data – their cars), and also data workers as they fill in various kinds of data about people/things into the Uber app. These data-based locations and identities can entitle them to a bundle of economic data rights. If appropriately institutionalized and implemented, such rights can help them get a better share of the expanding digital economy pie to which they contribute, thus ensuring economic justice.

---

[11] Lessig, L. (2000, January 1). Code is law. Harvard Magazine. https://www.harvardmagazine.com/2000/01/code-is-law-html

[12] Actors who are responsible for generating data increasingly figure in new legislative initiatives in the EU like the draft Data Governance Act and the draft Data Act.

All the discussed terms, concepts, frameworks, and tentative implications require a strong theoretical grounding, which is still missing. It will pave the way to their effective practical application in digital economy practices and regulation.

3. Platforms as sites of data-based economic injustice

The majority of the world's economically useful data is – or would soon be, as all sectors get platformised – controlled by a few platform corporations, increasingly in a monopolistic position in different sectors. Economic rights around data would therefore mostly address, and be exercised in relation to, these key 'data holders'.

The central role of platforms in any economic justice framework becomes clearer as one shifts from the theoretical landscape to venues of policy and law. Still taking an expediency-based situational view of what appears fair in a given economic context, as against a more theoretically-rooted approach, governments' action has focussed on governing data with platforms. It can be categorized as taking place at four levels:

1. User/consumer level: Concerning data about users and consumers with platforms.

2. Economic actors level: About data with platforms of platform-dependent economic actors, like traders, farmers, SMSs, small service providers, and gig workers.

3. Public sector level: Relating to platforms-held data that may be needed for public sector economic activities – economic regulation, welfare, infrastructures, and economic development.

4. Geopolitical level: Involving cross-border data with platforms that is leading to fears of data colonization (not just in developing countries but also in the EU).

Platforms employ users' own data to manipulate them and inter alia put them at economic disadvantage. For instance, an e-commerce platform may increase prices, or show only costlier items, for an individual for a good that data indicates they are desperate/impulsive to buy at a given moment; or, finding the battery of a commuter's phone close to dying, a ride-hailing platform may increase the fare. Laws try to protect users and consumers against such exploitation, but are unable to keep up with the ever more innovative and complex manipulations. One approach is to give users/consumers the right to withdraw their data from a platform and take it to another service provider. This puts competitive pressure on platforms to maintain good behavior. In Europe, the General Data Protection Regulation (GDPR) provides data portability rights as an economic lever to users and consumers on platforms. EU's Digital Markets Act (DMA) further extends data related user rights. While offering no new substantive rights, the draft Data Governance Act (DGA) sets up procedural enablements for implementing data rights provided in other laws, including through special, certified, neutral entities called data intermediaries.

Platforms similarly manipulate their business users, i.e. various economic actors dependent on them. After getting a data-based lock-in, they arbitrarily keep changing the terms of engagement in ways adverse to these dependent economic actors. The latter's data can be used by platforms to launch competing products of their own. The DMA provides rights to various economic actors on very large platforms to claim their data, including aggregate data, and port it to third parties of choice. The draft Data Act (DA) gives data access and portability rights to owners or operators of IoT devices. It applies to consumers using datafied devices at home (say, a smart refrigerator) and outside (connected cars), as well as to SMEs generating machine data. The draft Act further mandates cloud services to be interoperable to ensure portability rights can actually be employed (as against the original self-regulation approach of the GDPR).

The public sector has traditionally been the repository of most of society's data. This enabled it to perform key public functions like regulation, welfare, and providing various infrastructural services. With platforms increasingly becoming the main holders of society-wide data, they are displacing the public sector from many of its roles, one of which is indeed to regulate and diffuse concentration of economic power that is currently held by platforms. Mandated access to data for public agencies figures prominently in some laws and policy documents in the EU,[13] France[14] and India.[15] This is a very important area for ensuring economic justice in the digital age.

Ninety percent of capitalisation of the world's top 70 platforms is held in the US or China.[16] Never before has the world witnessed such concentration of economic power, which, as things stand, is slated to only get worse. As these platforms provide digital services globally, they become the biggest mines for data from across the world, which helps them provide even better and more digital services. This circular process is likely to soon create a situation where the two global digital poles of US and China would effectively colonize the rest of the world by controlling its data and intelligence.[17]

At the geopolitical level, a lot of data localization efforts have an economic self-reliance rationale. It is a different matter that these are yet to be coupled with clear and effective digital industrial policies that can actually employ such data for enhancing domestic digital industry. Such a thinking however has begun to emerge in general vision documents. The vehement opposition by major developing countries to discussing any digital trade agreement at the WTO – especially in terms of the global free flow of data doctrine, belongs to this fourth level of economic policy view about data (geopolitical level). National data sovereignty is a concept increasingly heard in the EU, as it also occurs in policy documents of developing countries like Rwanda,[18] South Africa,[19] and India,[20] and more recently of the African Union.[21] UNCTAD has called for developing a global data governance framework, which addresses both economic and non-economic aspects of data.

---

[13] European Commission. (2022). *Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act)*. https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113

[14] France's Law for a Digital Republic identifies 'data of general interest' that can be taken by public bodies from private businesses. See https://www.wipo.int/news/en/wipolex/2016/article_0014.html

[15] Government of India. (2019). Personal Data Protection Bill. https://www.meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill,2018.pdf

[16] UNCTAD. (2019). *Digital Economy Report 2019*. https://unctad.org/webflyer/digital-economy-report-2019

[17] Lee, Kai-Fu. (2017, June 24). *The Real Threat of Artificial Intelligence*. The New York Times. https://www.nytimes.com/2017/06/24/opinion/sunday/artificial-intelligence-economic-inequality.html
"So if most countries will not be able to tax ultra-profitable A.I. companies to subsidize their workers, what options will they have? I foresee only one: Unless they wish to plunge their people into poverty, they will be forced to negotiate with whichever country supplies most of their A.I. software — China or the United States — to essentially become that country's economic dependent, taking in welfare subsidies in exchange for letting the "parent" nation's A.I. companies continue to profit from the dependent country's users. Such economic arrangements would reshape today's geopolitical alliances."

[18] Republic of Rwanda. (2017). National Data Revolution Policy. https://statistics.gov.rw/file/5410/download?token=r0nXaTAv

[19] Republic of South Africa. (2021). Draft National Policy on Data and Cloud. https://www.gov.za/sites/default/files/gcis_document/202104/44389gon206.pdf

[20] Government of India. (2021). Draft Report of the Committee of Experts on Non-Personal Data Framework. https://static.mygov.in/rest/s3fs-public/mygov_160922880751553221.pdf

[21] African Union. (2022). AU Data Policy Framework. https://au.int/en/documents/20220728/au-data-policy-framework

Many of these policy and legal efforts correspond directly with the earlier discussed rights based framework. After all, they are both aimed at similar objectives – ensuring fairness and justice in relation to data and data-based outcomes in a digital society. The underlying objective of all the discussed policy/legal efforts is to ensure 'equitable sharing of benefits' arising from data of individuals, groups, enterprises, and communities (including national and indegenous communities), and 'avoidance of harms' to them. Harms related jurisprudence about group data however is yet to take shape, although it figures in the draft report of India's Committee of Experts on Non Personal Data Framework (hereinafter CoE on NPD framework).[22]

As a practical way to move forward on the benefit sharing objective, most current policy/legal efforts focus on 'access to one's data' that is held exclusively by a few platforms. 'Equitable representation in data' has not yet got much policy attention, other than in public sector data (e.g. Estonia allows citizens to view their data with the government, and therefore possibly correct it), and some emerging work on data based discrimination. As for 'participatory data governance', by individual and collective data subjects, EU's Data Governance Act provides very useful facilitating conditions for data subjects and other rights holders to together manage access to and the uses of their data. The GAIA-X project in the EU on developing data infrastructures for cloud and data sovereignty involves participatory governance by cloud and data contributors. The draft report of India's CoE on NPD framework recommends mechanisms for community governance of data trusts.

However, theoretical and conceptual work on economic data justice, to the extent it exists, exhibits little if any explicit linkage to the policy and legal initiatives being undertaken in this area. In the absence of appropriate theoretical grounding, policies and law largely develop and act in a fire-fighting mode, remaining mostly ineffective. They tend to get quickly overtaken by the fast moving digital phenomena, which is inadequately anticipated by them. The two sides, of theoretical developments and policy work, need to come together.

[22] Government of India. (2021). Draft Report of the Committee of Experts on Non-Personal Data Framework. https://static.mygov.in/rest/s3fs-public/mygov_160922880751553221.pdf

4. Data infrastructures: The key data institution for economic justice

If platforms are the main sites of data related economic injustice, data infrastructures are the appropriate institution to correct it, and ensure economic justice. As mentioned, the main basis of platforms' extraordinary and unsustainable power is their dual and reinforcing roles as providers of digital services, as well as the exclusive miners and hoarders of a large portion of society's economically useful data. This vicious spiral can be interrupted if society's data is mandated (or otherwise incentivised) to be put into a set of sector-based data commons (or semi-commons that allow simultaneous interplay of graduated commons and private rights[23]), or data infrastructures. Any such legal mandate can be based on considering the society, or, as applicable, its communities, as the 'primary economic rights holders' for the respective collective data (or loosely the 'owners' of such data). After all, this data is a common social resource (or community resource) reflecting the attributes of, and relationships, in the society or community.

To avoid complexity, with regard to this discussion, we will largely skip the 'social versus individual' political contest that underwrites a great deal of political philosophy and practice, and appropriately so. Most data infrastructures aim to only collect and dispense non-personal data – although this separation is not easy to make or sustain. Is farm data or IoT data from a small family owned business personal data or non-personal data?. Due to the important public interest involved, new laws like the draft Health Data Space Regulation of the EU, apart from non personal data, also intend sharing de-identified personal data for specific uses but in certified safe and protected processing environments. Regimes will make different choices and decisions on data sharing and data commons depending on their perception of public interest, and political ideologies. For the purpose of this document, we will consider the institution of data infrastructure as an important mechanism for data related economic justice, and mostly avoid getting any further into discussing the important intersecting issues of individual privacy and freedoms. These are dealt with in greater length in the complementary social justice primer.

In the industrial era, many economic structures and activities were considered as infrastructural, to be provided by the public sector – either directly or as closely-regulated private utilities. These infrastructures were equally available to all, as public goods, on the top of which various economic actors could undertake competitive, market-based activities. A key reason that digital platforms have such immense power is that they have extinguished this distinction between the infrastructural part of the economy and the competitive, market-based suprastructural, They have integrated and captured both spaces. This separation must be reclaimed, starting with public goods approaches like data commons or data infrastructures, which are made equally available to all. These are needed both at national and global levels.

[23] Gurumurthy, A., & Chami, N. (2022). Governing the Resource of Data: To What End and for Whom? Conceptual Building Blocks of a Semi-Commons Approach. Data Governance Network. https://itforchange.net/sites/default/files/1741/WP23-Governing-the-Resource-of-Data-AG-NC.pdf

Taking a data infrastructure approach achieves three objectives, all of which contribute to economic justice.

1. With society's data available as a common non-rival resource to a large competitive set of digital businesses, it should enhance innovation and economic output.[24] This however requires patient and somewhat challenging work of chipping at and gradually transforming the dominant digital economy model that is based on end-to-end vertically-integrated platforms (with digital ecosystems bringing further integrations along new dimensions).

2. Claiming the key resource of data as a commons, and making it equally available to all, ensures much better distribution of economic power, and therefore greater economic justice. This can help ameliorate all four levels of data based economic injustice mentioned in the previous section.

3. Publicly, or commonly, managed data infrastructures provide an all-important site and lever for regulatory and legal interventions. Such interventions have become increasingly difficult if not impossible for end-to-end closed, all-private, digital platforms or ecosystems.[25] These exist as opaque and impregnable techno-structures, even as they enclose huge swathes of socio-economic activity in any sector. Data infrastructures become the open/commons/public part of society's digital architecture, interrupting its otherwise complete privateness and closedness. It thus provides useful techno-social sites and means for legal and regulatory intervention in digital code and architecture governing our societies (from the dictum 'code is law').

There still exists little to no theoretical work that spans and ties up ideas and concepts of data related economic injustice, society's and community's rights over its data, and community/public data infrastructures. Some policy/legal and programmatic efforts however are becoming visible regarding development of data infrastructures, often called by different names. These initiatives have mostly responded to emergent situations, in theoretically unsupported ways. The next section relates some policy/legal and programmatic initiatives that aim at redistributing data power across a larger set of players as against its exclusive control by a few global platforms. In some way or the other, all these connect closely to the concept and institution of data infrastructures.

[24] Most, though not all, of the data with the platforms is produced as a by-product of service delivery. Having to share data therefore does not necessarily disincentivize and reduce overall data production. In cases and areas where it does so, it can be factored in for data-sharing regimes, and sufficient incentives maintained for adequate data production.

[25] India's Universal Payment Interface, effectively under the Reserve Bank of India, is a meta-platform that all payment platforms in India have to plug into. This can inter alia provide a very useful site for enforcing federal banks' regulation on the financial system.

5. Governing data for economic justice

This section reviews some legal/policy and programmatic initiatives in the area of data related economic justice. It so happens that most of the important ones are either in the EU or India. This may have to do with both being relatively digital advanced but still facing huge domination by the two digital superpowers. And in this background, they are attempting to take control of and shape their digital destinies. They may just be a few steps ahead of the curve, and most other nations would soon reach this stage, and need to formulate clear economic policies and governance for data. Meanwhile, even in the US, a slew of legislations are lined up to check the power of Big Tech – with some data related provisions. The recent African Union's Data Policy Framework presents many very significant higher level points and principles about economic issues around data. However, it is only in the EU and in India that data governance from an economic point of view has taken legal form, or exists as policy or programs with sufficient clarity and impact.

All these laws, policies or programs address data concentration, seeking its redistribution by providing data access and control privileges to a variety of actors. Except for one or two, they do not explicitly use the terminology of rights, basing their provisions for data access on general principles of fairness, economic productivity, competition, public interest, and public functions. Some contain specific provisions related to data commons and data infrastructures, in some form. Others provide access rights or privileges, including the right of portability to third parties of choice, that can indirectly be employed by data rights holders to provision data commons or infrastructures.

EU's Data Strategy centers on creating sectoral data spaces, as a form of data infrastructures. (The conception of data spaces include data markets that work in an open, universally accessible, transparent, neutral, and regulated manner, which is also a kind of infrastructure, if paid for.) EU's GAIA-X project is a practical attempt to develop such data spaces, from below through collaboration among interested actors, without legal and regulatory help as of now. It is based on the idea of data sovereignty, defined as the "possibility of independent self-determination by the state and by organizations"[26] for "use(ing) and structuring of digital systems themselves, the data produced and stored in them, and the processes depicted as a result". It implies "complete control over stored and processed data and also the independent decision on who is permitted to have access to it".[27]

The X in GAIA-X stands for different sectors, like health, agriculture, transport and tourism. GAIA-X has been able to create some kind of data infrastructures in all these sectors, and more. The data infrastructures thus created are governed collaboratively by the key contributors of cloud facilities and data. These however must be EU headquartered entities and not foreign companies, underlining the idea of local community ownership.

[26] Government of Germany, Federal Ministry for Economic Affairs and Energy. (2019). Project Gaia-X: A Federated Data Infrastructure as the Cradle of a Vibrant European Ecosystem. https://www.bmwk.de/Redaktion/EN/Publikationen/Digitale-Welt/project-gaia-x.pdf?__blob=publicationFile&v=5

[27] Ibid.

GAIA-X offers open protocols and collaborative governance systems for those willing to bring their cloud and data to it, enabling self-determination in terms of one's data (data sovereignty). The draft Data Governance Act (DGA) goes a step further by providing data spaces legally-binding processes, enablements and certification that guarantee for all parties their neutral and transparent working in the interest of data contributors. Such an institution is legally called a data intermediary. If required, data infrastructures can also be facilitated for co-governance by data contributors, as 'data collaboratives', a term mentioned in the draft Act. The draft DGA does not provide any substantive data access rights, but those who have such rights can employ its provisions to ensure safe, neutral, certified and self-controlled data sharing, towards building data spaces and data infrastructures.

Substantive data access rights (whether explicitly called 'rights' or not) are provided in other EU laws. The GDPR provides rights to individual data subjects to access and port their data. The Digital Markets Act (DMA) enables business users of dominant platforms – like traders on an e-commerce platform, or cab drivers on a ride-hailing one – to access data with the platforms, including in aggregated forms, that is related to their products and activities, and to respective consumer interactions (subject to privacy protections). The draft Data Act (DA) allows IoT device owners or users to access data generated through these devices that may have been collected by an application or platform. Both these laws mandate that such data shall be transferred on request to third parties of choice. These could be various service providers (including competing services in the case of DMA but not DA), or data intermediaries/ infrastructures. The DMA and draft DA employ or imply the category of 'data generator', which joins the category of 'data subject' in the GDPR as holders of some kinds of data access rights.

It is important to recognise that there is very little a data subject or data generator can themselves directly do with the data that they are able to access. There can be some, one-off, individual uses, like contributing to one's 'right to repair'. But generally data is a systemic resource. It means that, to be of value, data mostly requires to be fed into a system working at considerable scale. Such a large system can be integrated and centrally controlled, as is the dominant model, or it could consist of federated, distributed, and vertically/ horizontally separated entities. It is not evident that individual data subjects or generators have sufficient incentive to take all the trouble of getting their data ported to alternative systems or a commons. This is well illustrated by the fact that GDPR's portability rights have hardly been used. Business users of platforms, and IoT owners/users, may have some inhouse uses of their data, but they are as unlikely to port it around to help alternative services, or building of data infrastructures, as individual data subjects did with their personal data using the GDPR. Even if they were to show any such inclinations, they are generally up against very powerful platforms that would easily offer enough strong incentives and disincentives to nip them.

In any case, rather than individual pieces of data it is the aggregate data patterns that may be most useful (for developing AI), and therefore platforms often anonymise and aggregate important data to escape individual rights based law, including privacy law.[28] Individual data benefit and access rights, whether for persons or enterprises, are therefore inadequate. Platforms can continue to exclusively derive immense benefits from the highly valuable resource of anonymised or otherwise aggregated data, which obviously cannot get covered

[28] Google, for instance, plans to phase out personalized cookies on its web browser Chrome and replace it by 'person types' based non-personal advertising targets.

under individual rights. Even though such aggregate data may entirely be built from the data of an assemblage of data subjects/generators, they are denied the 'right to benefit' from, or the 'right to access', what is indeed their data, and to avoid harm from it. Obtaining such due benefit and access (and avoiding harm) requires some kind of collective rights and agency for the relevant groups and communities; say Uber drivers in a city, patients of a rare disease in a region, or farmers of a rural district. Such groups/communities become the collective data subject of the aggregate data concerned, with corresponding rights on it. A practical means to exercise such a collective right will be to use it to collectively port aggregate data to a data commons or data infrastructure. This can support a competitive set of data-based services, serving the group/community much better than data monopolies do.

The draft report of the Indian CoE on NPD Framework recommends such a collective right of groups and communities to their data. These collective rights can be exercised to mandate large platforms to port relevant non personal data, pertaining to a particular group/community, to respective 'community trusts'. (The report lists many conditions and protections to safeguard genuine competitive advantages of data holders.) Acting as data infrastructures, these trusts make such data available to a wide range of actors – researchers, public and other social work agencies, and digital start-ups and small businesses. Community trusts will have some mechanisms for community based governance, but this remains inadequately outlined. The report also proposes a 'duty of care' for holders of non personal data, that can be legally invoked, and derelictions challenged, by any member of the community concerned.

The EU is now realizing that voluntary or self-regulatory approaches, or even formal rights of data access by themselves, do not ensure better distribution of data benefits. Data porting requires cloud interoperability, which the GDPR left to industry self regulation. The recent draft Data Act recognises that voluntarism does not work especially when at stake is the most valuable resource of the times, i.e. data. It now mandates cloud interoperability. But it is the EU's draft Health Data Space Regulation that takes the most far reaching steps towards an effective data infrastructure, with provisions that mandate sharing of data by all actors, private or public. This is the first sectoral regulation for enabling EU's data spaces under its Data Strategy, and others can be expected in the future. The draft Regulation mandates linking all health data repositories and exchanges, including of private service providers, to one grid managed by a set of public bodies. This ensures, actual and effective, universal real-time portability of data between different health service providers. It also creates a data infrastructure for inferred, anonymised, or otherwise non-personal health data that has immense system-wide value and use in the health sector. Data seekers can access such data against a data permit for specific laid-out purposes, with some uses clearly barred. Any data use and processing is to be undertaken within 'safe processing environments' provided by public bodies. What is envisaged is a fully-operative and really effective data infrastructure, distributing the benefits of data across the society, in an equitable manner.

It will be interesting to watch what path is taken by data space regulations for other sectors, each sector being quite different in terms of its requirements, public versus private/market nature, kind of public interest involved, etc. However, some kind of legally supported data infrastructures are expected to be created in all sectors.

[29] Holla, N. (November 2021). Democratising Technology for the Next Six Billion: India's 'Digital Public Goods' Innovation. ORF Issue Brief, 506. https://www.orfonline.org/research/democratising-technology-for-the-next-six-billion/

India has a very successful Digital Public Goods (DPGs)[29] approach that connects well to the concept of data infrastructures. The Data Empowerment and Protection Architecture provides an infrastructure for consent based personal data access. What is however most interesting and promising is the progresssive adaptation of the DPGs approach to different sectors, addressing their unique needs and challenges. Unified Payment Interface (UPI) is immensely successful as the national infrastructure for digital payments that is publicly owned and managed, and used by all digital payments services in India. It has enabled India to buck the global trend whereby digitalisation of payments systems has meant their end to end privatization and compartmentalization (like with WeChat and AliPay in China). This model of public digital network, or meta-platform, for nationwide digital interactions is now being taken to other sectors, like health, e-commerce and agriculture. Existing or new private platforms can plug into such a public network, both gaining from it and contributing to it

As this model is applied to sectors more complex than payment, it is facing the need for public/commons based data exchanges and data infrastructures. Some work in this regard is underway in India in the areas like health and agriculture. The National Open Digital Ecosystems (NODE) model of the central government also envisages common data pools in different sectors, with some kind of principles of fairness and reciprocity in data access, as well as collaborative governance by stakeholders. But this may not mean governance by data subjects or data generators, or otherwise the weaker parties in the data value chains, but by private sector holder-contributors and users of data. (A similar defect may also characterize the EU's GAIA project's collaborative governance model.) This underlines the need for evolving and relying on rights-based frameworks for participation of appropriate data rights holders in governance of data and data based socio-economic systems.

Lastly, considering economic justice at the geopolitical level raises issues of data sovereignty and data localisation. They are seen as policy instruments to ensure rule of law over a country's (or national community's) data. Data related economic justice is evidently difficult to ensure without data being subject to rule of law. Laws still remain mostly at the national level, and data may therefore need to be kept within their reach. The other option is appropriate international agreements that can assure countries the right to govern their globally-mobile and -distributed data. This has to be assured not just in terms of privacy and security, which concerns are now being internationally recognised,[30] but also with respect to economic and social justice, which are still not.

Developing countries at the WTO have been calling for recognition of data as a national resource, like natural resources, that can be devoted to strategic use for digital industrialisation.[31] The EU is putting conditions on global flow also of non personal data, which cannot be on privacy grounds.[32] It may also be questioned how, if there were to be free and unhindered global flow of data, would the EU enforce its data related economic regulations that were discussed. These include guaranteeing access for different domestic actors to data that is very likely to be collected by foreign platforms.

---

[29] Holla, N. (November 2021). Democratising Technology for the Next Six Billion: India's 'Digital Public Goods' Innovation. ORF Issue Brief, 506. https://www.orfonline.org/research/democratising-technology-for-the-next-six-billion/

[30] During its G20 Presidency in 2019, Japan introduced the concept of 'Data Free Flow with Trust', acknowledging the concerns of privacy and security with regard to cross-border data flows. See https://www.weforum.org/agenda/2022/05/cross-border-data-regulation-dfft/

[31] As can be seen in various submissions especially of the Africa Group to the WTO on ecommerce over the last few years. See https://docs.wto.org/dol2fe/Pages/SS/directdoc.aspx?filename=q:/Jobs/GC/144.pdf&Open=True

[32] Such restrictions figure in the draft Data Governance Act, and are also implied in the draft Data Act, both, of the European Commission.

UNCTAD recently proposed developing a global framework for data governance, covering both economic and non-economic issues.[33] The recent African Union's Data Policy Framework recommends "data to flow on the continent while safeguarding human rights, data protection, upholding security and ensuring equitable sharing of the benefits".[34] The Indonesian presidency of G20 this year has called for employing principles of fairness, transparency, lawfulness and reciprocity to cross border data flows.[35] G20's civil society group called the C20 recommended to the G20 leaders that discussions on cross-border data flows take into account not only the first generation of rights (civil and political, expressed in privacy and security), but also second generation rights (economic and social, like share in benefits from ones data), and third generation rights (collective rights, including the right to digital development).[36]

[33] UNCTAD. (2021). *Digital Economy Report 2021*. https://unctad.org/webflyer/digital-economy-report-2021
[34] African Union. (2022). AU Data Policy Framework. https://au.int/en/documents/20220728/au-data-policy-framework
[35] *Minister of Communications and Informatics Emphasizes Digital Sovereignty in Cross-Border Data Flow*. (2022, July 20). G20 Presidency of Indonesia. Retrieved from https://www.g20.org/minister-of-communications-and-informatics-emphasizes-digital-sovereignty-in-cross-border-data-flow/
[36] C20 Indonesia. (2022). *Civil 20 Policy Pack 2022*. https://civil-20.org/2022/wp-content/uploads/2022/10/POLICY-PACK-C20-2022.pdf

6. Future agenda for research and governance

This section lays out some theoretical gaps and practical challenges in the area of data and economic justice. It thereby suggests the needed agendas for future research, and for evolution of data governance .

Anchoring data laws in rights-based framework for economic justice

The need for moving away from ad hoc data laws and to anchor them in rights based frameworks for economic justice can be illustrated with regard to the institution of data infrastructures. It is the EU's draft Health Data Space Regulation that will establish the most advanced form of such an institution.The regulation lists two most important objectives, that are connected; (1) to ensure individual's right to benefit from their data, not just individual data but also from insights arising out of aggregate data, and (2) to provide effective, practical means for individual health data portability, which can actually implement the formal rights given in the GDPR for this purpose. These two objectives neatly map onto the first two elements of the theoretical framework for data related economic justice that was proposed in section two; namely, the right to benefit from one's data, and the right to access one's data including porting it to third parties of choice. The other two elements of the theoretical framework, the right to representation, and the right of participation, however do not figure in the Health Data Space Regulation. Further instantiations of the draft regulation should consider including these rights as well. It is required for the best, democratic, benefits from health data for all.

In general, it will be worthwhile to employ such a rights-based economic justice framework for developing and evaluating various data regulations. But this requires further development, substantiation, and honing of such rights-based frameworks.

Collective economic rights to data

The two 'rights' occuring in the draft Health Data Space Regulation exist in individual terms, as laid out in its introductory text. Much of the benefits in the context of health data space/infrastructure arise from secondary uses of data, and are collective. These cannot be adequately accounted for only in individual terms, as the draft Regulation attempts to. The draft Act does invoke the collective basis of 'public interest'. But that may be too loose and weak a concept (and perhaps legally inadequate as a basis of resource takeover), in a context that involves mandated sharing and accessing of the most important resource of our times. A collective rights based framing is much more appropriate, for getting benefits and accessing the important resource of data. Communities could also exercise a collective right to (appropriate) representation (and correcting data), and to participation in governance, directly or through their representatives. The draft regulation mentions national level public bodies, but the right to collective governance should also go to sub-national levels, and to appropriate interest-based communities, like communities of patients of a rare disease.

The disinclination of EU policymakers to step into the territory of collective rights, even when dealing with what is considerably a collective resource of social/community/group data, is equally evident in its horizontal data-related economic laws, like the DMA, the DGA and the DA. All these only provide data access rights to

individuals or individual enterprises, and correspondingly have significant limitations with regard to meeting their stated objectives of actually being able to redistribute data power in meaningful ways.

It was discussed earlier how it is extremely unlikely that individual holders of data rights will have the needed incentives to take the trouble of individually porting their data into any kind of data infrastructures, even if that would ultimately be beneficial to all. This is a kind of Prisoners' Dilemma, when individuals acting separately make choices that are sub-optimal for all, a well-known phenomenon of social behavior. GDPR's scantily used data portability right provides an adequate testimony to this. Business users of platforms and IoT owners/users would not act any differently. Traders on an ecommerce platform may employ the DMA to pull their individual/enterprise data and benefit from it individually to some extent. But what would really challenge the platform's data power, and greatly benefit all the traders on a platform, is if all traders were to have collective rights, and agency, over all their individual and aggregate traders' data on a platform. Employing collective agency, such individual and collective data can be taken to alternative platforms, or put into a commons/infrastructure, or simply used as a bargaining tool with the platform to improve the terms of engagement. It could even be employed to earn a role in co-governing the platform.

In fact, as per current and planned EU laws, it is theoretically possible that, say, all or most traders on an ecommerce platform employ the provisions of the DMA to separately pull their individual data, and then separately put it into a 'data collaborative', using the enabling law of the DGA.[37] But this is very unlikely to happen in practice at any appreciable scale, rendering the data access and use provisions of these laws largely ineffective in challenging platforms' monopoly data power. Theories and practice of trade unionism are instructive to show how institutionalized collective actions and bargaining succeeds best where the individual kind may not. This is even truer for data, with the marginal value of any single individual's or enterprise's data with a platform being extremely low. But the aggregate data of all or most together is indeed the most valuable resource that a platform has. It is this collective resource that has to be leveraged, that requires collective rights.

Confronting platforms' data power, therefore, requires collective data rights apart from individual ones. Such rights are also salient for different kinds of groups and communities that may want self-determination to be collectively represented in data in certain ways, derive certain kinds of collective data benefits, and avoid certain types of common harms. As mentioned earlier, the draft report of the Indian CoE on NPD Framework proposes collective data rights for groups and communities over their data. It also proposes some basic means of community participation in governing its data. Much more theoretical exploration is needed in these areas.

Data subject vs data holder – Who has the primary economic rights?

The EU also remains hesitant to clearly identify the primary economic rights holder for data (who can also be called the 'owner', in a very specialized sense as applying to a unique resource that is non-rival and usable in multiple, graduated, and complex ways – like the use of word 'property' in 'intellectual property'). The EU and other jurisdictions still address any situation of data-related economic unfairness through expedient solutions, without laying out basic rights and obligations of different kinds of data actors. Such a legal patchwork

[37] Some initiatives do employ individual rights to try and get data into common data trusts, for the benefit of data subjects. See https://www.workerinfoexchange.org/

can lead to long term structural distortions, entrenching default incumbent powers. This will only cause and exacerbate various types of injustices as a digital society matures. Ownership or property is not just as provided by law, it could be based on de facto power and private arrangements.[38] For all intent and purposes, platforms today own most of our data, It is this that forms the basis of their immense and growing power.

This default may be getting further strengthened by an unthinking and loose use in EU's data related legislations of the term 'data holder', defined as the actor who has the right and/or capability to provide access to data. This is done without providing any basis or definition of rightful possession and use, which seems to just imply not having done or doing anything illegal. With very little data law around, that begs the question as to what is legal or illegal, and the collective will of the society with regard to 'data ownership and control'. Such a legally perpetuated ambiguity begins to provide legal cover to the near blanket powers for data holders over society's data (especially the most valuable aggregate kind). Even as the current laws afford some data access rights to the weaker actors, they may be doing more damage in the long term. They are setting normative and structural conditions for a digital society where the collector and hoarder of data (as against the subject, and generator, of data – generally the weaker parties) gradually gets de facto and de jure established as the principal rights owner in relation to data, other than the rights specifically allocated to other actors by various laws in a somewhat ad hoc manner (and mostly without practical ways of actually employing them). Such legal patch-ups would always remain a losing proposition.

It may be better to take the converse route; the individual and collective data subjects be established by a foundational law as the primary holders of economic (and other) rights to the respective data. They can, loosely, be considered the owners of such data about them (like the appropropriation of the term 'property' in IP laws which also relates to intangible and nonrival resources). Some such rights may also be specifically allocated to data generators, as the primary owners or users of things, resources, or processes about which the relevant data is. Data worker is another potential rights-holding category that we discussed. Data holders can be allowed certain default privileges of data use (subject to privacy and other laws) that remain valid unless and until the ownership rights get invoked and employed by individuals or communities in a certain predetermined manner (which is appropriately cautious). This has to be undertaken without disrupting things in ways that are overall harmful to all or most.

Fears about the entire structure of digital society falling like ninepins if the dominant digital/data system is fiddled with are vastly overstated. Change towards greater data-based economic justice can be steered in gradual and pragmatic ways, preserving, and perhaps enhancing, digital effectiveness and productivity. Developing the social and political will to step out of the morass of current paralysis is the starting point here. But it also requires considerable theoretical work to look into questions like who should possess the primary economic rights to data, what justifications and precedents – moral, economic, and jurisprudential – can they be based on, and how can such rights be operationalised while protecting a high level of efficiency and productivity.

---

[38] Brown, N.D., Martens, B., & Muller-Langer, F. (2017). The economics of ownership, access and trade in digital data. European Commission. https://joint-research-centre.ec.europa.eu/publications/economics-ownership-access-and-trade-digital-data_en *"In his paper "Might makes right", Umbeck (1981) explains how private land property rights emerged during the California Gold Rush in the complete absence of any legal and state-enforced property rights. An equally distributed violence technology among miners (they all had guns) ensured that none could dominate the others and resulted in a fair distribution of mining parcels (Skaperdas, 1991 & 1992)."*

The foundational principle of data subjects (and to a lesser extent, data generators) as, individually and collectively, holding primary rights to data pertaining to them, can be duly legislated, and even constitutionalised. It can then be elaborated into a more detailed framework for its implementation. The four part rights-based framework proposed in section two – consisting of the right to benefit, right to access, right to representation, and right to participate – could directly result from such a foundational primary right over one's data. Laws and regulation for data related economic justice addressing different specific situations and relationships should be based on such rights, and evaluated against them.

The right to participate in governing data-driven economic systems

An important right for data subjects or data generators is the right to participate not just in data governance but also in the governance of data-driven digital systems, which control and cause actual data-based outcomes. This is of cardinal importance to protect the interests of weaker economic actors in a digital society – traders, farmers, small/individual service providers, MSMEs, workers, etc, and of marginalized groups and communities. As discussed earlier, alluring possibilities exist for these economic actors to claim co-governance of platforms (to the extent practically possible and pragmatic), which may already be implicit in some current and planned data related EU legislations. These need to be strengthened through collective data rights based explicit legal formulations that directly aim at and serve such a participatory governance objective. Adequate practical means need to be provided for implementation of various rights. This is what can best diffuse and redistribute platforms' digital and data power, to the benefit of not just the platform-dependent economic actors but all.

Most of the concern for platform-based or gig workers currently focuses on getting them employment status. But one must also explore their economic data rights. If indeed gig workers are not employees but independent parties (or are of an in-between status), they can leverage data rights like those provided in the DMA for business users of platforms. As discussed earlier, this opens up possibilities for participatory governance of the respective platforms by data rights wielding workers. Facing a lot of legal and political pressure, but not willing to give employment status to platform workers, some platform corporations have proposed a new category for these workers between employees and contracted independent parties.[39] Alternatively, what about considering platforms as an entirely new kind of economic actor, somewhere in between traditional private businesses and cooperatives? This can be achieved by employing innovative institutional designs that appropriately address and balance the two imperatives of innovation/productivity, on one hand, and distributive/economic justice, on the other. After all, the two kinds of resources that give rise to platforms – network resource and data/intelligence resource – are both a form of commons, provisioned respectively by interactions between, and data of, those who undertake business or work on the platform.

[39] Feiner, L. (2020, August 10). Uber CEO advocates for 'third way' to classify gig workers while fighting California labor lawsuit. CNBC. https://www.cnbc.com/2020/08/10/uber-ceo-op-ed-pushes-for-third-way-to-classify-gig-workers.html

Appropriate data rights for workers who are data subjects, and/or data generators (a concept that needs expanding to include various data work), need to be conceptualized with a view to ensuring economic justice for them. This can provide them a strong basis for collective bargaining, and even co-governance, as discussed. Separately, and in addition, it may also be considered if data workers should have rights on data and AI that they help produce that are akin to moral rights under IP law for creators to their intellectual creations, even if arising as a part of employment. When Google employees walked out against their work being used by the US military in the very sensitive area of drone warfare (and Google eventually relented), were these employees exercising moral rights over their data and intelligence work? This event inspired the formation of the Alphabet Workers Union. How can traditional institutions like trade unions be brought together with new formulations of workers' data rights, to shape new theoretical constructs and practical means for ensuring economic justice for workers?

The inalienable social/community embeddedness of data resource

With the digital society and economy representing a fundamentally new socio-economic paradigm, it is important to develop universal high level principles and rights for economic justice in this new context, as we had for the industrial society. The need is even higher because the digital economy is born global, and considerably organized in a global manner. This primer has focussed on such higher level, universal, issues, principles and rights. Like with global human rights, they are necessary to provide the appropriate foundation, and start, for ensuing various kinds of justice.

However, it remains important to recognise that data is a very special kind of artifact, and its resource nature is at best partial. It is in many ways an inalienable social and cultural part of the respective groups, communities and societies, as it is a part of personhood of individuals. It cannot, ordinarily, be removed from such social, community, and individual mooring or embedding.[40] Although in exploring precedents for establishing primary economic rights of communities over their data we took the example of natural resources, (most) data can never be removed from its social context and mooring as perhaps natural resources could be. Even the economic value, and harm, of data (mostly) remain linked to the social context of its creation. Economic theories and frameworks around data cannot just abstract it as simply a resource; they need to always remain cognizant of the inalienable social and personal aspect of data, even in relatively distant levels and spaces of its processing. This is not to deny that many kinds of data do have more general, and abstracted, applications as well. The point here is to underline the social and community embeddedness of data, which often gets ignored or under-appreciated in the digital economy discourse.

Taylor has recently argued about the unique nature of data, and of data markets, and a different kind of regulation that they therefore require. It is needed to keep a close eye on data's social context, especially the possibility of individual and collective harm. She recommends that data market regulation moves "away from a prioritization of the economic value of data and toward a more nuanced approach that aims to align the uses of data with the needs and rights of the communities reflected in it."[41]

---

[40] To note once again that 'data' is indeed a very broad term, and data is of many kinds. We are mainly speaking here about data that is most relevant to the digital economy, which generally concerns various kinds of social relationships.

[41] Taylor, L., Mukiri-Smith, H., Petrocnik, T., Savolainen, L., & Martin, A. (2020). (Re)making data markets: an exploration of the regulatory challenges. https://doi.org/10.31235/osf.io/pv98s

Economic justice frameworks therefore need to be highly contextualized to the needs and priorities of the particular communities and societies involved, and to different kinds of social and economic relationships, in all their specificities and nuances.

How to construct and address the resource nature of data as the first step towards ensuring an equitable distribution of its value, while maintaining its individual, community and society embeddedness and specificity is an area requiring much theoretical work.

Continuity with traditional institutions of economic justice

The welfare state is an indispensable institution for economic justice. Addressing the data needs of the public sector, and employing collective rights of communities and the society for this purpose, is an important area for theoretical and legal development. While the dangers of data-based authoritarian states developing all around us is absolutely real, it cannot lead to throwing the welfare state out along with the proverbial bath water. A strong data and digital intelligence based welfare state, that can adequately perform its regulatory, distributive, social security, infrastructural, and other functions in a digital age, needs to be theorized, and enabled by law.

In the opening section we mentioned various traditional institutions devoted to economic justice, like public infrastructures; welfare state; taxes; employment protections and trade unions; economic regulation like competition law and consumer protection; specific interventions in support of economic actors like SMEs, farmers, traders, artisans, etc; and, macro-economic measures like fiscal, monetary, industrial and trade policies. Most of this primer was devoted to the institution of public infrastructures in relation to data, with some other institutions like welfare state and trade unions touched upon. Data governance needs to be reimagined in relation to all the above listed traditional institutions of economic justice. Institutions can and must surely change, be transformed, or even abandoned in light of new possibilities, but sufficient normative and structural continuity is required, and transitions need to be planned well.

For instance, one needs to examine what effect the so-called distributed management of data through Web3 technologies, chiefly blockchain, has on economic justice. These are described as improvements over Web 2.0 that concentrated power by centralizing data controls. Data controls, it is claimed, would now be decentralized and distributed, thus democratizing economic power. More robust institutional thinking needs to be brought to such tall, and often unsubstantiated, claims. What effect, for instance, crypto-currencies would have on monetary and fiscal policies, and thereby on economic justice? Web3 is touted as being inherently participatory. Taking from the long traditions of participatory governance and democracy, if robust theories and practices are developed around the right to participate in governance of data and data based systems, it may help instruct Web3 enthusiasts on how participation based on the amount of resource-intensive technical work done (‘proof of work’) or the extent of assets held (‘proof of stake’) is really not participation but more like collaborative elitism.

Despite grand declarations to the effect, the globe is still not flat. As an expression of its ‘right to development’, each country has the right to frame and implement its digital industrial policies, an important traditional means for economic development in an uneven geo-economic landscape through appropriate interventions by the state. Key to such policies will be suitable data rights and data governance regimes, which therefore need to be theorized and developed also from such a perspective.

Global economic data rights for a global digital economy

Lastly, on geo-economics of data, and global trade frameworks, it may be required to get past convenient labeling of developing countries as seeking data localisation only out of authoritarian statist motives (which as well might also be there). The dangers of digital colonization are very real and extremely scary, attested nowadays by regular affirmations by EU leaders and policy documents about the need for digital and data sovereignty. If even Europe feels it, one can well understand what it would be for developing countries. It is not a threat just to the state, as some like to see it, but an existential danger for every single person, group, community, organization and enterprise in the countries facing this prospect, which happens to be all countries other than the US and China. Southern experience of industrial age colonization, with devastating economic and other kinds of injustices, is not that far back in time. Digital intelligence based dependencies are likely to be even worse than those that were based on uneven industrialisation.

Digital systems are globally organized, and there is a great merit and value in them remaining considerably so. This certainly requires easy global flow and exchanges of data. But theorists and researchers need to get out of their comfort zones and address upfront the apparently conflicting interests and requirements around global data flows. As discussed earlier, what may be relevant in this regard are international agreements that look at cross-border data flows not just from the standpoint of first generation human rights (civil and political), but also second generation rights (social and economic), and third generation rights (collective, including the right to development). Another theoretical area to explore is whether, due to the special nature of data as being socially rooted even as it is instantaneously globally mobile and distributed, the world can move in the direction of considering the place of origin of data (and thereby the community which is the primary rights holder) as exercising the primary jurisdiction over data, wherever data might physically be.

To end; economic justice in the digital age, and within it data related economic justice, needs to develop into a full blown independent area of research and theoretical exploration. It has to maintain close connections with rapid policy, regulatory and legal developments taking place across the world in this area. All such policy /legal work correspondingly needs to become more holistic and grounded in some foundational principles and rights. What is at stake here is the whole design of the new paradigm of a digital society, and the place of economic justice in it.