



Task Force 05

INCLUSIVE DIGITAL TRANSFORMATION

Equitable Data Governance in the Age of AI: A Data Justice Perspective

Alison Gillwald, Executive Director, Research ICT Africa (SA)

Carla Rodrigues, Research Coordinator, Data Privacy Brasil (BR)

Jai Vipra, Researcher, IT4Change (IN)

Jamie Fuller, Researcher, Research ICT Africa (SA)

Mariana Rielli, Director, Data Privacy Brasil (BR))



Abstract

Dominant approaches to governing advanced data-driven systems, like Artificial Intelligence (AI), adopt a negative regulatory perspective: they focus on preventing first-generation rights violations – particularly those of privacy and security – through a compliance and penalty regime. Although necessary, this approach alone cannot produce just results. Rather, a data justice approach to the regulation of data-driven systems is necessary to arrest the perpetuation of injustice and historical inequalities present in the data itself. Applied in the context of economics, a data justice approach provides a rationale for positive discrimination to deal with the differential impact of harms and the uneven distribution of opportunities associated with data-driven technologies. This brief assesses regulations for equitable data value creation, both private and public, with a focus on AI. Drawing on scholars like Diane Coyle, for governing data for broad economic benefit and social well-being, and Brett Frichmann, to discuss the economic characteristics of data, the brief will analyse effective data governance that is needed for fairness in availability, accessibility, usability, and integrity of data in this context, also addressing ownership concentration and its effects on trade and competition.

Introduction

AI is increasingly deployed in all areas of social and economic activity, including in the provision of public services, the management of labour, and in the mediation of social and economic engagement. As AI policies, plans, and frameworks develop at the highest levels of the UN and become the centerpieces of numerous international agendas, we must remember that at the core of most current uses of AI (e.g. machine learning techniques), is data. There is a plethora of poorly understood and ill-defined individual and collective data-related risks. Deploying AI technologies without addressing these risks will result in widespread harms to humanity, notably in the perpetuation of historic socio-economic injustice.

In economic terms, data can be understood as a public good in that it is inherently non-rivalrous and non-excludable¹ This means that, at the technical level, data is infinitely usable without detracting from another person's ability to use it, and there are no natural barriers to multiple people using the same data at once. The data upon which AI is built needs to be treated as a public good to avoid the exacerbation of data-related harms and economic inequalities²

¹ Frischmann, Brett M. *Infrastructure: The Social Value of Shared Resources*. (Oxford, New York: Oxford University Press, 2012).

² Coyle, Diane. *Markets, State, and People: Economics for Public Policy*. (Princeton University Press, 2020).

Diagnosis of the Issue

Advanced data-driven technologies such as machine learning have ushered in global monopoly platforms that extract massive amounts of data from users across the globe. The real-time granular intelligence collected from millions of online transactions, communications and interrelations daily, has relevance in every sector to create huge new efficiencies³. Currently, however, the power to amass and convert data into intelligence is highly concentrated by a few global tech corporations⁴. Big tech corporations are the key sites of data-based power, increasingly replacing public agencies as the main holders of society-wide data, and without the same levels of accountability⁵. Over 90% of the market capitalisation value of the 70 largest platforms is estimated to be concentrated in China and the United States of America. By contrast, Europe accounts for 4% and Africa and Latin America together account for 1%⁶.

³ OECD. *Artificial Intelligence in Society*. (Paris: OECD Publishing, 2019). doi.org/10.1787/eedfee77-en

⁴ Gillwald, Alison, Dewey Murdick, Kelle Howson, Parminder Jeet Singh, and Anita Gurumurthy. *GPAI Data Justice Policy Brief: Putting Data Justice into Practice*. (Global Partnership on AI, November 2022). <https://gpai.ai/projects/data-governance/data-justice-policy-brief-putting-data-justice-into-practice.pdf>.

⁵ Coyle, Diane, and Stephanie Diepeveen. *Creating and Governing Social Value from Data*. (SSRN: November 28, 2021). <https://doi.org/10.2139/ssrn.3973034>.

⁶ Gillwald, Alison, Dewey Murdick, Kelle Howson, Parminder Jeet Singh, and Anita Gurumurthy. *GPAI Data Justice Policy Brief: Putting Data Justice into Practice*. (Global Partnership on AI, November 2022).

The concern raised by this concentration of power is highlighted when viewed alongside growing evidence that supports there being a widening of digital inequalities, between and within countries, especially in the Majority World⁷. Such inequalities exist not only between those online and those offline (as is the case in a voice and basic text environment). There is also a significant disparity between those who have the technical and financial resources to use the internet actively and productively, therefore having access to meaningful connectivity, and those who are “barely” online, passively using tiny bits of data to communicate intermittently⁸.

The exclusion of people from online financial services, remote work and digital production makes them invisible, underrepresented and/or misrepresented in the data extracted by global monopoly digital platforms. The misrepresentation of individuals in the data sets of advanced data-driven systems has significant implications for their economic, political and social engagement⁹. Given the geographic patterns associated with the large swathes of people without, or with limited access to digital services,

⁷ Research ICT Africa, Digital Futures: South Africa’s Digital Readiness for the ‘Fourth Industrial Revolution’, Research ICT Africa, 2020. https://researchictafrica.net/wp/wp-content/uploads/2021/01/021220_Digital-Futures_SAs-Digital-Readiness-for-4IR_01.pdf.

⁸ UNCTAD. *Digital Economy Report 2019. Value Creation and Capture: Implications For Developing Countries*. (UNCTAD, 2019). https://unctad.org/system/files/official-document/der2019_en.pdf.

⁹ Timcke, Scott. *Capital, State, Empire: The New American Way of Digital Warfare*. (London, University of Westminster Press, 2017).

misrepresentative data sets in AI systems similarly have large impacts on nations as a whole, with respect to their global competitiveness or geopolitical positioning¹⁰. AI systems that are biased and unrepresentative lead to decisions and policies that fail to address or even worsen the needs and conditions of those misrepresented. Consequently, opportunities for economic advancement, education, and social mobility are diminished, perpetuating a lack of opportunity, irrevocably deepening economic disparities and entrenching cycles of poverty¹¹. Those at the intersection of multiple inequalities, and particularly rural, black women, are especially vulnerable to discrimination in algorithmic decision-making¹²¹³

Gaps in Existing Data Governance Policies

Existing data governance frameworks have so far failed to recognise the differential impacts of AI on individuals, communities and countries. These differential impacts

¹⁰ OECD. *Artificial Intelligence in Society*. (Paris: OECD Publishing, 2019).

¹¹ Singh, Parminder Jeet, and Anita Gurusurthy. *Economic Governance of Data: Balancing Individualist-Property Approaches with a Community Rights Framework*. SSRN Scholarly Paper. (Rochester, NY: Social Science Research Network, January 1, 2021). <https://doi.org/10.2139/ssrn.3873141>.

¹² Gillwald, Alison, and Partridge, Andrew. *Assessing the Gender Dimensions of Digital Inequality for Policy Action*. (New York: UN Women, 2022).

https://www.unwomen.org/sites/default/files/2022-10/_Background%20Paper_Alison%20Gillwald_Digital%20Inequality.pdf

¹³ OECD. *Bridging the Digital Gender Divide*. (OECD, 2018). <https://www.oecd.org/digital/bridging-the-digital-gender-divide.pdf>

require different kinds of risk mitigation. The G20 has an important role in remedying this situation. Policy reforms designed in the Global North may translate into different outcomes if merely imported into the Majority World. Further, policy reforms evolved in the Majority World – such as Digital Public Infrastructure, data cooperatives, or indigenous data stewardship – may find resonance in the Global North.

There are several key gaps in current dominant data policy frameworks, which ought to be addressed if there is to be a fairer distribution of the opportunities and risks associated with the roll-out of advanced data-driven systems. One key limitation is their almost exclusive focus on personal data, leaving non-personal data close to completely unregulated¹⁴. Personal data is information referring to an identified or identifiable individual; it is, in that sense, a contextual definition whereby pieces of information that may not directly correlate to a persons' identity, such as their ID or biometric information, can still be considered personal data. When information falls out of that scope, however, or when it is anonymized, it is non-personal and will be largely unprotected. A striking example of non-personal data that can be highly impactful to individuals and communities, possibly furthering inequalities, is environmental or spatial data used to inform public policies and services ranging from housing distribution to air quality control¹⁵. At the same time, as issues of hyperconnectivity and the possibility of re-

¹⁴ Gillwald, Alison, Dewey Murdick, Kelle Howson, Parminder Jeet Singh, and Anita Gurumurthy. *GPAI Data Justice Policy Brief: Putting Data Justice into Practice*. (Global Partnership on AI, November 2022).

¹⁵ Barreneche, Carlos, and Lombana-Bermudez, Andres. “Another Infrastructure Is Possible: Grassroots Citizen Sensing and Environmental Data Justice in Colombia.”,

identification become more pervasive, the very distinction between these categories is called into question and gives way to a more nuanced approach that considers all data and their economic and social impact.¹⁶

Another limitation of data governance frameworks within the multilateral system, and endorsed by the G20, stems from their focus on ensuring that data governance is people-centered and rights-preserving. AI frameworks, such as the UNESCO AI Principles and OECD AI guidelines, go a long way in establishing strong first-generation rights protections and frameworks for ethical design. However, such frameworks by themselves do not produce just outcomes. This is because first-generation rights, like freedom of speech, assembly, and voting cannot be fully realized without the support of second-generation rights. So far, second-generation rights, which address issues of social justice and economic equality, have not received adequate attention in dominant data governance frameworks. Second-generation rights include the right to education, health care, and an adequate standard of living, which can be significantly undermined by inequitable and unjust outcomes of automated decision-making systems that rely on data. At the same time, the significant economic opportunities made possible through technology are missed by those who lack digital access or the requisite literacy to leverage digital tools and services productively, as is the case for millions in the Majority World. A sole focus on first-generation rights in the employment of AI is ineffectual in the context of

International Journal of Communication 17 (22). ((University of Southern California, 2023). <https://ijoc.org/index.php/ijoc/article/view/18821>

¹⁶ Taylor, Linnet, Hellen Mukiri-Smith, Tjaša Petročnik, Laura Savolainen, and Aaron Martin. 2022. “(Re)Making Data Markets: An Exploration of the Regulatory Challenges.” *Law, Innovation and Technology* 14 (2): 355–94. doi:10.1080/17579961.2022.2113671.

pervasive socio-economic injustice at a structural level. Rather, it is necessary to ensure that second-generation rights are recognised to advocate for state responsibility in providing essential services and creating conditions for collective well-being, as well as remedying discrimination and injustice stemming from data-driven systems in order to address the root causes of inequality.

Moreover, the practice of data governance has mostly been approached from a negative regulatory perspective: current approaches prioritise preventing first-generation rights violations of data subjects – particularly privacy and security- through a compliance and penalty regime. And yet, there are many areas of data governance (including data availability, accessibility and integrity, as well as concerns about ownership), that require positive regulatory or governance intervention¹⁷. Positive discrimination that is sensitive to marginalised categories (including but not limited to gender, race, and age) is necessary to redress intersectional inequality, particularly with respect to meaningful access to and use of advanced data-driven tools, like AI. The notion of meaningful access looks beyond “mere” access, to take into consideration various other aspects enabling an individual to productively employ digital solutions, including the affordability and quality of such services, while also accounting for the protection of basic rights like privacy¹⁸. To

¹⁷ Coyle, Diane, and Stephanie Diepeveen. *Creating and Governing Social Value from Data*. (SSRN: November 28, 2021).

¹⁸ World Bank. *Data for Better Lives*. International Bank for Reconstruction and Development (Washington DC, 2021).

<https://www.worldbank.org/en/publication/wdr2021#:~:text=World%20Development%20Report%202022>

promote more equitable and just outcomes, positive discrimination in accessing data, consumer protection, data protection, public procurement and data access and sharing is required. Ensuring historically marginalised groups gain access to foundational digital and data infrastructures and services in order to be better represented is the primary way to deal with bias in the giant datasets that dominate commercial activity.

At the same time, it is important to recognize that another concern overlooked in dominant data governance frameworks centers on their limited focus on the collective harms of data specifically on indigenous communities. Indigenous data sovereignty has been identified as a key concern by indigenous scholars, in advancing a decolonial data policy agenda. Such communities may require limitations on access to their data and knowledge systems, in order to guard against commercial and scientific exploitation of knowledge extraction¹⁹. This calls for nuance in advocacy around open data sharing which attempts to respond to the concentration and enclosure of data assets. Without responsible stewardship mechanisms, open data carries risks not only for personal privacy, but for the appropriation of indigenous knowledge systems and for furthering other types of injustices that may be embedded in the data itself.

1%3A%20Data%20for%20Better%20Lives%20explores%20the,individuals%2C%20bu
sinesses%2C%20 and%20societies.

¹⁹ Gillwald, Alison, Dewey Murdick, Kelle Howson, Parminder Jeet Singh, and Anita Gurusurthy. GPAI Data Justice Policy Brief: Putting Data Justice into Practice. (Global Partnership on AI, November 2022).

Recommendations

1. In addition to upholding first-generation rights of privacy, freedom of expression and access to information in the regulation of data, G20 positions on data governance frameworks should more actively promote second-generation economic rights, and third-generation, planetary rights in the regulation of data to rectify the uneven impact of harms associated with data-driven digital technologies, in particular AI. Several areas of data governance such as data availability, accessibility, usability, and integrity, as well as concerns about ownership, impacts on unfair trade and competition require positive regulatory and/or governance intervention.

2. Economic regulation is needed to redress the uneven distribution of opportunities currently associated with data-driven technologies such as AI. By ensuring access to data and digital skills, economic regulation can create more equitable outcomes for intensifying global processes of datafication, creating both private and public value. This includes, for instance, globally coordinated antitrust action, mandatory data sharing policies, and enforceable data minimisation rules. For all these actions, global coordination among regulators is imperative, without which large multinational technology firms can force countries towards a deregulatory race to the bottom.

3. To support more public-interested and equitable outcomes from these processes, G20 could promote the more active balancing of the current dominant commercial, supply-side valuation of data used in the allocation of global resources, which has produced these uneven outcomes, with more demand-side valuation in the allocation of resources that recognises their social value, including as common goods and downstream inputs.

4. Global cooperation is needed in the implementation and enforcement of data governance, which respects national sovereignty and avoids the involuntary incorporation of states into multilateral systems of governance in which they have no say and are perceived not to serve their national interest. With the representation of large numbers of non-G20 states in the organisational membership of the African Union in the G20, many of those whose voices are absent from forums of global governance or even in all-embracing initiatives such as the Global Digital Compact or the World Summit on the Information Society (this year in its 20-year review, in which digital inequality remains a central issue) there is an opportunity to forge common ground on harnessing AI for humanity in ways that will be truly transformative.

Possible Scenarios

In Scenario 1 the status quo continues with the G20 member states aligned with the multilateral guidelines on human-centred, rights-preserving data governance frameworks, applied to AI in a ‘responsible and ethical’ way. A minority of people are able to exercise the rights and opportunities arising from better-regulated environments with institutional capabilities that can reap the benefits of AI for their citizenry. But without global institutional arrangements to implement these frameworks or enforce them outside of the European Union, the largely unmitigated harms associated with power and information asymmetries underpinning AI systems across the globe continue. Consequently, the majority of the people in the world continue to be unevenly impacted both by exclusion and by their invisibility, underrepresentation or discrimination in the millions of automated AI decisions made every day on platforms and applications largely owned by big tech companies.

In Scenario 2 - G20 are able to provide leadership and secure global commitments to redressing digital inequality and data injustices by operationalising a global system of governance for the realisation of global digital public goods (including data) at the national and regional level where scope and scale are required. This would provide the foundations for the development of digital public infrastructure proposed by the G20 in response to the concentration of data infrastructure in the hands of a few giant corporations, accountable to no state, not even for purposes of taxation, national security or the common good. In order for this to succeed in creating more just outcomes, global commitments require unparalleled investments (and other forms of solidarity such as the writing off of sovereign debt, that like vaccine solidarity never materialised during the COVID-19 pandemic) in human development necessary to harness advanced data-driven technologies for public and private value creation and to accelerate progress towards SDGs.

Conclusion

As AI technologies become ever-pervasive, efforts to ensure universal and meaningful digital inclusion become ever more pressing, and more complex. Digital inequalities are layered over structural inequalities. Without addressing fundamental structural inequalities that are reflected in the data that feed AI systems, the roll-out of AI technologies will deepen existing inequalities. To ensure an equal distribution of the opportunities and risks associated with data-driven systems, policy interventions must address gaps in the foundational infrastructure of data-systems, and more importantly, in the human development gap that prevents the meaningful use and deployment of data-driven systems.

References

- Barreneche, Carlos, and Alejandra Lombana-Bermudez. 2023. "Another Infrastructure Is Possible: Grassroots Citizen Sensing and Environmental Data Justice in Colombia." *International Journal of Communication* 17: 22.
- Coyle, Diane. *Markets, State, and People: Economics for Public Policy*. Princeton University Press, 2020.
- Coyle, Diane, and Stephanie Diepeveen. "Creating and Governing Social Value from Data," November 28, 2021. <https://doi.org/10.2139/ssrn.3973034>.
- Frischmann, Brett M. "Infrastructure: The Social Value of Shared Resources". Oxford, New York: Oxford University Press, 2012).
- Gillwald, Alison, Dewey Murdick, Kelle Howson, Parminder Jeet Singh, and Anita Gurumurthy. "GPAI Data Justice Policy Brief: Putting Data Justice into Practice." Global Partnership on AI, November 2022 <https://gpai.ai/projects/data-governance/data-justice-policy-brief-putting-data-justice-into-practice.pdf>.
- Gillwald, Alison, and Partridge, Andrew. "Assessing the Gender Dimensions of Digital Inequality for Policy Action." New York: UN Women, 2022. https://www.unwomen.org/sites/default/files/2022-10/_Background%20Paper_Alison%20Gillwald_Digital%20Inequality.pdf.
- OECD. "Artificial Intelligence in Society." Paris: OECD Publishing, 2019. doi.org/10.1787/eedfee77-en.
- . "Bridging the Digital Gender Divide," 2018. <https://www.oecd.org/digital/bridging-the-digital-gender-divide.pdf>.
- RIA, Alison. "Digital-Futures_SAs-Digital-Readiness-for-4IR_01.Pdf." Policy Paper. Research ICT Africa, 2020. <https://researchictafrica.net/wp/wp->

content/uploads/2021/01/021220_Digital-Futures_SAs-Digital-Readiness-for-4IR_01.pdf.

Singh, Parminder Jeet, and Anita Gurumurthy. “Economic Governance of Data: Balancing Individualist-Property Approaches with a Community Rights Framework.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, January 1, 2021. <https://doi.org/10.2139/ssrn.3873141>.

Taylor, Linnet, Hellen Mukiri-Smith, Tjaša Petročnik, Laura Savolainen, and Aaron Martin. 2022. “(Re)Making Data Markets: An Exploration of the Regulatory Challenges.” *Law, Innovation and Technology* 14 (2): 355–94. doi:10.1080/17579961.2022.2113671.

Timcke, Scott. *Capital, State, Empire: The New American Way of Digital Warfare*. London: University of Westminster Press, 2017.

UNCTAD. “Digital Economy Report 2019. Value Creation and Capture: Implications For Developing Countries.” UNCTAD, 2019. https://unctad.org/system/files/official-document/der2019_en.pdf.

World Bank. “Data for Better Lives.” International Bank for Reconstruction and Development / The World Bank. Washington DC, 2021. <https://www.worldbank.org/en/publication/wdr2021#:~:text=World%20Development%20Report%202021%3A%20Data%20for%20Better%20Lives%20explores%20the,individuals%2C%20businesses%2C%20and%20societies>.



Let's **rethink** the world

