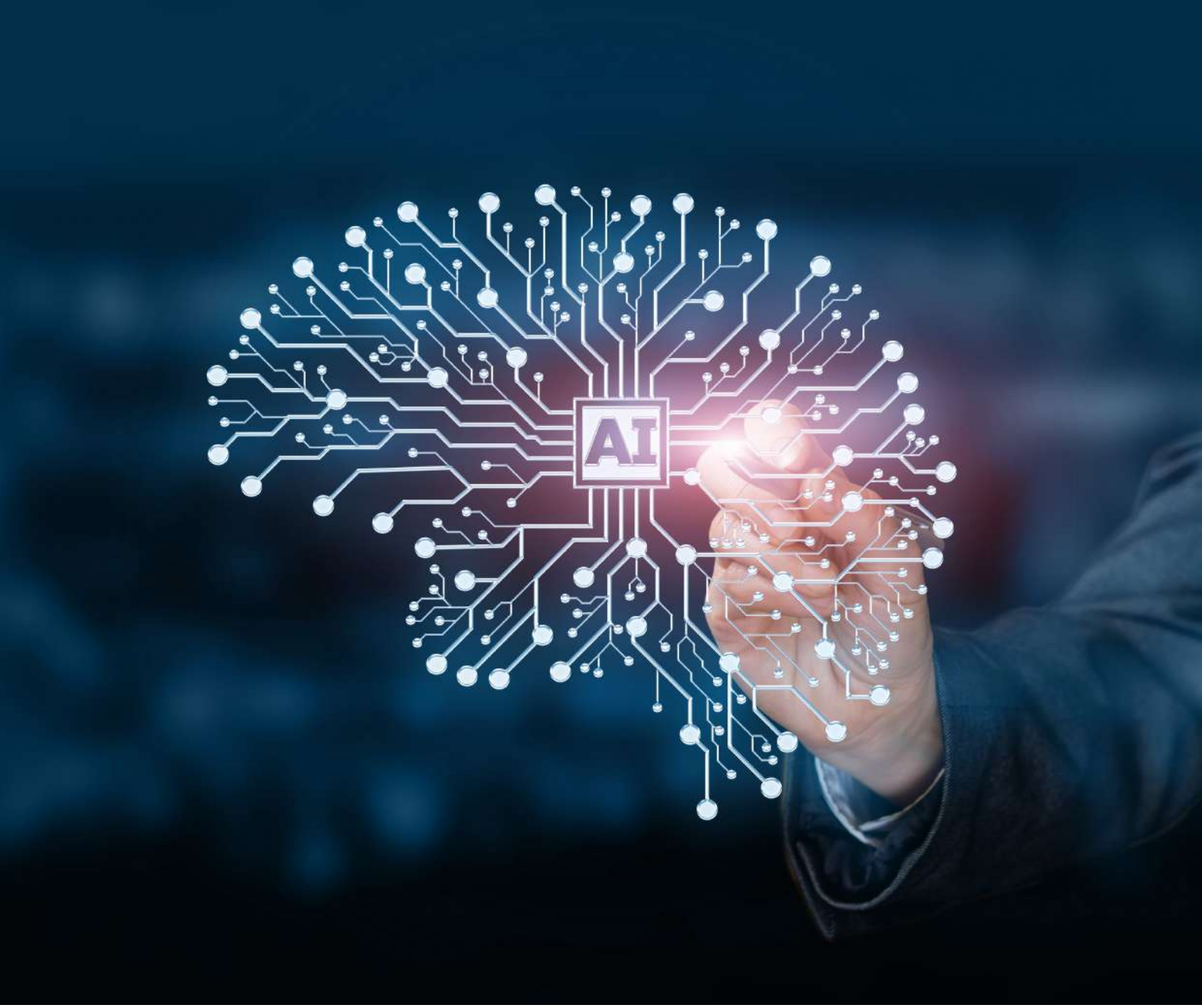


BEYOND THE NORTH-SOUTH FORK ON THE ROAD TO AI GOVERNANCE:

An Action Plan for Democratic & Distributive Integrity

REPORT | MARCH 2022

PARIS PEACE FORUM de PARIS sur la PAIX

In a world requiring more collective action, the Paris Peace Forum is a platform open to all seeking to develop coordination, rules, and capacities that answer global problems.

Year-round project support, policy development activities, and an annual event each November help better organize our planet by convening the world, boosting solutions, and incubating initiatives.

Initiate: Digital Rights in Society enables the establishment of a global multi-stakeholder process to address issues raised by algorithmic governance. This effort engages governments and policymakers, civil society organizations, and academic scientists and engineers from across Europe, North America, and, crucially, the Global South.

The organization is working to chart possible pathways toward an international digital rights framework, engaging civil society in determining the trajectory of automated technologies and their application, as well as helping to develop regulatory approaches that strengthen collective rights everywhere in the world.

Initiate is a non-profit organization based in Cambridge, Massachusetts incubated by the Paris Peace Forum and supported by a generous grant from Luminate.

Initiate
DIGITAL RIGHTS IN SOCIETY

PARIS PEACE FORUM
de PARIS
sur la PAIX

BEYOND THE NORTH-SOUTH FORK ON THE ROAD TO AI GOVERNANCE:

An Action Plan for Democratic & Distributive Integrity[1]

REPORT | MARCH 2022

[1] Acknowledging that the categories of South and North are not watertight, this paper argues for situating geopolitical and geo-economic power within the history of post-colonial development.

ABOUT THE AUTHORS

This publication is informed by discussions and consultations within a 21-member expert working group exploring the convergences and divergences in algorithmic governance approaches between the Global South and North. The group convened regularly through the second half of 2022. Its members are:

Writing Committee

Anita Gurumurthy	IT for Change (India)
Chioma Agwuegbo	TechHer (Nigeria)
Nandini Chami	IT for Change (India)
Virgilio Almeida	Berkman Klein Center for Internet and Society, Harvard University

Working Group

Abdul-Hakeem Ajijola	National Cybersecurity Policy and Strategy Review Committee (Nigeria)
Abhishek Gupta	Montreal AI Ethics Institute, Microsoft, Green Software Foundation
Anita Gurumurthy	IT for Change (India)
Baobao Zhang	Maxwell School of Citizenship and Public Affairs, Syracuse University
Chioma Agwuegbo	TechHer (Nigeria)
Derya Özkul	Refugee Studies Centre, University of Oxford
Emre Eren Korkmaz	Department of International Development, University of Oxford
Francisco Brito Cruz	InternetLab (Brazil)
Jean F. Queralt	IO Foundation (Malaysia)
Juan Carlos Lara	Derechos Digitales (Chile)
Jun-E Tan	Independent Researcher (Malaysia)
Malavika Jayaram	Digital Asia Hub (Hong Kong)
Moussa Kondo	Accountability Lab (Mali)
Nagla Rizk	American University in Cairo
Nandini Chami	IT for Change (India)
Nighat Dad	Digital Rights Foundation (Pakistan)
Peggy Hicks	Office of the United Nations High Commissioner for Human Rights
Roxana Radu	Blavatnik School of Government, University of Oxford
Shashi Jayakumar	Centre for Excellence for National Security (Singapore)
Urvashi Aneja	Digital Futures Lab (India)
Virgilio Almeida	Berkman Klein Center for Internet and Society, Harvard University

Research Team

Sadhana Sanjay	IT for Change (India)
Tanay Mahindru	IT for Change (India)

Secretariat

Adrien Abécassis	Paris Peace Forum
Merih Angin	Initiate: Digital Rights in Society
Elina Noor	Koç University (Turkey)
Jack Loveridge	Initiate: Digital Rights in Society

TABLE OF CONTENTS

Executive Summary 6

Recommendations 8

I. AI Governance at a Crossroads: Fragmentation vs. Coordination 9

II. Erosion of the Civic-Public Space: Why AI Governance Needs a Paradigm Shift 12

III. Justice in the AI Economy: Fair Value Distribution & Implications for Development 16

IV. AI Constitutionalism: Advancing Cooperation in Governing the Data 19

Conclusion 21

Bibliography 24

EXECUTIVE SUMMARY

Artificial intelligence (AI) is transforming the world faster than the world can mitigate intensifying geopolitical divisions and socio-economic disparities. As technological change outpaces regulatory policy, no common platform has yet emerged to coordinate a variety of governance approaches across multiple national contexts. The concerns and interests of the citizens and civil society of the Global South – broadly, the post-colonial nations of Latin America and the Caribbean, Africa, the Middle East, South and Central Asia, and the Asia-Pacific – must be prioritised by policy makers to reverse increasing fragmentation in the governance of algorithmic platforms and AI-powered systems worldwide. Particular attention must be given to the varied ways in which national governments and transnational corporations deploy such systems to monitor, manage, and manipulate civic-public spaces across the Global South.

The Global South represents a major source of the human-generated data and, indeed, the very raw materials upon which complex computing networks and AI systems rely. It therefore follows that the societies of the Global South are entitled to both equitable economic benefits and meaningful protections from powerful platforms and tools largely controlled by



corporations based in the Global North and the great powers, particularly the United States (US) and the People's Republic of China, but also the European Union (EU). This equity must be predicated upon what we define as an 'AI constitutionalism' that approaches AI and big data as fundamental resources within the modern economy akin to electricity and water, essential components for economic and social development in the 21st century.

Redirecting ongoing AI ethics discussions toward a rights-based paradigm with concrete principles for policy across

national contexts offers the best prospect for an international governance framework that places the interests of Global South on equal footing with those of the great powers. Informed by discussions within a 21-member expert working group convened regularly by Initiate: Digital Rights in Society and the Paris Peace Forum through the second half of 2021, this paper proposes that AI constitutionalism and a rights-based approach should guide the development of high-level international protocols and conventions that will set policymaking standards for AI's development and deployment, worldwide.

Special protections must be developed, whether at an international level or by national governments, to safeguard civic-public spaces across the Global South and outline best practices for assessing the potential impacts of AI-based services before their deployment. More generally, international norms and agreements must be established to ensure the equitable distribution of the benefits of AI-powered platforms and algorithmic systems, avoiding locking Southern countries into loops of dependency. A 'fair value distribution' regime between the Global South and the rising AI powers must be pursued and achieved.



RECOMMENDATIONS

To work toward a less fragmented AI policy landscape that incorporates the interests and concerns of Global South countries, the working group recommends several key actions*, including:

1 | **Sustaining a running global discussion of digital rights concerns and standards in AI policy that prioritises civil society, expert, and policy maker representation from the Global South.**

2 | **Building a global database of proposed and implemented AI legislation, focusing especially upon the nations and subnational units of the Global South.**

3 | **Global South countries should consider implementing the requirement of AI impact assessment reports.**

4 | **Create incentives to retain young AI scientists and engineers in the Global South.**

* See more information on the recommendations on page 22 of this report.

I. AI GOVERNANCE AT A CROSSROADS:
FRAGMENTATION VS. COORDINATION

Today's emerging artificial intelligence (AI) governance landscape is highly fragmented.[2] Over 160 sets of artificial intelligence ethics and governance principles currently exist, but no common platform brings these different initiatives together (Report of the Secretary-General, 2020; Radu, 2021). Private sector and governments have relatively even input in these AI governance initiatives while civil society organisations have less robust representation (Ulnicane et al., 2021). Further, there is an overwhelming geographic disparity in norm-setting around AI.[3]

Notably, most of these guidelines originate from wealthy Organisation for Economic Cooperation and Development (OECD) nations while voices from the Global South remain poorly represented (Haas et al., 2020). Reviews of existing frameworks suggest that equality and non-discrimination, transparency, accountability, safety, social well-being, privacy, human dignity, and autonomy constitute the common core of normative concerns in the global conversation on AI governance (Fukuda-Parr et al., 2021).

While initial conversations on AI governance mostly unfolded in silos, with technologists focusing on solutionism in "the machine learning model, the inputs, and the outputs" (Aizenberg et al., 2020), key recent events[4] have paved the way for an ethical turn in which not only technologists, but also public policy actors, civil society activists and Big Tech corporations actively participated. Unfortunately, in the absence of enforceable standards and accountability measures, the moral values embodied in the human rights discourse too often end up being deployed as mere rhetorical devices within these guidelines (Fukuda-Parr et al., 2021) – resulting in an open-ended, anything-goes, ethical practice.

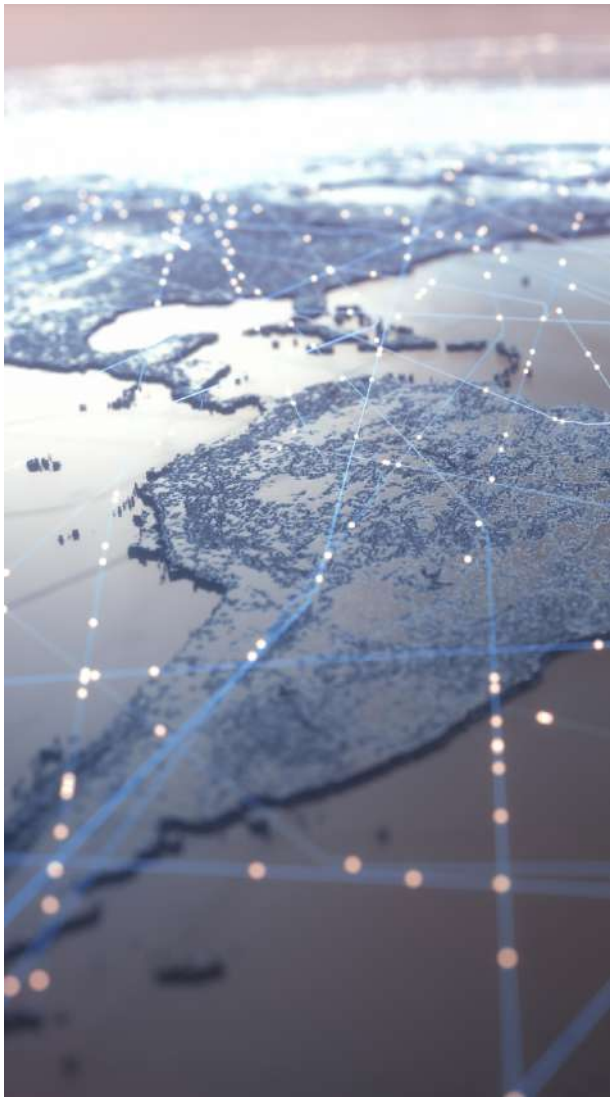
[2] See https://oecd.ai repo.

[3] See https://www.technologyreview.com/2020/09/14/1008323/ai-ethics-representation-artificial-intelligence-opinion.

[4] Such as the Cambridge Analytica scandal (2016) which lifted the lid off the risks of the algorithmified public sphere for democracy; the exposés of Project Maven and Project Dragonfly (2018-19) that alerted the wider public to the new military-industrial complex, and growing disquiet about algorithmic discrimination in welfare systems and the UN Special Rapporteur, Philip Alston's, investigation on the digital welfare state (2019).

Fortunately, recent conceptual explorations in AI governance reflect a necessary techno-social interdisciplinarity, albeit from a select few industrialised contexts, connecting, for instance, intelligent automation and the future of work; algorithmic public sphere and democratic life and citizens' rights and the digital welfare state (Gurumurthy et al., 2019). Yet, without a corresponding institutional arrangement for clear and enforceable obligations and commitments in the AI governance ecosystem, the policy impacts of this ethical turn may well be limited. A rights-based AI governance paradigm[5] with workable remedies for consumers and citizens – especially, vulnerable individuals and groups implicated in AI systems across the world – is thus an urgent imperative.

Rising socio-economic inequality and the intensification of the labour-capital divide in the structural transformation wrought by the current hyper-capitalist AI paradigm pose twin concerns for the socio-economic rights of the majority the world over (Acemoglu et al., 2020; Bughin et al., 2019). Emerging evidence also shows that the histories and geographies of colonialism have structured the international politico-economic order of the AI age (Mohamed et al., 2020), indelibly influencing the right to development for nations and peoples across the Global South. In today's AI economy, most developing countries are mere sources of the new raw material of data, while also proving dependent on the Global North for AI infrastructure and services (Feijóo et al., 2020). Critically, these countries are also sources of physical raw materials that are used to create and power AI systems.



[5] It is important to recognise the limitations of rights-based regimes in countries with weak institutional and regulatory capacities. Rights-based perspective may also not be able to adequately deal with structural and collective harms.



Critiques of algorithmic systems in the context of the North-South problematic have been varied, including: the overwhelming 'whiteness' of algorithmic decision systems (Cave et al., 2020); intensification of global labour hierarchies in the transnational data value chains that power AI business models; and the export of dubious, rights-violating AI product-testing to countries with less robust legislative frameworks are all manifestations of an 'algorithmic coloniality' (Mohamed et al., 2020), representing the exploitation and dispossession of the Global South in the emerging AI-driven international order. A rights-centred AI governance system must therefore be particularly attentive to socio-economic rights as they arise in the international political economy of development, straddling all generations of human rights.

The compact between the state and the market under global data capitalism is an important political arena wherein contestations for a just world order are already emerging. This paper argues for reclaiming the AI paradigm and shifting it towards democratic and distributive integrity, tracing common concerns as well as identifying fault lines to which progressive civil society in the Global North and South must attend.

II. EROSION OF THE CIVIC-PUBLIC SPACE: WHY AI GOVERNANCE NEEDS A PARADIGM SHIFT

AI is transforming the structures of collective choice through which social policy outcomes are generated in contemporary democracy, refashioning the state's exercise of political power (Risse, 2021). This transformation holds the potential of concentrating ever greater power in fewer hands. The automated public sphere is a fount of disinformation, hate speech, computational propaganda, and information warfare. There is copious evidence that user engagement-maximising algorithms at the heart of the social media business model are amplifying highly polarising content and hate speech (Dasgupta, 2021). Hate, xenophobia, and incitement to violence on social media platforms are on the rise. As the United Nations (UN) Special Rapporteur on Minority Issues observed in early 2021, three-quarters or more of the victims of online hate speech are members of minority communities (Office of the United Nations High Commissioner for Human Rights, 2021). Online and sexist hate has also snowballed to unprecedented levels during the global COVID-19 pandemic (Dehingia et al., 2021).

Platform self-governance dependent on a combination of human and AI moderation has fared poorly with respect to ensuring



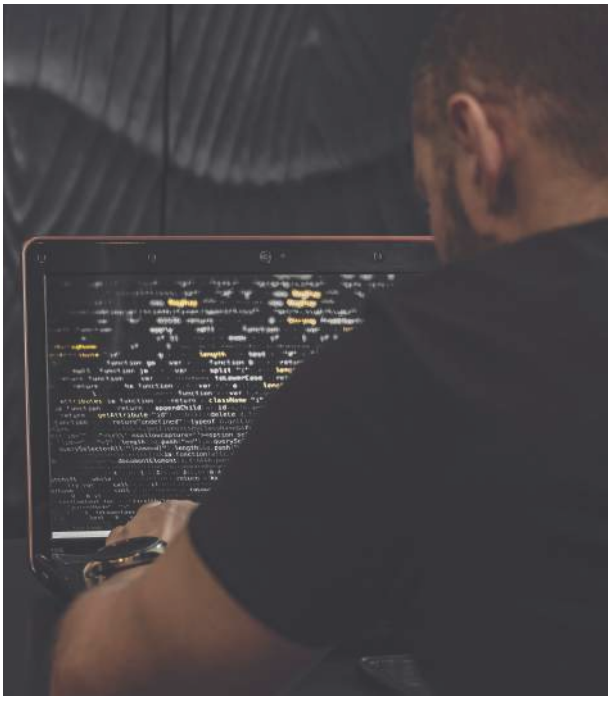
expedient removal of harmful content (Lyons, 2021). Jurisdictions throughout the Global South are at additional risk in this respect. The Facebook Files released by Frances Haugen through the Wall Street Journal in September 2021 suggest that the company has failed to establish effective terms and conditions of service, revise existing business models, and invest in the development of AI systems to filter local language hate speech and misinformation in developing countries, even when internal teams have flagged these as high-risk contents (Elliot et al., 2021). Facebook, however, is by no means unique among Global North corporations facing scrutiny for algorithms and practices seemingly harmful to Global South citizens and civil society.

Social media manipulation and digital surveillance tactics of governments and political parties are also to blame for undermining public discourse in digitally mediated forums (Neudert et al., 2019). A 2019 research study by the Oxford Internet Institute shows that politicians and political parties had deployed cyber propaganda, spreading manipulated media to amass fake followers and garner voter support in 45 democracies (Bradshaw et al., 2019). Also consider the case of the Israeli cyber-arms company NSO Group's Pegasus spyware deployed globally since at least 2011 to perform surveillance upon politicians, journalists, and activists, for a variety of motivations and with a broad range of harmful results (Marczak et al. 2018). Such cases reveal the broad vulnerability of digital systems and should inform how algorithms generally and AI platforms specifically might be abused by unchecked governments and nefarious actors alike.

Further, the abuse of AI surveillance technology is hardly confined to illiberal states. Carnegie's AI Global Surveillance Index (2019) that mapped 176 countries around the world found that 75 of the countries surveyed, including 51 percent of advanced democracies, were engaging in AI surveillance practices. The study showed that 56 countries had deployed smart city/safe city platforms, while 64 had rolled out facial recognition systems, and 52 had adopted smart policing practices (Feldstein, 2019). The deployment of facial recognition technology without safeguards by law enforcement agencies has emerged as a major bone of contention not just in the Global South – India (IFF, 2020), Uruguay (Datysoc, 2020), Brazil (Network Rights Coalition, 2019 & 2020), and South Africa (Lekabe, 2021) – but equally, in the North – the United States (US) (New America, 2021), the United Kingdom (UK) (Privacy International, 2021), the European Union (EU)[6]. Despite being a proponent of a 'trustworthy human-rights based approach' to AI governance, the EU has a wide berth for AI-based surveillance by law enforcement agencies (Vincent, 2021).



[6] See https://panoptic.in/central/FRT-000025; https://reclaimyourface.eu.



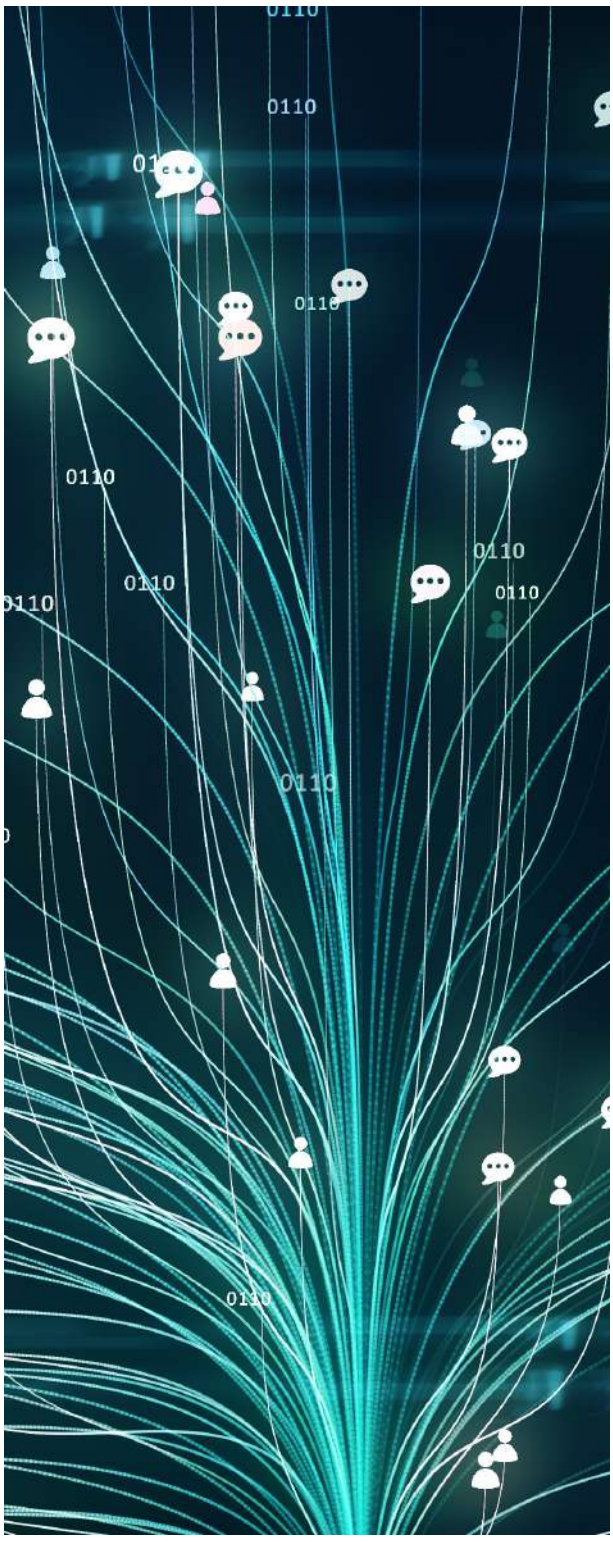
The US and the EU are guilty of what China is frequently criticised in international policy discourse – exporting AI surveillance technology that could threaten civic and political freedoms in other countries (Greco, 2021). A 2020 Privacy International study found that the EU has been directing aid funds to build mass-scale, high-risk biometric identity systems across the African continent to manage migration flows, without any data protection and human rights impact assessments (Privacy International, 2020). Foreign influence operations on social media are another threat, with social media companies having detected the presence of cyber troops engaged in such practices in at least seven countries: China, India, Iran, Pakistan, Russia, Saudi Arabia, and Venezuela (Bradshaw et al., 2019). The deployment of troll-farms and bots makes such propaganda warfare harder to trace and address (Barsotti, 2018).

Another emerging concern in both the Global North and South, as noted in the 2019 report of the UN Special Rapporteur on Poverty and Human Rights, is the *algorithmification* of the welfare state (Secretary-General, 2019). The algorithmic ranking and sorting of citizens to determine eligibility to access benefits and the digital-by-default welfare state has started to be rolled out without consideration for a rights-based approach and protection in this scenario: an upgrade of the Victorian poorhouse for the digital age, automatically sorting impoverished citizens into those 'deserving' and 'undeserving' of state largesse (Eubanks, 2018). This trend adds a layer of long-term vulnerability for consumers, which emanates from the need to create and maintain one or multiple online identities over a long period of time to access digital-by-default services (Kira et al. forthcoming). Citizens in the Global South are additionally disadvantaged as their governments' AI systems are frequently imported from the Global North and deployed without consideration of contextual factors (Secretary-General, 2019).[7]

[7] There are a few exceptions, such as India's domestic use and exportation of Aadhar to other countries.

The lack of a global agreement on social media governance has largely enabled the corporations that own the platforms to operate with impunity, particularly across the Global South. The Christchurch Call (Christchurch Call, 2019) on how online content should be moderated is perhaps the closest statement to any global consensus on the issue. However, the Christchurch Call is still not a multilateral agreement, lacking any legally binding obligations for digital companies (Pandey, 2020). A manipulated and weaponised cyberspace can erode substantive democracy, obscuring the collusion of the state and the market in the brazen disregard of human rights and the rule of law. A stalemate on an international covenant on cybersecurity (Clarke, 2021) also means that political sovereignty and national security interests are threatened in an international order where clandestine AI-enabled information warfare by foreign states is becoming the norm (Ördén et al., 2021). The adoption of AI in national welfare systems without appropriate tests for necessity, proportionality and legality may herald a crisis for citizenship rights with no recourse or remedy in international human rights benchmarks.

The status quo signals the inadequacy of current institutional frameworks to protect and nurture the democratic content of society through appropriate political mediation of the meaning, use and limits AI. The immediate task for AI governance thus centres on restoring the democratic integrity of the social order in the current conjuncture.



III. JUSTICE IN THE AI ECONOMY:
FAIR VALUE DISTRIBUTION & IMPLICATIONS FOR DEVELOPMENT

AI is to our digital epoch what electricity was to the industrial revolution: a paradigm-shifting general-purpose technology whose diffusion brings an exponential increase in productivity. This exponential increase derives from the augmentation of fixed capital and human capabilities in the production process, labour substitution, and product and service innovation (Bughin et al., 2018; Zuboff 2018). AI is estimated to add anywhere between USD 13 trillion and USD 15.7 trillion to global economic output by 2030 (Rao & Verweij, 2017). As the United Nations Conference on Trade and Development's (UNCTAD) 2021 Digital Economy Report observes, business models revolving around AI cannot exist without control over the data that feeds such models (UNCTAD, 2021). The generation of "intelligence premium" (Gurumurthy et al., 2019) is predicated on the ceaseless capture of social data.[8] This explains why the first-mover digital platforms from the US and China that control huge data enclosures are also leading investments and research in AI.[9]

If global AI adoption continues along the same trajectory, it might widen performance gaps, not just at the firm level and the individual worker level, but also the country level. Front-runner AI companies are likely to benefit disproportionately and may double their returns by 2030 while companies that delay adoption will be left far behind (Bughin et al., 2018). Similarly, at the worker level, demand for jobs and wages may grow for a few knowledge workers with digital and cognitive skills and with expertise in tasks that are hard to

[8] Mass digitization, which expanded with the Internet in the 1990s and escalated with data centres in the 2000s, has made available vast resources of data. A regime of knowledge extraction – built on Big Data – gradually employed efficient algorithms to extract 'intelligence' by capturing these open sources of data, mainly for the purpose of predicting consumer behaviour and selling ads. The knowledge economy has morphed into a novel form of capitalism in which unilateral control over data-based intelligence is the source of profit.

[9] As UNCTAD (2021) observes: Regarding mergers and acquisitions (M&A) of start-ups active in the AI segment, during the period of 1 January 2016–22 January 2021, there were 308 M&A deals worth $28.4 billion. The top five companies in the world, by number of acquired AI start-ups in the same period, were the Big Tech companies from the United States, followed by Baidu (sixth) and Tencent (eighth) from China. Apple led this ranking, followed by Google and Microsoft.

automate but will shrink for the majority performing repetitive and low digital skill jobs (Acemoglu et al., 2020). The US and China dominate the entire global AI economy: the two countries account for over 94 percent of all funding of AI start-ups in the past five years, 70 percent of the world's top AI researchers (UNCTAD, 2021) and 90 percent of the market capitalisation value of the world's 70 largest digital platform companies that control a significant proportion of cross-border data flows on the Internet (UNCTAD, 2019). American and Chinese participants are also better represented in the industry bodies developing standards, creating long-term dependency on basic technical protocols for the whole world.

With big data drawn from the Internet of Things becoming crucial, the EU, South Korea, and Japan, with their strong manufacturing base, associated computing power, and human resource capabilities, stand a very good chance of catching up (UNCTAD, 2021). The winners may well dominate the coming decades geo-economically and geo-politically (Feijóo, et al., 2020).

The acquisition of effective domestic AI capabilities depends upon three factors: big data, computing power, and the work of prominent AI researchers and engineers. Unfortunately, developing countries, disadvantaged both by the adverse terms of their integration into the Internet economy of user-generated data flows and limited industrial capacity to shift to smart manufacturing, are at high risk of being relegated in perpetuity to the low value parts of the AI economy. As currently configured, the AI race threatens to leave sub-Saharan Africa and most developing countries behind (UNCTAD, 2021), with an unprecedented concentration of wealth in the hands of a few companies in China and the United States. The competitive advantage in their 'cheap labour' that developing countries historically enjoyed may thus be rendered completely irrelevant (Lee, 2018).



On a fine-grained level, the concentration of data ownership in the hands of big technology multinationals feeds into local inequalities in countries of the Global South where they operate. This asymmetry in data ownership represents a barrier to entry for smaller homegrown start-ups and feeds into market concentration in contexts where local legislative infrastructure is weak and laws on competition and data protection, if present, are still nascent (Rizk, 2019). This exacerbates inequality and results in further exclusion for the less fortunate in countries of the Global South.



The lack of a globally accepted economic resource governance regime for data aids economic concentration and deepening of inequalities in the AI paradigm.[10] The rules for cross border data flows in the global economy are determined by a few powerful countries whose corporations enclose data from far and wide as trade secrets, (James, 2021) asserting de facto ownership rights over these holdings in perpetuity (Fia, 2021). In this intelligence economy, countries and communities of the Global South lacking in data processing and AI capabilities face a dangerous and untenable paradox. Not only must they relinquish any claims to their own data now locked up in AI systems of transnational capital, but they also have no means to legitimately derive a fair share of benefits generated therein. This results in gross economic unfairness in the global digital economy. Algorithmic coloniality is thus naturalised (Gurumurthy & Chami, 2021).

[10] That said, there may be other rights-based regimes we need to establish before we start institutionalising a regime for data as an economic resource.

IV. AI CONSTITUTIONALISM:
ADVANCING COOPERATION IN GOVERNING THE DATA ECONOMY

The governance deficit that marks the AI paradigm today has produced a crisis of democratic and distributive integrity. It cannot be mended without a holistic vision that puts people and the planet at the centre, something that is not without precedent. Evolving this AI governance roadmap requires that we eschew both techno-pessimism and techno-fundamentalism. Blanket bans and knee-jerk reactions to AI may not be the solution (Schwartz & Sheard, 2021; Paz, 2021) and disproportionate attention to a hypothetical technological singularity (Lacker, 2021) can distract from the crucial narrative of power in the AI governance debate. At the same time, a business-as-usual optimism will not be sustainable.

What we have seen with the steady erosion of the civic-public space in AI-mediated sociality and the unjust order of the AI-enabled economy that is inimical to a peaceful and just future. As the UN Special Rapporteur on Freedom of Expression (2018) has highlighted, the



universal human rights framework can provide a robust starting point for evolving an effective AI governance approach. The needed shift, however, cannot just be a nominal rearticulation of global constitutionalism for the AI age. A global AI constitutionalism[11] must carry the aspirations of multiple communities, privileging a future society where manipulation, loss of autonomy, exploitation and injustice have no place. It must be based on a rebalanced multilateralism for a renewed intelligence paradigm that sets a high standard for state obligations towards freedoms.

[11] By 'global AI constitutionalism' we refer to a yet-to-be formulated set of universal principles that can serve as a reference point for AI regulatory policy and legislation across national contexts. In this context, the UNCTAD Digital Economy Report (2021) underscores the need for a new global governance framework for data flows that will determine "who has access to data, under which conditions and for which use" (UNCTAD, 2021).

Humanity faces an emergency in the abuse of AI by unaccountable power – the weaponization of dual use AI technologies, attacks on the sovereignty of state parties, and an unchecked aggrandisement of corporate power, particularly within the Global South. States must urgently embrace their duty to protect the rights of their citizens and uphold their extraterritorial obligations towards the human rights of all peoples. Beyond this, however, what is at stake is more than a state commitment. It is a willingness to act now to set the right foundations, build consensus on a framework for upholding rights in the age of AI – with adjustments in place to fill the gaps in international human rights law– and develop capacity to implement remedies.

Committing to a global AI constitutionalism is not the same as calling for a universal formula to address bias, discrimination, and inequality in the design of AI systems. A meaningful interpretation of the moral values that human rights represent in different contexts still needs to be adopted. For instance, privacy concerns emerge in unique ways across the Global South stemming from complex and situated notions straddling autonomy over personal information, the realities of shared use of digital artefacts, and communal identity (Ahmed et al., 2017).



CONCLUSION

With AI's decisive role in determining pathways to economic prosperity and development, and an evident and growing AI divide, particularly in the wake of the Covid-19 pandemic, there is an urgent need to redefine the AI governance debate. There has been much discussion on the 'AI divide' between the Global North and South. The appeals for appropriate application of AI for social transformation and the achievement of Sustainable Development Goals (SDGs), thus hinges on the prospect of tackling global inequality in the AI paradigm.

We argue that a global governance framework for AI cannot sidestep the political economy of data and AI's emerging role in the world economic order. 'AI for good' should be about distributing the gains of the AI paradigm equitably to encourage locally embedded development innovations, not merely the unidirectional sharing of knowledge and innovation from the Global North to the South. In the context of international development, the ownership and distribution of AI-powered technologies and relevant data must not be allowed to generate loops of dependency between Southern countries and the now dominant powers in the space: China, the US, and, to some degree, the EU. The terms of AI's distribution and benefits must be favourable to the Global South. By highlighting the convergences and divergences that mark the political field in the AI governance debate, we aim to draw attention to elements of a new agenda for progressive actors in civil society – activists, scholars, technologists – grounded in the notion of 'AI for fair value'.



The objective of this paper, however, is not to articulate specific policy proposals for adoption at an international level, but rather to draw attention to the problem of AI policy fragmentation and to call for well-informed efforts toward remedying it through international dialogue. That said, the working group recommends several key actions to work toward a less fragmented AI policy landscape that incorporates the interests and concerns of Global South countries. These include:

Working Group recommendations for a less fragmented AI policy landscape, incorporating Global South perspectives:

- ☑ Sustaining a running global discussion of digital rights concerns and standards in AI policy that prioritises civil society, expert, and policy maker representation from the Global South, with the goal of arriving at concrete regulatory principles that can be adopted at the highest international level. To be effective, this dialogue must engage policy makers and corporate leadership, but civil society's role must be robust and representative to drive that discussion and achieve tangible policy outcomes.

- ☑ Building a global database of proposed and implemented AI legislation, focusing especially upon the nations and subnational units of the Global South. Such a database could also be used to monitor potentially problematic applications, policy blind spots, and loopholes around the world, facilitating the establishment of worldwide norms in AI policy. The OECD's Artificial Intelligence Policy Observatory offers a vital and laudable start to that end.[12] Still, a more dedicated Global South focus would ensure that this effort is comprehensive, active, and broadly-embraced across Southern civil society. This would represent something of a first alert system for abuses and a clearinghouse for best practices with direct local input.

- ☑ Global South countries should consider implementing the requirement of AI impact assessment reports, before the AI-based services are made publicly available by private – and especially extranational – corporations.

- ☑ Create incentives to retain young AI scientists and engineers in the Global South. One suggestion would be to incentivise global tech companies to create research and development facilities in the South. That initiative would also contribute to create diversity in AI development teams.

[12] See https://oecd.ai/en/

What is especially needed is a clear and unanimous rejection of anti-citizen abuses of AI technologies, algorithmic systems, and software. The prospect of AI being harnessed to similarly infringe upon citizens' rights, privacy, and open public discourse around the world is a dangerous one. Particularly in the context of outcomes for Global South countries, these critical cybersecurity issues involving AI must be addressed at the highest level, with the goal of codifying protections through conventions, protocols, and other binding processes via existing intergovernmental organisations and international agencies, such as the UN.

Ungoverned, AI holds the potential to exacerbate socio-economic inequality, erode civil rights, and undermine peace and economic development around the world. When developed, monitored, and applied in cooperative ways that correspond with widely-held standards of human rights, AI holds the potential to stimulate sustainable economic growth and reduce inequalities, promoting a durable peace for the whole world – South and North.



BIBLIOGRAPHY

Acemoglu, D., & Restrepo, P. (2020). The Wrong Kind of AI? Artificial Intelligence and the Future of Labour Demand. *Cambridge Journal of Regions, Economy and Society*, *13* (1), 25–35.

Ahmed, S. I., Haque, M., Guha, S., Rifat, M. R., & Dell, N. (2017). Privacy, Security, and Surveillance in the Global South: A Study of Biometric Mobile SIM Registration in Bangladesh. *the 2017 CHI Conference*.

Aizenberg, E., & van den Hoven, J. (2020, July). Designing for human rights in AI. *Big Data & Society*, https://doi.org/10.1177/2053951720949566.

Allmann, K. & Radu, R. (forthcoming). Missteps on the path to digitizing welfare: 'digital footprints' as access currency for e-government services. *Global Policy*.

Barsotti, S. (2018). *Weaponizing Social Media: Heinz Experts On Troll Farms And Fake News*. Retrieved from Heinz College of Information Systems and Public Policy: https://www.heinz.cmu.edu/media/2018/October/troll-farms-and-fake-news-social-media-weaponization

Bissio, R. (2018). *Spotlight on Sustainable Development 2018*. Civil Society Report, IT for Change.

Bradshaw, S., & Howard, P. N. (2019). *2019 Global Inventory of Organised Social Media Manipulation*. The Computational Propaganda Project at the Oxford Internet Institute.

Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). *Notes From the AI Frontier: Modeling The Impact Of AI On The World Economy*. Discussion Paper, McKinsey Global Institute.

Bughin, J., Seong, J., Manyika, J., Hämäläinen, L., Windhagen, E., & Hazan, E. (2019). *Notes From the AI Frontier: Tackling Europe's Gap in Digital and AI*. Discussion Paper, McKinsey Global Institute.

Cave, S., & Kanta, D. (2020). The Whiteness of AI. *Philosophy & Technology*, *33*, 685-703.

Christchurch Call. (2019). *Christchurch Call*. Retrieved from https://www.christchurchcall.com/call.html

Clarke, L. (2021, April 22). *UN countries agreed to a more peaceful cyberspace – but it won't stop state-sponsored attacks*. Retrieved from TechMonitor: https://techmonitor.ai/policy/geopolitics/un-countries-cybersecurity-deal-state-sponsored-attacks

Dasgupta , B. (2021, October 23). *Twitter says its algorithm amplifies right wing political content*. Retrieved from Hindustan Times: https://www.hindustantimes.com/india-news/twitter-says-its-algorithm-amplifies-right-wing-political-content-101634926182240.html

Datysoc. (2020, November 17). *Organizaciones de la sociedad civil y académicas expresan su preocupación por reconocimiento facial en el Proyecto de Ley de Presupuesto de Uruguay*. Retrieved from Datysoc: https://datysoc.org/2020/11/17/organizaciones-de-la-sociedad-civil-y-academicas-expresan-su-preocupacion-por-reconocimiento-facial-en-el-proyecto-de-ley-de-presupuesto-de-uruguay/

Dehingia, N., Lundgren, R., Dey, A., & Raj, A. (2021). *Trends in online misogyny before and during the COVID-19 pandemic: Analysis of Twitter data from five South-Asian countries*. Center on Gender Equity and Health, University of California San Diego.

Elliot, V., Christopher, N., Deck, A., & Schwartz, L. (2021, October 26). *The Facebook Papers reveal staggering failures in the Global South*. Retrieved from Rest Of World: https://restofworld.org/2021/facebook-papers-reveal-staggering-failures-in-global-south/

Eubanks, V. (2018). *The Digital Poorhouse*. Retrieved from Harper's Magazine: https://harpers.org/archive/2018/01/the-digital-poorhouse/

Feijóo, C., Kwon, Y., Bauer, J., Bohlin, E., Howell, B., Jain, R., . . . Xia, J. (2020). Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy. *Telecommunications Policy*, *44* (6).

Feldstein, S. (2019). *The Global Expansion of AI Surveillance*. Working Paper, Carnegie Endowment for International Peace.

Fia, T. (2021). An Alternative to Data Ownership: Managing Access to Non-Personal Data through the Commons. *Global Jurist*, *21* (1), 181-210. https://doi.org/10.1515/gj-2020-0034.

Fukuda-Parr, S., & Gibbons, E. (2021). Emerging Consensus on 'Ethical AI': Human Rights Critique of Stakeholder Guidelines. *Global Policy*, *12*, 32-44. https://doi.org/10.1111/1758-5899.12965.

Greco, M. (2021, October 7). *What Biden Needs to Say About Surveillance Tech and Foreign Policy*. Retrieved from Just Security: https://www.justsecurity.org/78467/what-biden-needs-to-say-about-surveillance-tech-and-foreign-policy/

Gurumurthy, A., & Chami, N. (2019). *The Wicked Problem of AI Governance*. Friedrich-Ebert Stiftung India Office. Artificial Intelligence in India.

Gurumurthy, A., & Chami, N. (2021, May 3). Towards a Global Digital Constitutionalism: A Radical New Agenda for UN75. *Development*, *64*, 29-38.

Gurumurthy, A., & Chami, N. (2021). *Governing the Abstract Object of Data: Towards a Distributive Integrity Framework*. IT for Change.

Gurumurthy, A., Bharthur, D., Chami , N., Vipra, J., & Anwar, I. A. (2019). *Platform Planet: Development in the Intelligence Economy*. IT for Change.

Haas, L., & Gießler, S. (2020, April 28). Retrieved from Algorithm Watch: https://algorithmwatch.org/en/ai-ethics-guidelines-inventory-upgrade-2020/

IFF. (2020). *Introduction to Facial Recognition Projects in India*. Retrieved from Internet Freedom Foundation: https://internetfreedom.in/facial-recognition-in-india-part-i/

James, W. (2021, October 22). *G7 countries reach breakthrough on digital trade and data*. Retrieved from Reuters: https://www.reuters.com/world/g7-countries-agree-shared-position-digital-trade-communique-2021-10-22/

Lacker, K. (2021, October 25). *Thoughts On AI Risk*. Retrieved from Kevin Lacker's blog: https://lacker.io/ai/2021/10/25/thoughts-on-ai-risk.html

Lee, K.-F. (2018). *AI superpowers: China, Silicon Valley, and the new world order*. Houghton Mifflin Harcourt.

Lekabe, T. (2021, April 23). *AfriForum wants assurance on Sanral facial recognition tech*. Retrieved from TheCitizen: https://www.citizen.co.za/news/south-africa/government/2478929/afriforum-wants-assurance-on-sanral-facial-recognition-tech/

Lyons, K. (2021, October 17). *Facebook disputes report that its AI can't detect hate speech or violence consistently*. Retrieved from The Verge: https://www.theverge.com/2021/10/17/22731214/facebook-disputes-report-artificial-intelligence-hate-speech-violence

Mhlambi, S. (2020). *From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance*. Carr Center for Human Rights Policy.

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*, *33*, 659-684. https://doi.org/10.1007/s13347-020-00405-8.

Muchiri, C. W. (2020). *Impact of Artificial Intelligence on IP Policy: Call for Comments- Response from a Global South Perspective*. Strathmore Law School.

Network Rights Coalition. (2019, November 27). *IGF 2019: Open Letter*. Retrieved from Coalizão Direitos na Rede: https://direitosnarede.org.br/2019/11/27/igf-2019-open-letter/

Network Rights Coalition. (2020, November 17). *Open Letter from Brazilian Civil Society on the occasion of the 15th edition of the United Nations Internet Governance Forum*. Retrieved from Coalizão Direitos na Rede: https://direitosnarede.org.br/2020/11/17/open-letter-from-brazilian-civil-society-on-the-occasion-of-the-15th-edition-of-the-united-nations-internet-governance-forum/

Neudert, L.-M., & Howard, P. (2019, November 18). *Online politics needs to be cleaned up – but not just by Facebook and Twitter, say Oxford academics*. Retrieved from Oxford Internet Institute: https://www.oii.ox.ac.uk/news-events/news/online-politics-needs-to-be-cleaned-up-but-not-just-by-facebook-and-twitter-say-oxford-academics/

New America. (2021, June 3). *Civil Rights Concerns Regarding Law Enforcement Use of Face Recognition Technology*. Retrieved from New America: https://www.newamerica.org/oti/briefs/civil-rights-concerns-regarding-law-enforcement-use-of-face-recognition-technology/

Office of the United Nations High Commissioner for Human Rights. (2021, March 23). *Report: Online hate increasing against minorities, says expert*.

Ördén, H., & Pamment, J. (2021). *What Is So Foreign About Foreign Influence Operations?* Carnegie Endowment for International Peace.

Pandey, P. (2020, August 3). One year since the Christchurch Call to Action: A Review. *ORF Issue Brief*(389).

Paz, A. W. (2021, July 7). *We need to regulate mind-reading tech before it exists*. Retrieved from Rest Of World: https://restofworld.org/2021/chile-neuro-rights/

Privacy International. (2020, November 10). *Here's how a well-connected security company is quietly building mass biometric databases in West Africa with EU aid funds*. (International Privacy) Retrieved from Privacy International: https://privacyinternational.org/news-analysis/4290/heres-how-well-connected-security-company-quietly-building-mass-biometric

Privacy International. (2021, August). *Civil Society Groups: Live Facial Recognition Technology should not be used in public spaces*. Retrieved from Privacy International: https://www.privacyinternational.org/sites/default/files/2021-08/LFRT%20Open%20Letter%20Final.pdf

Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective. *Policy and Society*, *40* (2), 178-193.

Rao, A., & Verweij, G. (2017). *Sizing the prize What's the real value of AI for your business and how can you capitalise?* PWC.

Report of the Secretary-General. (2020). *Roadmap for Digital Cooperation*. United Nations.

Risse, M. (2021). *Artificial Intelligence and the Past, Present, and Future of Democracy*. Discussion Paper, Carr Center for Human Rights Policy.

Rizk, N. (2019). Artificial Intelligence and Inequality in the Middle East: The Political Economy of Inclusion. *The Oxford Handbook of Ethics of AI*.

Schwartz, A., & Sheard, N. (2021, January 20). *Why EFF Doesn't Support Bans On Private Use of Face Recognition*. Retrieved from Electronic Frontier Foundation: https://www.eff.org/deeplinks/2021/01/why-eff-doesnt-support-bans-private-use-face-recognition

Marczak, B., Scott-Railton, J., McKune, S., Abdul Razzak, B., & Deibert, R. (2018). Hide and seek: tracking NSO Group's Pegasus spyware to operations in 45 countries. *Citizen Lab Research Report*. No. 113, University of Toronto.

Secretary-General. (2019). *Report of the Special rapporteur on extreme poverty and human rights*. Advance Unedited Version.

Ulnicane, I., Knight, W., Leach, T., Stahl, B. C., & Wanjiku, W.-G. (2021). Framing governance for a contested emerging technology:insights from AI policy. *Policy and Society, 40* (2), 158-177.

UNCTAD. (2021). *Digital Economy Report 2021*.

UNCTAD. (2019). *Digital Economy Report 2019*.

UNCTAD. (2021). *Technology and Innovation Report*.

Vincent, J. (2021, April 21). *EU outlines wide-ranging AI regulation, but leaves the door open for police surveillance*. Retrieved from The Verge: https://www.theverge.com/2021/4/21/22393785/eu-ai-regulation-proposal-social-credit-ban-biometric-surveillance-exceptions

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.