GLOBAL INFORMATION SOCIETY WATCH 2019

Artificial intelligence:
Human rights, social justice and development

Association for Progressive Communications (APC),
Article 19, and Swedish International Development Cooperation Agency (Sida)

Global Information Society Watch

2019

Global Information Society Watch 2019
Artificial intelligence: Human rights, social justice and development

Operational team
Valeria Betancourt (APC)
Alan Finlay (APC)
Mallory Knodel (ARTICLE 19)
Vidushi Marda (ARTICLE 19)
Maja Romano (APC)

Project coordination team
Valeria Betancourt (APC)
Cathy Chen (APC)
Flavia Fascendini (APC)
Alan Finlay (APC)
Mallory Knodel (ARTICLE 19)
Vidushi Marda (ARTICLE 19)
Leila Nachawati (APC)
Lori Nordstrom (APC)
Maja Romano (APC)

GISWatch 2019 advisory committee
Namita Aavriti (APC)
Rasha Abdul Rahim (Amnesty International)
Alex Comninos (Research ICT Africa)
Malavika Jayaram (Digital Asia Hub)
J. Carlos Lara (Derechos Digitales - América Latina)
Joy Liddicoat (Centre for Law and Emerging Technologies, University of Otago)
Andrew Lowenthal (EngageMedia)
Micaela Mantegna (Geekylegal/Machine Intelligence Lab, Center for Technology and Society, San Andres University)
Valeria Milanes (Asociación por los Derechos Civiles)

Project coordinator
Maja Romano (APC)

Editor
Alan Finlay (APC)

Assistant editor and proofreading
Lori Nordstrom (APC)

Publication production support
Cathy Chen (APC)

Graphic design
Monocromo

Cover illustration
Matías Bervejillo

We would like to extend a special note of thanks to a number of authors who have made ad honorem contributions to this edition of GISWatch.
We gratefully acknowledge the following:

Philip Dawson and Grace Abuhamad (Element AI)
Anita Gurumurthy and Nandini Chami (IT for Change)
Rasha Abdul Rahim (Amnesty International)

APC would like to thank the Swedish International Development Cooperation Agency (Sida) and ARTICLE 19
for their support for Global Information Society Watch 2019.

Published by APC
2019
Printed in USA

Global Information Society Watch 2019 web and e-book
ISBN 978-92-95113-13-8
APC Serial: APC-201910-CIPP-R-EN-DIGITAL-302

Radicalising the AI governance agenda[1]

Anita Gurumurthy and Nandini Chami
IT for Change
www.ITforChange.net

What's missing in mainstream global debates on AI governance

Advances in artificial intelligence (AI) present human civilisation with challenges that are unprecedented. As a class of technologies[2] that simulate human intelligence processes for learning, reasoning and self-correction, AI disrupts the way societies define, organise and use knowledge, thus radically recasting social and economic systems. Understanding and deconstructing AI systems that are self-learning and self-correcting is not easy. In fact, experts in the field have even stated that it is impossible. The widespread diffusion and adoption of AI, even if much of it for now is so-called "narrow AI",[3] is therefore as terrifying as it is exciting – something that Bill Gates has compared to the complexity of nuclear technology. Quite naturally, a vibrant debate on the governance of AI has been gathering momentum, involving governments, multilateral institutions, technology companies, the technical community and global civil society. The search is on for the right combination of legal-regulatory, ethical and technological approaches that constitute effective AI governance.

Mainstream debates on AI governance take note of violations of the human rights considerations of privacy, equality and non-discrimination, uncertain futures of work, and erosion of democracy in the emerging AI paradigm. They do not, however, fully address the entanglement of AI in neoliberal capitalism and what this means for the life-chances of individuals and communities. Because of this, AI governance debates tend to carry critical blind spots.

Blind spot 1: Collective autonomy and choice in the debate on AI and human rights

Across stakeholders, there is growing acknowledgement of how AI systems could undermine human rights. A systematic mapping of the over 32 sets of influential AI principles/guidelines in existence today by the Cyber Harvard project reveals that informational privacy, equality, fairness and freedom from discrimination are critical concerns shared by all stakeholders involved in the development and deployment of AI technologies: governments, multilateral organisations, advocacy groups and technology companies.[4] The inscrutability of AI systems means that the subjectivity of their creators can reinforce the very biases that create an unequal society, leading to a due process failure. Inherent biases in input/training data sets as well as in definitions of output parameters produce unfair outcomes.

Institutional and techno-governance mechanisms to address bias in AI are indeed necessary to tackle inequality and discrimination. However, existing proposals in this regard, whether from multilateral agencies (such as the global legal framework mooted by the UN Special Rapporteur on freedom of expression in his 2018 report),[5] or plurilateral bodies (the OECD Council's Recommendation on Artificial Intelligence),[6] or governments (the European Commission's Ethics Guidelines for Trustworthy AI),[7] or civil society (the Toronto Declaration[8] for protecting equality and non-discrimination in AI systems), or the technical community (such as IEEE's project on evolving an open standard on algorithmic bias), tend to focus exclusively on addressing misrecognition.

1 This report has been adapted from "The Wicked Problem of AI Governance", which will be published by FES-India in October 2019.

2 Ranging from computer vision, natural language processing, virtual assistants and robotic process automation to advanced machine learning. See: Bowles, J. (2018, 18 September). McKinsey warns that AI will further divide the world economy into winners and losers. *Diginomica*. https://diginomica.com/mckinsey-warns-that-ai-will-further-divide-the-world-economy-into-winners-and-losers

3 AI used for a narrowly defined task, as opposed to the more complex general or strong AI.

4 Fjeld, J., et al. (2019, 4 July). Principled Artificial Intelligence: A Map of Ethical and Rights-Based Approaches. *Berkman Klein Center for Internet & Society*. https://ai-hr.cyber.harvard.edu/primp-viz.html

5 https://undocs.org/A/73/348

6 https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

7 https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

8 https://www.accessnow.org/the-toronto-declaration-protecting-the-rights-to-equality-and-non-discrimination-in-machine-learning-systems

They fail to imagine redress to individuals and communities caught in relationships of exploitation that are based on uneven and unfair distribution of intelligence capital. In the AI-led economy, algorithmic intelligence extracted from data resources is the "secret sauce"[9] that enables the disruption of the economic status quo and the attainment of new levels of efficiency. At present, such "intelligence capital" is concentrated in the hands of a few transnational corporations, which have enclosed valuable data resources in order to cement their market dominance by foreclosing the possibility of competing AI innovations emerging in the future.

Because of their failure to address the unequal distribution of intelligence capital and the resultant inequality in opportunity structures, existing AI and human rights proposals ignore the changing structures of choice. We urgently need framings about equality and non-discrimination in relation to AI that are attentive to "equality of autonomy"[10] – the spread across society of the ability and means of people to choose their life course. Our response to safeguarding human rights in the AI paradigm must move beyond identity-based discrimination, and tackle AI-based economic exploitation through new governance approaches for the AI economy that expand individual and collective choices.

Blind spot 2: Economic self-determination in the debate on AI

In the race towards the "Fourth Industrial Revolution", an ideology of AI-frontierism is widely evidenced in policy circles. Not wanting to be left behind, developing country governments are caught up in the language of "innovation" and "entrepreneurship", authoring national plans and road maps for their digital start-up ecosystem and upskilling of workers. These efforts view AI-led development as a simplistic aggregate of individual efficiencies that will somehow magically add up to national productivity gains. They completely ignore the fact that development is a "competitive and global undertaking", characterised by a sustained and continuing effort to capture opportunities for higher value knowledge and technological capabilities.[11] In the current context, strides in development are possible only for countries that can harness AI at a socio-structural level for higher growth and redistributive gains. Developing countries urgently need to use AI to create and/or deepen national capacity for moving out of low value locations in the global value chain. However, the debate so far[12] seems to flatten the global political economy of development with a broad brush stroke, and even glib prescriptions exhorting countries of the South to build their domestic AI capabilities and upskill their populations.

How can these prescriptions be met if access to and ownership of data and digital intelligence are denied to these countries? The AI-led global order is entrenched firmly in what activists and scholars have argued is a form of neocolonisation.[13] Today, economic power is a function of how AI technologies are employed in networked systems organised around incessant data processing. As data started flowing on a planetary scale with the advent of the internet, creating and multiplying social and economic connections, predatory capitalism found a new lease of life. The value of the global network of connections has since grown exponentially with the emergence of the platform model, the network-data infrastructures that mediate and organise production and exchange on a global scale. In the emerging global AI economy, competitive advantage is determined by the ability to reach higher levels of efficiency through the intelligence capital generated by processing data.

Moving to the higher value segments of the global economy is, however, inordinately difficult in the current global economic order, where corporations and countries who have enjoyed a first-mover advantage in harvesting data for digital intelligence systematically reinforce their position of dominance. As the United Nations Conference on Trade and Development (UNCTAD) Trade and Development Report[14] cautions, the restructuring of global value chains by the platform business model has

9 Morozov, E. (2018, 28 January). Will tech giants move on from the internet, now we've all been harvested? *The Guardian*. https://www.theguardian.com/technology/2018/jan/28/morozov-artificial-intelligence-data-technology-online

10 Sen, A. (2001). *Development as Freedom*. Oxford University Press.

11 Mann, L., & Iazzolino, G. (2019). *See, Nudge, Control and Profit: Digital Platforms as Privatized Epistemic Infrastructures*. IT for Change. https://itforchange.net/platformpolitics/wp-content/uploads/2019/03/Digital-Platforms-as-Privatized-Epistemic-Infrastructures-_5thMarch.pdf

12 Smith, M., & Neupane, S. (2018). *Artificial Intelligence and Human Development: Toward a Research Agenda*. International Development Research Centre. https://idl-bnc-idrc.dspacedirect.org/handle/10625/56949 and World Economic Forum. (2017). *Accelerating Workforce Reskilling for the Fourth Industrial Revolution: An Agenda for Leaders to Shape the Future of Education, Gender and Work*. www3.weforum.org/docs/WEF_EGW_White_Paper_Reskilling.pdf

13 Avila, R. (2018). *Resisting Digital Colonialism*. Mozilla. https://internethealthreport.org/2018/resisting-digital-colonialism and Couldry, N. & Mejias, U. (2018). *Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject*. LSE Research Online. https://eprints.lse.ac.uk/89511/1/Couldry_Data-colonialism_Accepted.pdf

14 UNCTAD. (2018). *Trade and Development Report 2018: Power, Platforms and the Free Trade Delusion*. https://unctad.org/en/PublicationsLibrary/tdr2018_en.pdf

coincided with the appearance in global economic statistics of a "widening gap between a small number of big winners in global value chains and a large collection of participants, both smaller companies and workers, who are being squeezed."[15]

The United States (US) and its allies have also sought to use trade negotiations to assert their advantage and maintain the status quo on unrestricted cross-border data flows to protect US platform monopolies. Similarly, they have been stalling demands of developing countries for disclosure of source code/algorithms by transnational digital corporations, even though such technology transfer conditionalities for market access are currently permissible under the Agreement on Trade Related Investment Measures (TRIMs). Without the sovereign right to control the terms on which the data of their citizens or the data generated in their territories flows across jurisdictions and/or the means to build the digital intelligence capabilities to boost their economies, countries in the developing world cannot create the endogenous conditions for their citizens to reap the AI advantage. They will never be able to create the intelligence capital for reaching higher value knowledge capabilities. On the contrary, their vulnerabilities could potentially be accentuated, as the systematic flight of data from their territories for exogenous AI infrastructure models creates economic and political dependencies.

The terms of the debate therefore need to shift away from individualist solutions to secure the future of the economy towards governance frameworks that invoke the economic right of nation states and communities to have sovereignty over data – which may be seen as "a new form of wealth"[16] – to self-determine their development pathways.

Blind spot 3: The realpolitik of algorithmic scrutiny in the debate on norms for digitally mediated democracy

The early consensus on internet exceptionalism linked to free speech seems to be giving way to a realisation that a hyper-extractive algorithmic regime needs new norms that can hold platform intermediaries accountable for preserving democracy in digitally mediated times. There is thus an increasing acknowledgement about the need for public scrutiny of the algorithmic tools used by platforms for content curation, user profiling and targeting.[17]

In the past year, the European Union (EU) has been at the helm of this debate, with members of the European Parliament calling for an algorithmic audit of the profiling practices of Facebook in October 2018 and the establishment of an EU Committee of Ministers to deliberate on safeguards against algorithmic manipulation by platforms, including digital communication services.[18] While the EU – as a politically powerful and economically relevant bloc – may well be able to create the regulatory structures and enforce accountability mechanisms *vis-à-vis* transnational platform companies within its territory, most countries in the global South lack such clout and the institutional wherewithal for regulatory oversight. As mentioned, the US and its allies have also sought to protect the intellectual property interests of their digital corporations in trade-related negotiations, insisting that no country can make market access contingent on source code/algorithmic disclosure.[19] Most developing countries therefore face a Hobson's choice: they must give in to opaque and unilateral AI-enabled content governance policies and practices of transnational platform companies in order to have access to the essential communications infrastructure that they depend on the latter to provision.

These geo-economic and geo-political dynamics as well as the absence of a binding international framework on the obligations of transnational corporations render the plausibility of effective regulatory intervention by developing countries moot. Ideas of self-regulation tend to gain currency, furthering a user-centred approach that depoliticises the problem, replacing democratic oversight with corporate largesse.

A two-pronged response is necessary to prevent the degeneration of the digitally mediated public sphere. Firstly, the deleterious consequences of "AI-gone-wrong" for democracy cannot be tackled without a right for all countries to scrutinise the algorithmic apparatus shaping social interactions in

15 Ibid.

16 PTI. (2019, 28 June). Data 'new form of wealth', needs to take into account developing nations' needs: India. *New Indian Express*. www.newindianexpress.com/world/2019/jun/28/data-new-form-of-wealth-needs-to-take-into-account-developing-nations-needs-india-1996614.html

17 Garton Ash, T., Gorwa, R., & Metaxa, D. (2019). *GLASNOST! Nine ways Facebook can make itself a better forum for free speech and democracy*. Reuters Institute for the Study of Journalism and University of Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-01/Garton_Ash_et_al_Facebook_report_FINAL_0.pdf

18 Koene, A., et al. (2019). *A Governance Framework for Algorithmic Accountability and Transparency*. European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf

19 Ibid.

their territory. The proposed international treaty on business and human rights is a highly pertinent instrument[20] through which corporate violations that undercut democracy and human rights can be addressed by governments. Additionally, the health of public spheres in digital times hinges on a global agreement, a binding normative framework on data and AI that prescribes duties of states *vis-à-vis* national and global democracy. A reinterpretation of human rights obligations of state and non-state actors in the age of AI, therefore, is not optional: it is an urgent need. A global normative framework for data and AI must also address the issue of data extractivism, setting limits on individual profiling in the online communications sphere.

A radical agenda for AI governance: Building blocks

Violations of the foundational human rights principle of equality and non-discrimination and the thwarting of political and economic democracy in the AI paradigm are, evidently, a result of data imperialism – the control that algorithmic circuits of digital intelligence confer on the already powerful who own the data. Surprisingly though, this facet of AI is hardly alluded to in the debates on AI governance, which – as demonstrated above – propose liberalist, structural interventions (focusing on correcting misrecognition but not maldistribution) at best and neoliberal, individualistic fixes (that transfer burdens of navigating the digital economy on individuals) at worst. When viewed from this standpoint, the contours of the AI governance debate shift significantly. It becomes apparent that transforming the political economy of data ownership and control that is deepening global development fault lines is the critical missing link. The AI governance agenda therefore needs to be transformed and radicalised, embracing a focus on data and AI constitutionalism.

Two critical steps need to be accomplished for such a radical departure:

(a) Acknowledging data sovereignty as part of the right to development

In the AI paradigm, without a national-level strategy to leverage data resources for inclusive innovation and social transformation, the right and duty of nation states to formulate appropriate national development policies as envisaged in the Declaration on the Right to Development cannot be realised. For example, in order to safeguard strategic economic interests, countries must be able to build and strengthen public data pools, mandating private firms to relinquish their exclusive rights over data collected and processed as part of their business where such data is assessed to be of national importance. They must also be able to prevent the enclosure and expropriation of cultural/knowledge commons or community data by transnational digital companies. But in a context where the bulk of data resources of developing countries are in the hands of transnational digital companies headquartered elsewhere, such national-level policy measures can be enforced only by re-asserting jurisdictional sovereignty over data resources through the introduction of restrictions and controls on cross-border data transfers, and data localisation measures. It is this policy space that is currently being taken away by advanced AI nations who are utilising trade policy avenues to push for the maintenance of the status quo on unrestricted data flows and protect the interests of their corporations. Such tactics also promote a myth that any national-level conditionalities on data flows are likely to impede global flows of information on the internet.

The sovereign right of nation states to the data on their citizens or collected within their territories needs to be articulated through a binding global normative framework on data and AI. Norms about putting AI to the service of human rights and development justice must embrace the cutting-edge wisdom about the inalienability, indivisibility and interdependence of human rights, with a futuristic outlook for the 21st century. To fulfil their human rights obligations in the AI paradigm, states need to implement various measures, balancing multiple interests and priorities in the national context. A sophisticated governance framework for access to and use and control of data is needed that effectively balances the rights of data principals with the rights of those investing in the resources that enable the creation of digital intelligence, the rights of affected individuals/communities, and the broader public interest.[21]

20 For more details, see Zorob, M. (2019, 30 September). The Lengthy Journey towards a Treaty on Business & Human Rights. *Business & Human Rights Resource Centre*. https://www.business-humanrights.org/en/the-lengthy-journey-towards-a-treaty-on-business-human-rights

21 British Academy, Royal Society, & techUK. (2018). *Data Ownership, Rights and Controls: Reaching a Common Understanding*. https://royalsociety.org/-/media/policy/projects/data-governance/data-ownership-rights-and-controls-October-2018.pdf and Scassa, T. (2018). *Data Ownership*. Centre for International Governance Innovation. https://www.cigionline.org/sites/default/files/documents/Paper%20no.187_2.pdf

(b) Reining in transnational digital corporations

Given that the bulk of AI innovation is currently being spearheaded by transnational corporations, norms and rules at the national level are necessary to protect the interests of domestic businesses and enterprises (across a wide spectrum that includes not-for-profits and cooperatives). Policy measures will need to straddle: FRAND (Fair, Reasonable and Non-Discriminatory Access) provisions in technology patenting to prevent digital corporations from locking in essential building blocks of algorithmic innovation;[22] foreign direct investment controls in the digital start-up sector to prevent extractivist investments that cannibalise domestic enterprises;[23] regulation for algorithmic audit and scrutiny to protect the rights to privacy, equality and non-discrimination; and limits on the use of personally identifiable data for hyper-profiling. But the rapacious greed of digital transnational corporations for data, their opacity about algorithms and brazen non-compliance with domestic regulation are issues that require an international mechanism to enforce corporate accountability. Although some progress has been made in deliberating a legally binding instrument on transnational corporations and business enterprises with respect to human rights, this process has not gathered momentum owing to the clout that transnational corporations enjoy. The need for progress on this front cannot be overemphasised.

22 4iP Council. (2018). *A FRAND Regime for Dominant Digital Platforms? Contribution by 4iP Council to the European Commission's Workshop on Shaping Competition Policy in the Era of Digitisation*. https://ec.europa.eu/competition/information/digitisation_2018/contributions/4ip_council.pdf

23 Ciuriak, D. (2018, 15 November). Industrial-era Investment Strategies Won't Work in a Data-driven Economy. *Centre for International Governance Innovation*. https://www.cigionline.org/articles/industrial-era-investment-strategies-wont-work-data-driven-economy

GLOBAL INFORMATION SOCIETY WATCH 2019

Artificial intelligence:
Human rights, social justice and development

Artificial intelligence (AI) is now receiving unprecedented global attention as it finds widespread practical application in multiple spheres of activity. But what are the human rights, social justice and development implications of AI when used in areas such as health, education and social services, or in building "smart cities"? How does algorithmic decision making impact on marginalised people and the poor?

This edition of Global Information Society Watch (GISWatch) provides a perspective from the global South on the application of AI to our everyday lives. It includes 40 country reports from countries as diverse as Benin, Argentina, India, Russia and Ukraine, as well as three regional reports. These are framed by eight thematic reports dealing with topics such as data governance, food sovereignty, AI in the workplace, and so-called "killer robots".

While pointing to the positive use of AI to enable rights in ways that were not easily possible before, this edition of GISWatch highlights the real threats that we need to pay attention to if we are going to build an AI-embedded future that enables human dignity.

GLOBAL INFORMATION SOCIETY WATCH
2019 Report
www.GISWatch.org

APC
ARTICLE 19
Sida