



**APC Submission to the UN Special Rapporteur
on contemporary forms of racism, racial
discrimination, xenophobia and related
intolerance:
Thematic report on new information
technologies, racial equality, and non-
discrimination**

*Association for Progressive Communications (APC)
December 2019*

Table of contents

Introduction.....	3
a. Most pressing forms and mechanisms of structural racial discrimination and inequality associated with new information technologies, including algorithmic discrimination and bias, and automated and interactive machine decision making.....	3
Predictive policing.....	3
Facial recognition.....	4
Neocolonialism.....	4
Privacy.....	5
Labour.....	5
Poverty.....	6
Gender discrimination.....	7
Inclusion.....	7
b. Mapping the political economy and other structural forces driving the respective patterns of racial discrimination and exclusion.....	8
c. Outlining the appropriate human rights legal, policy and advocacy responses rooted in global human rights equality and non-discrimination norms.....	9
d. Preferred definitions of various AI, ML, and new information/digital tech.....	12
About APC.....	13

Introduction

The Association for Progressive Communications welcomes the focus of the UN Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance on acute and structural threats that new information technologies, including artificial intelligence (AI), pose to the rights to non-discrimination and racial equality, human rights principles and standards, and also welcomes the opportunity to contribute to her report on this important topic.

Contrary to popular belief that AI is neutral, infallible and efficient, it is a socio-technical system with significant limitations.¹ One possible explanation is that the data used to train AI systems “emerges from a world that is discriminatory and unfair, and so what the algorithm learns as ground truth is problematic to begin with.”² Humans building these systems have their biases and train systems in a way that is flawed.

But there is another explanation that focuses on the global power relations in which these systems are built. AI systems are flawed because they amplify some voices at the expense of others, and are built by a few people and imposed on others. “In other words, the design, development, deployment and deliberation around AI systems are profoundly political.”³ The impact of AI is significant and unique, depending on the context in which these systems are deployed, and the purposes for which they are built. It is a matter of reckoning with the imperfect, discriminatory and unfair world from which these systems arise, and the underlying structural and historical legacy in which these systems are applied.

The 2019 edition of the Global Information Society Watch (GISWatch) report, produced by APC in partnership with ARTICLE 19, focuses on the impacts of AI from the perspectives of human rights, development and social justice, with a specific focus on the global South. This submission draws heavily on GISWatch 2019, extracting elements that are most relevant to the topic of this consultation.⁴

a. Most pressing forms and mechanisms of structural racial discrimination and inequality associated with new information technologies, including algorithmic discrimination and bias, and automated and interactive machine decision making

Predictive policing

Increasingly, law enforcement agencies use AI for predictive policing – mainly, for building “predictive maps of crime”, predicting what areas are more prone to crime or even which persons are more likely to be perpetrators or victims of a crime. In his report, “AI policing of people, streets and speech”,⁵ Luis Fernando García Muñoz explains that these tools rely on multiple sources of data such as criminal records, crime statistics, and the demographics of people or neighbourhoods, among others. “Many of

¹Marda, V. (2019). Introduction. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

²Ibid.

³Ibid.

⁴The full edition is available at: https://giswatch.org/sites/default/files/gisw2019_artificial_intelligence.pdf

⁵García Muñoz, L. F. (2019). AI policing of people, streets and speech. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

these data sets are flawed and biased in ways which can reinforce racial and other types of discrimination,” says García Muñoz, yet predictions made by AI systems trained with skewed data are often seen as “neutral” or “objective”, further ingraining discriminatory and abusive practices, he notes. Moreover, these systems are implemented without transparency, accountability or community participation in the decisions around their implementation or in the evaluation and oversight of their impacts, further limiting the detection and remedy of undesired outcomes.

In recent years, for example, the use of big data for predictive policing seems to be a popular trend in Latin America, according to Paz Peña and Joana Varon.⁶ The authors outline AI systems that are being used (or are meant to be deployed) in Argentina, Brazil, Chile, Colombia, Mexico and Uruguay, among others. As Peña and Varon say, critics point to the negative impacts of these systems on poorer neighbourhoods and other affected communities, including police abuse.

The Thailand country report⁷ looks at the use of predictive technology to combat human trafficking and exploitation, examining potentially positive use cases. The authors acknowledge the critiques that these systems can reproduce existing patterns of discrimination, reflect the widespread biases in society, or even exacerbate existing inequalities by suggesting that historically disadvantaged groups actually deserve less favourable treatment.⁸ However, the report also explores how machine learning could support exploited workers in vulnerable situations: for example, to create targeted education and awareness-raising campaigns; support “frontline responders”⁹ to proactively screen against current practices of exploitation; and inform evidence-based policy to support the prosecution of exploiters. The report introduces Apprise, an expert system that frontline responders are currently using in Thailand to support the initial screening stage of victim identification.¹⁰

Facial recognition

Biometric-based other data-intensive systems are being deployed around the world in ways that reinforce and exacerbate structural racism and inequality associated with new technologies, in particular for people who are in positions of vulnerability and marginalisation for their multiple and intersecting forms of identity. Biometric-based identity systems such as facial recognition pose a particular risk, as is raised in the sections below on poverty, inclusion and gender discrimination, and also highlighted in the Brazil country report¹¹ and the thematic report on the use of AI for surveillance and policing.¹²

⁶Peña, P., & Varon, J. (2019). Decolonising AI: A transfeminist approach to data and social justice. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

⁷Thinyane, H., & Puthawong, M. (2019). Apprise: Using AI to Unmask Situations of Forced Labour and Human Trafficking”, p. 221. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

⁸Ibid.

⁹The report uses the term “frontline responders” to refer to the broad range of stakeholders that assess working conditions and help potential victims access help or remediation channels – including police, labour inspectors, auditors and NGOs. Ibid.

¹⁰Ibid.

¹¹Canto, M. (2019). “We don’t need no observation”: The use and regulation of facial recognition in Brazilian public schools. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

¹²García Muñoz, L. F. (2019). Op. cit.

Neocolonialism

In their thematic report, Anita Gurumurthy and Nandini Chami¹³ state that the AI-led global order is “entrenched firmly in what activists and scholars have argued is a form of neocolonisation.” Economic power is a function of how AI technologies are employed in networked systems organised around incessant data processing, explain the authors. According to them, “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.”

The authors argue that debates on AI governance propose liberalist, structural interventions (focusing on correcting misrecognition) at best and neoliberal, individualistic fixes (that transfer burdens of navigating the digital economy on individuals) at worst. Hence, Gurumurthy and Chami propose a shift in AI governance towards a framework that considers the political economy of data ownership and control.

Privacy

Several reports reflect on the relationships between the right to privacy, data protection and algorithms.¹⁴ For example, in her report on “The weaponisation of AI: An existential threat to human rights and dignity”, Rasha Abdul Rahim contends that the mass collection and profiling of personal data could also have an impact on the right to equality and non-discrimination. According to Abdul Rahim, “Systems employing machine-learning technologies can vastly and rapidly reinforce or change power structures, as the data sets used to teach algorithms contain historical biases which are then reproduced and amplified.”¹⁵ Abdul Rahim’s report also addresses how the masses of data collected to train targeting algorithms to profile personal data and create patterns on the basis of which Autonomous Weapons Systems (AWS) would make decisions on when to use force and against whom, undermine the right to privacy and the right to equality and non-discrimination. AWS, explains the report, could therefore fuel the bulk collection of data and result in indiscriminate mass surveillance, which is never a proportionate interference with the right to privacy.¹⁶

Labour

Especially in the global South, says Noopur Raval’s report on AI and labour, we are witnessing not total and complete automation of/in work but rather a “heteromation”, a reorganisation in the division of labour between humans and machines.¹⁷ The report calls for a human-centric orientation to the questions

¹³Gurumurthy, A., & Chami, N. (2019). Radicalising the AI governance agenda. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

¹⁴See for example, Comminos, A., Muller, E. S., & Mutung’u, G. (2019). Artificial intelligence for sustainable human development, and Perkov, B., & Kalezic, P. (2019). Living under the watchful eye: Implications of facial recognition surveillance in Serbia. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

¹⁵Abdul Rahim, R. (2019). The weaponisation of AI: An existential threat to human rights and dignity. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

¹⁶Ibid.

¹⁷Raval, N. (2019). Automating informality: On AI and labour in the global South. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

of AI and labour in the global South, and recentring human work alongside machine intelligence. Such a shift also means that we are not necessarily talking about “machines replacing humans” but rather displacing traditional work roles and thus calling for a re-imagination of human work.

The report addresses the increasing platformisation and algorithmic management of work in India and illustrates the socio-technical effects of AI implementation in work, with a focus on prevalent informality and vulnerability as well as social hierarchies of caste, gender and class in India.

Raval focuses on “algorithmic platforms”, the service intermediaries that manage pools of workers in real time, such as Uber, Olacabs, food-delivery platforms and e-commerce platforms. The report reflects on the granular surveillance used to track platform workers, worker and customer communications, rest times, and the consequences of surveillance technologies in the workplace and workers’ rights to privacy and to enjoy and be fulfilled by work.¹⁸ “It is worth noting that it is not a coincidence that such technologies are being both deployed on and refined through their use on informal and contract workers,” says Raval, “given the overall lack of transparency and monitoring of work conditions among these groups.”

Platforms produce complex new realities for work in the global South. While AI-embedded work platforms widen participation for some actors, they have also been known to leverage and reinforce the existing socio-cultural hierarchies that shape certain forms of work themselves.¹⁹ The datafication of already vulnerable worker/citizen subjects produces a kind of “double marginalisation”, similar to the datafication of refugees and asylum seekers.²⁰

Another strand of AI, work and surveillance in the global South relates to the “hidden ghost work” of data cleaning, image labelling, text processing and content moderation being performed by back-end workers across developing economies.²¹ And the gender aspects of this. A recent paper by APC’s Namita Aavriti addresses how moderation is also done voluntarily by women of colour, queer and trans people, and racial minorities – “guerilla moderation” and labour, ultimately feminised, devalued, and offshored.²² As Aavriti points out, on the one side we have humans, their prejudices and the exhaustion of their labours, and on the other side are algorithms and (not) sophisticated computational and deep learning models that learn largely from the assumptions and biases that humans have, and then reproduce them on a massive scale.

Poverty

AI systems are increasingly used by governments to deploy anti-poverty programmes. These systems are being used in Latin America, for instance, raising concerns regarding data collection and management, and their potential to increase social injustice in the region.²³ Peña and Varon explain some of the risks of the datafication imposed by these systems and the quantification of the self and bodies, and the lack of

¹⁸Ibid.

¹⁹Ibid.

²⁰Ibid.

²¹Ibid.

²²Aavriti, N. (2019, 23 September). Are we any better at judging right from wrong? Automation in content moderation. *GenderIT.org*. <https://www.genderit.org/articles/are-we-any-better-judging-right-wrong-automation-content-moderation>

²³Peña, P., & Varon, J. (2019). Op. cit.

space for communities for re-negotiation. In other words, these systems replace “social identity” with “system identity”.²⁴

This is the case with the Plataforma Tecnológica de Intervención Social and Alerta Infancia in Chile, applied to minors in poor communities. The system assigns risk scores to communities, generating automated protection alerts, which then allow “preventive” interventions. Civil society groups working on child rights declared that, beyond surveillance, the system encourages forms of stigmatisation, discrimination and even criminalisation. Peña and Varon explain in their report that this also affects indigenous peoples, migrant populations and those with lower economic incomes, ignoring that a growing cultural diversity demands greater sensitivity, visibility and respect, as well as the inclusion of approaches with cultural relevance to public policies.

The Brazil country report addresses the use of facial recognition technology in public schools in that country and its connection with the Bolsa Família programme, a direct income transfer programme aimed at families living in poverty and extreme poverty.²⁵ The author, Mariana Canto, points out that much of the peripheral and vulnerable population in Brazil is being registered in this “experiment” – that is, data is being collected on vulnerable and marginalised groups – and highlights the risks of biased machine learning algorithms that police, profile and punish minorities, and the deprivation of civil rights of certain groups of society as a result of social inequalities and power relations in society.²⁶

Gender discrimination

Through the case of the use of algorithms to “predict teenage pregnancy” (a programme called Plataforma Tecnológica de Intervención Social in Salta, Argentina), Peña and Varon show how AI solutions claim to be neutral and objective, and instead have been increasingly deployed to support potentially discriminatory public policies that undermine human rights of unprivileged people such as poor women and girls.²⁷

This platform includes monitoring and censoring women and their sexual and reproductive rights and the system is now being deployed in other Argentinian provinces, such as La Rioja, Tierra del Fuego and Chaco, and has been exported to Colombia and implemented in the municipality of La Guajira.²⁸

Facial recognition surveillance also impacts differently on women and non-binary individuals. For example, explains García Muñoz, a study of types of facial analysis software showed that while the error rate in determining the gender of light-skinned men was 0.8%, the error rate for darker-skinned women reached up to 34% in some cases. “This gender and racial bias creates an aggravated risk of perpetuating the discriminatory effects that policing and the criminal justice system have been found to be responsible for,” says the report.²⁹

In their country report “Feminist or not? Canada’s challenges as it races to become a leader in artificial intelligence”, Heath, Molnar and Poetranto discuss Canada’s AI efforts and the challenges to ensure that

²⁴Ibid.

²⁵Canto, M. (2019). Op. cit.

²⁶Ibid.

²⁷For the gender implications of AI, please also see the work of Dr. Rachel Adams: <https://link.springer.com/article/10.1007/s00146-019-00918-7>

²⁸Peña, P., & Varon, J. (2019). Op. cit.

²⁹García Muñoz, L. F. (2019). Op. cit.

this does not contradict the country's commitments to human rights and gender equality.³⁰ Some of the challenges outlined by the authors include the gender and racial imbalance in science, technology, engineering and mathematics (STEM), and the lack of accountability and transparency in the government's use of emerging technologies, including in immigration and policing.

Inclusion

Comninos, Muller and Mutung'u discuss in their report the application of AI for development with a focus on the use of these systems in the roll-out of digital ID in the global South. On the one hand, the potential of digital ID has spurred the uptake of official state ID programmes and welfare and social protection schemes. However, digital ID programmes present several nuanced challenges, say the authors.

For instance, India's digital ID project, the Unique Identification Authority of India (UIDAI, or Aadhaar), implemented since 2008, aims to provide every Indian resident with an ID number linked to their demographic and biometric data. Aadhaar is linked to welfare and social protection delivery.

Controversially, an Aadhaar ID is only available to Indian residents who are registered in the National Register of Citizens (NRC) and not to refugees or stateless persons. "There is a possibility that over 1.9 million people in the north Indian state of Assam could be excluded from the NRC, and thus from citizenship, Aadhaar ID, state services and social welfare," explains the report. Hence, Aadhaar demonstrates that new ways of counting citizenry will intersect with structures of exclusion, possibly creating new layers of exclusion or amplifying existing ones.³¹

The report also looks at Kenya's digital ID programme, the National Integrated Identity Management System (NIIMS) – also known as Huduma Namba – proposed in August 2019, which has three components: a centralised database, a unique identifier for each person, and a card to be carried for mandatory use in accessing services. Unique identifiers include fingerprints, hand geometry, earlobe geometry, retina and iris patterns, and voice waves. This system, say Comninos, Muller and Mutung'u, intersects with existing topologies of societal inclusion and exclusion, ethnic cleavages, and statelessness in Kenya. For example, in Kenya, the Nubian, Shona and Makonde communities, which have historically lived in areas that became borders during colonialism, are subjected to long vetting processes before they can acquire identity documents. "When governments link digital ID to determination of citizenship, it puts at risk populations who for historical reasons lack primary identification documents," conclude the authors.

b. Mapping the political economy and other structural forces driving the respective patterns of racial discrimination and exclusion

A number of GISWatch authors shed light on the multiple and varied structural forces that drive respective patterns of racial discrimination and exclusion in the application of AI. For example, according to Gurumurthy and Nandini Chami, "In the race towards the 'Fourth Industrial Revolution' developing countries are caught up in the language of 'innovation' and 'entrepreneurship', authoring national plans

³⁰Heath, V., Molnar, P., & Poetranto, I. (2019). Feminist or not? Canada's challenges as it races to become a leader in artificial intelligence. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

³¹Comninos, A, Muller, E. S., & Mutung'u, G. (2019). Op. cit.

and road maps for their digital start-up ecosystem and upskilling of workers.” These efforts, say the authors, 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.”³² Gurumurthy and Chami call for new governance approaches to the AI economy with a focus on “equality of autonomy”: expanding individual and collective choices.³³ Developing countries need to use AI to create and/or deepen national capacity for moving out of low-value locations in the global value chain, building domestic capabilities and upskilling their populations, state Gurumurthy and Chami. However, this will be difficult to achieve if access to and ownership of data and digital intelligence are denied to these countries, the authors argue.³⁴

Peña and Varon also discuss the positive connotation around the idea that exploitation of vulnerable groups’ data, both by governments and private companies, will only benefit the population. Anti-poverty government programmes, say the authors, “reflect a positivist framework of thinking, where reality seems to be better understood and changed for good if we can quantify every aspect of our life.” This logic also promotes the vision that what humans shall seek is “progress”, which is seen as a synonym of augmented production and consumption, and ultimately means exploitation of bodies and territories. All these numbers and metrics about unprivileged people’s lives are collected, compiled and analysed under the logic of “productivity” to ultimately maintain capitalism, heteropatriarchy, white supremacy and settler colonialism.³⁵

As mentioned above, digital ID programmes in the global South often involve international companies. The growing “discourse on big data as a resource for development”, say Alex Comninos, Emily Shobana Muller and Grace Mutung’u, indicates a shift from the predominance of state-collected data to a big-data model where data is primarily collected and processed by corporations and only secondarily accessed by governments.³⁶

The transformation of agriculture and rural economies with digital development, automation and other computing technology and inequality at different levels are additional factors. Machine learning can be used to better understand plant diseases and AI can gather and analyse environmental information in real time. This provides many job opportunities, but exacerbates inequalities faced by smallholder farm owners who cannot afford technologies, enforcing the competitive advantage of industrial, corporate-controlled agriculture;³⁷ and food supply chains will be further integrated and concentrated, as pointed out by GRAIN.³⁸

³²Gurumurthy, A., & Chami, N. (2019). Op. cit.

³³Ibid.

³⁴Ibid.

³⁵Peña, P., & Varon, J. (2019). Op. cit.

³⁶Comninos, A, Muller, E. S., & Mutung’u, G. (2019). Op. cit.

³⁷Ibid.

³⁸GRAIN. (2019). Defending food sovereignty in the digital era. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

c. Outlining the appropriate human rights legal, policy and advocacy responses rooted in global human rights equality and non-discrimination norms

Due to the increasing impact of AI across the world, there has been a significant push towards thinking about the ways these systems should be governed, with various frameworks of reference arising.

These various governance frameworks take different forms. Multiple United Nations mechanisms are currently addressing the implications of AI from a human rights and development perspective, including the High-level Panel on Digital Cooperation, the Human Rights Council, UNESCO's World Commission on the Ethics of Scientific Knowledge and Technology, and the International Telecommunication Union's AI for Good Summit. Regional bodies like the European Union High-Level Expert Group on Artificial Intelligence also focus on questions of human rights and principles of social justice like fairness, accountability, bias and exclusion. Technical and multistakeholder bodies like the Partnership on AI and the Institute of Electrical and Electronics Engineers (IEEE) also invoke principles of human rights, social justice and development. All of these offer frameworks that can guide the design, development and deployment of AI by governments, and for companies building AI systems.³⁹ More focused on the data trust dimension and ethics dimensions, the Organisation for Economic Co-operation and Development adopted the OECD Principles on Artificial Intelligence in 2019, which recommend data trusts as a way to support the safe, fair, legal and ethical sharing of data. In June 2019, the G20 Digital Economy Ministers incorporated the OECD's recommendation on data trusts into their "human-centred AI Principles."⁴⁰

As Gurumurthy and Chami summarise, a systematic mapping by the Berkman Klein Center at Harvard 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.⁴¹

The human rights framework is a minimum requirement to which AI systems must adhere, given the impact on privacy, freedom of expression and freedom of assembly, among others. In the GISWatch introduction, Vidushi Marda proposes that this can be done by conducting thorough human rights impact assessments of AI systems prior to deployment, including assessing their legality against human rights standards, and by the private sector affirming commitment to the UN Guiding Principles on Business and Human Rights.⁴²

In the report on automated weapons and war, the accountability gap regarding these systems and lack of access to effective remedy for victims are addressed. "Since it is of course not possible to bring machines to justice, who would be responsible for serious violations?" stresses Rasha Abdul Rahim. For instance, nearly 250 tech companies, including XPRIZE Foundation, Google DeepMind and Clearpath Robotics, and over 3,200 AI and robotics researchers, engineers and academics have signed a Lethal Autonomous Weapons Pledge committing to neither participate in nor support the development, manufacture, trade or use of autonomous weapons. Given the high risk that AWS pose to human rights, Amnesty International

³⁹Marda, V. (2019). Op. cit.

⁴⁰Dawson, P., & Abuhamad, G. (2019). Towards data governance that empowers the public. In A. Finlay (Ed.), *Global Information Society Watch 2019 – Artificial intelligence: Human rights, social justice and development*. APC and ARTICLE 19. <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>

⁴¹Gurumurthy, A., & Chami, N. (2019). Op. cit.

⁴²Marda, V. (2019). Op. cit.

is calling for a legally binding instrument to ensure that meaningful human control is retained over the use of force by prohibiting the development, production, transfer and use of AWS.⁴³

While human rights provide an important minimum requirement for AI systems to adhere to, social justice is another lens through which AI systems should be understood and critiqued. Finally, a third strand of governance emerges from a development perspective, to have the UN Sustainable Development Goals (SDGs) guide responsible AI deployment and how AI could contribute to achieve the SDGs, and to leverage AI for economic growth, particularly in countries where technological progress is synonymous with economic progress.⁴⁴

The Open Data Institute (ODI) announced a partnership with the UK Office for Artificial Intelligence and Innovate UK to run three data trust pilots focusing on tackling illegal wildlife trade, reducing food waste and improving municipal public services. Canada, which published its first national strategy for AI in 2017,⁴⁵ announced in May 2019 a new Digital Charter that referenced data trusts as a possible way to facilitate data sharing in a privacy and security-enhancing manner for research and development purposes in areas such as health, clean technology or agribusiness.⁴⁶

The extent to which existing regulations in national, regional and international contexts apply to these technologies is unclear, although a closer analysis of data protection regulation, discrimination law and labour law is necessary.⁴⁷

It is vital to acknowledge that AI governance suffers from structural inequalities, says Marda. Jurisdictions from developing countries do not form part of the evidence base on which AI policies are built. "Narratives around AI that inform governance models need to be driven in a bottom-up, local-to-global fashion that looks at different contexts with the same level of granularity in the global South as was afforded to the global North," she states. "Much like AI systems operate in societies that have underlying structural inequalities, the deliberation around AI suffers from a similar underlying structural problem," she concludes.⁴⁸

According to Gurumurthy and Chami, mainstream debates on AI governance address human rights considerations connected with privacy, equality and non-discrimination, the uncertain future of work, and challenges regarding democracy. However, they do not fully address the "entanglement of AI in neoliberal capitalism and what this means for the life-chances of individuals and communities."⁴⁹

Gurumurthy and Chami identify critical blind spots in the AI governance discussions:

- Collective autonomy and choice in the debate on AI and human rights: Existing institutional and techno-governance mechanisms fail to imagine redress to individuals and communities caught in relationships of exploitation that are based on uneven and unfair distribution of intelligence capital.⁵⁰

⁴³Abdul Rahim, R. (2019). Op. cit.

⁴⁴Comninos, A, Muller, E. S., & Mutung'u, G. (2019). Op. cit.

⁴⁵Heath, V., Molnar, P., & Poetranto, I. (2019). Op. cit.

⁴⁶Dawson, P., & Abuhamad, G. (2019). Op. cit.

⁴⁷Marda, V. (2019). Op. cit.

⁴⁸Ibid.

⁴⁹Gurumurthy, A., & Chami, N. (2019). Op. cit.

⁵⁰Ibid.

- Economic self-determination in the debate on AI: Developing country governments fail to understand development is a “competitive and global undertaking”, characterised by a sustained and continuing effort to capture opportunities for higher value knowledge and technological capabilities.⁵¹

Transforming the political economy of data ownership and control that is deepening global development fault lines is the critical missing link, argue Gurumurthy and Chami: “The AI governance agenda therefore needs to be transformed and radicalised, embracing a focus on data and AI constitutionalism,” they propose. For this, two critical steps need to be accomplished: acknowledging data sovereignty as part of the right to development, and an international mechanism to enforce corporate accountability, reining in transnational digital corporations.⁵²

In their report, Peña and Varon also argue for the necessity to connect a social and environmental justice agenda to the data revolution. They propose to “build a transfeminist critique and framework that offers not only the potential to analyse the damaging effects of AI, but also a proactive understanding on how to imagine, design and develop an emancipatory AI that undermines consumerist, misogynist, racist, gender binarial and heteropatriarchal societal norms.”⁵³ Beyond even a human rights framework, the authors propose a decolonial and tranfeminist approach as “tools to envision alternative futures and overturn the prevailing logic in which AI systems are being deployed.” Transfeminist values need to be embedded in AI systems, argue Peña and Varon.⁵⁴

To put this decolonial feminist approach into practice, the NGO Coding Rights, in partnership with MIT’s Co-Design Studio, developed a game they called the “Oracle for Transfeminist Futures” to collectively think about “what kind of transfeminist values will inspire and help us envision speculative futures.”⁵⁵ Values such as agency, accountability, autonomy, social justice, non-binary identities, cooperation, decentralisation, consent, diversity, decoloniality, empathy, security, among others, emerged in workshops organised around the game.

Heath, Molnar and Poetranto’s report also points out gaps regarding AI’s deployment and the impact on vulnerable groups such as migrants who identify as women or gender non-binary. “Without oversight to ensure diversity and proper impact assessments, the benefits of new technologies like AI may not accrue equally,” conclude the authors.

d. Preferred definitions of various AI, ML, and new information/digital tech

In developing GISWatch 2019, we discovered that there is not one, single definition of AI that is widely accepted, and that a lot of civil society organisations are struggling with the definitions, let alone how to integrate and respond to AI, and are farther still from knowing how to approach governance issues around them. The lack of cohesion and diverse interpretations are as revealing as any concrete definition. That said, the publication used the following definitions:

⁵¹Ibid.

⁵²Ibid.

⁵³Peña, P., & Varon, J. (2019). Op. cit.

⁵⁴Ibid.

⁵⁵<https://www.transfeministech.codingrights.org>

- “Artificial intelligence” is a blanket term that could refer to varying levels and kinds of big data and algorithmic innovations.⁵⁶ AI is broadly defined as the ability of computers to exhibit intelligent behaviour. Much of what is referred to as “AI” in popular media is one particular technique that has garnered significant attention in the last few years – machine learning (ML).⁵⁷
- “Machine learning” is the process by which an algorithm learns and improves performance over time by gaining greater access to data. Given the ability of ML systems to operate at scale and produce data-driven insights, there has been an aggressive embracing of its ability to solve problems and predict outcomes.⁵⁸
- “Autonomous weapons systems” (AWS) encompass both lethal and less-lethal systems. AWS can be defined as weapons capable of selecting and applying force against targets without meaningful control. Autonomy in weapons systems should be understood as a continuum, and AWS should not be confused with unmanned aerial vehicles (UAVs), commonly referred to as drones, which are remotely piloted by a human operator. By contrast, AWS would incorporate software and algorithms which, on their own, would be able to make critical determinations about life and death.⁵⁹

About APC

APC is an international network of civil society organisations founded in 1990 dedicated to empowering and supporting people working for peace, human rights, development and protection of the environment, through the strategic use of information and communication technologies (ICTs). We work to build a world in which all people have easy, equal and affordable access to the creative potential of ICTs to improve their lives and create more democratic and egalitarian societies.

⁵⁶Raval, N. (2019). Op. cit.

⁵⁷Marda, V. (2019). Op. cit.

⁵⁸Ibid.

⁵⁹Abdul Rahim, R. (2019). Op. cit.