

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 Comminos (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 Milanés (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

Creative Commons Attribution 4.0 International (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0/>

Some rights reserved.

Global Information Society Watch 2019 web and e-book

ISBN 978-92-95113-13-8

APC Serial: APC-201910-CIPP-R-EN-DIGITAL-302

Disclaimer: The views expressed herein do not necessarily represent those of Sida, ARTICLE 19, APC or its members.

Country and regional reports introduction

Alan Finlay

Flawed digital technologies are increasingly at the core of our daily activities, and they interact with us. – Franco Giandana (Creative Commons Argentina/Universidad Nacional de Córdoba)

The 43 reports published here show there are few areas where the potential of artificial intelligence (AI) is not being explored. Even in so-called “least developed countries”, AI experiments and programmes are proliferating. For example, in Rwanda, “innovation companies [are] attracted by [it] being a ‘proof-of-concept’ country where people who are thinking about setting up businesses are offered a place to build and test prototypes before scaling to other countries.” In Benin, among several AI pilots including big data labs, training drones to work in areas such as health, agriculture and conservation, and an annual contest to combine algorithms with local games such as *adji* (dominoes), at least two initiatives in the country focus on empowering women and girls in the use of robotics and AI. “Despite the lack of an enabling environment,” writes Abebe Chekol (Internet Society – Ethiopian Chapter), “the country is becoming a thriving centre for AI research and development.”

The authors take a loose definition of AI, and in doing so cast a relatively wide net on what they consider relevant for discussion. What all of the reports have in common, however, is a focus on when AI – variously defined – meets the intersection of human rights, social justice and development, and “shocks” this intersection; sometimes for the better, but also often raising critical issues that demand the attention of human rights advocates. While the focus in these reports is on perspectives from the global South, reports from countries such as Canada, Germany, Russia, the Republic of Korea and Australia are included, offering a useful counterpoint to countries where the application of AI is only just emerging. Three regional reports are also included: largely the result of authors feeling the need to take a regional perspective on the theme, rather than focusing on developments in a

particular country. Taken together, these reports offer a snapshot of AI-embedded future/s at different stages of development, and a useful opportunity to identify both the positive potential and real threats of AI deployment in diverse contexts.¹

Several reports are concerned with the digitalisation of the workplace, and the impact of AI and automation on worker rights. If predictions of job losses are anything to go by, economies are set to be reshaped entirely. In a country like Ethiopia, for example, about 85% of the workforce is said to be vulnerable to technological replacement, while a similar percentage of those currently employed in Argentina are predicted to need reskilling. In Bangladesh, women working in the ready-made garment sector, “who are at the bottom of the production process and are often engaged in repetitive tasks,” are the mostly likely to suffer the results of automation.

The claim that AI, while shedding menial and repetitive jobs, will create a newly skilled and re-employable workforce currently lacks evidence to support it. This is the “elephant in the room” Deirdre Williams writes in her regional discussion on the Caribbean: “[W]hile there is also insistence that the same new technology will create new jobs, few details are offered and there is no coherent plan to offer appropriate re-training to those who may lose their jobs.” Given the high cost of “retooling” workers, they will instead be “pushed into lower-wage jobs or become unemployed,” writes Chekol. “[I]f the outcome is not mass unemployment, it is likely to be rising inequality.”

In many countries, a reinvigoration of the union movement is necessary. In Argentina, for example, unions report being unprepared to cope with the inevitable changes in the workplace:

Unions are behind in the debate on AI. [They] are disputing basic issues such as salary, health, loss of employment, with no economic stability

¹ Although not usual for GISWatch editorial policy, two country reports were included for India given the number of good proposals we received for that country. We also included a second report on Australia – on AI in the creative industries – because we felt that a focus on AI and the creative sector was a unique consideration not discussed in other country reports.

and pendular changes of government. We started to think in terms of emerging issues such as AI, but suddenly a new government destroyed even the ministry of work.

In that country it was necessary to create a union specifically focused on digital platforms – one that was able to offer collective voice and action for isolated, “on-demand” workers who face new challenges in demanding their rights.

As authors suggest, automation in the workplace is not inherently a bad thing, and can result in meaningful improvements in worker rights, such as assigning robots to do dangerous jobs, or relieving workers from the need to work in unhealthy workspaces. Yet the socioeconomic benefits and costs of workplace change need to be properly understood for their potential impact on society overall – and with the views of workers firmly embedded in policy design and decisions – rather than simply the result of a micro-focus on efficiency and more exact profit, with assumptions made about worker needs.

Authors also show how algorithmic design can perpetuate systemic discrimination – whether due to race, caste, class, gender, or against differently marginalised individuals, groups and communities. In her discussion of automation in the Australian welfare system, Monique Mann calls this a “structural and administrative *violence* [my italics] against those who are socially excluded and financially disenfranchised.” New forms of discrimination are also created (for example, by profiling the unemployed in Poland, what others have called a “double marginalisation” is felt), and the opportunities for discrimination are increased – through, for example, mass surveillance using facial recognition technologies.

Automated facial recognition (AFR) technology receives some attention in these reports, including its use in the persecutory surveillance of the Uyghur ethnic minority in Xinjiang in China, and in Brazilian schools to monitor (and ostensibly improve) attendance. But such a technological response to improve school drop-out rates among lower-income students, Mariana Canto from Instituto de Pesquisa em Direito e Tecnologia do Recife (IP.rec) argues, does not address the structural reasons for this – such as the relevance of the curriculum design, the need for students to work to support their families, and even the levels of crime and violence they are likely to experience on their way to school. Moreover, she adds:

It is important to remember that as systems are being implemented in public schools around the country, much of the peripheral and

vulnerable population is being registered in this “experiment” – that is, data is being collected on vulnerable and marginalised groups.

Mathana Stender from the Centre for the Internet and Human Rights (CIHR) points out in their report on the rise of automated surveillance in Germany that AFR “can [also] lead to automated human rights abuses.” And these abuses are indiscriminate:

With biased assumptions built into training of models, and flawed labelling of training data sets, this class of technologies often do not differentiate between who is surveilled; anyone who passes through their sensor arrays are potential subjects for discrimination.

The implication is that automated surveillance throws the net for potential discrimination wider, increasing the likelihood of global incidences of discrimination being experienced.

Beyond the effect of systemic bias in algorithmic decision making is the question of the quality of the data fed into AI systems. As Malavika Prasad and Vidushi Marda (India) put it, machine learning is “a process of generalising outcomes through examples” and “data sets have a direct and profound impact on *how* an AI system works – it will necessarily perform better for well-represented examples, and poorly for those that it is less exposed to.” For example, census or other socioeconomic data used to train AI or for automated decision making may be varied, and involve questionable methodologies or uneven research processes. This poses challenges for countries where this data is not “clean” or there is a lack of skills and resources to produce the necessary data. In Chile, write Patricia Peña and Jessica Matus from Instituto de la Comunicación e Imagen and Fundación Datos Protegidos, there is a need for “a chain of quality [control] from its collection, capture, use and reuse, especially when it is taken from other databases, so that no bias is generated,” while Ethiopia, “like most other African countries, has the lowest average level of statistical capacity. The lack of data, or faulty data, severely limits the efficacy of AI systems.”

Authors also raise concerns about the access to private data by businesses – especially given that private-public partnerships are seen as necessary to finance much public sector AI development (for example, think of the number of service-level arrangements necessary for smart cities to exist). But questions such as “What access do private companies developing AI technology have to private data?” and “Do they store the data, and for how

long?” largely go unanswered. In the Ponto iD surveillance system set up in Brazilian schools, there is a “lack of information that is included in the company’s privacy policy, or on city halls’ websites.” In its investigation into the introduction of AI in health care in Cameroon, Serge Daho and Emmanuel Biko-bo from PROTEGE QV write:

While patients’ data is collected by the Bonassama hospital and transferred to Sophia Genetics [a company based in the United States and Switzerland] using a secured platform, we could not determine how long this data is stored. [...] Is the confidentiality of Bonassama hospital patients a priority to Sophia Genetics? Hard to answer. Nor have we been able to find out whether or not the patients’ informed consent was requested prior to the data gathering process (the nurses we interviewed could not say).

Korean Progressive Network Jinbonet offers a practical account of policy advocacy in this regard – for example, explaining the legal difference between “pseudonymised” and “anonymised” data – and the litigating temperament necessary from civil society. As it found, not only did guidelines for the de-identification of personal data offer the opportunity for a lively trade in personal data between companies, but the state-run Health Insurance Review and Assessment Service had sold medical data from hospital patients to a life insurance company, and the data of elderly patients to Samsung Life. In Costa Rica, specific legal addenda are needed to oversee and secure the national medical database there, considered “one of the most important information resources in the country.”

The country reports suggest a mixed policy response to AI. A number of countries still do not have adequate data protection laws in place – an essential prerequisite for the roll-out of AI technology. If policies governing AI exist, they are often too broad to account for the real-life implications of the technologies on the rights of people and citizens, or they can become quickly outdated, leaving what Anulekha Nandi from Digital Empowerment Foundation (India) describes as a “governance vacuum over a general-purpose technology with unquantifiable impact on society and the economy.”

In this lacuna, a number of authors (e.g. Rwanda, Pakistan, Jamaica) reference the EU’s General Data Protection Regulation (GDPR) as a template for good governance that can be applied in their own country. Authors point out that a regional perspective on legislation is necessary – but not necessarily easy to achieve. In Latin America, for example,

despite the regional roll-out of Prometea in the judicial system in Buenos Aires, the Constitutional Court in Colombia, and at the Inter-American Court of Human Rights in San José, digitalisation plans in countries like Argentina tend to focus on building a country “brand” as a regional leader in the sector, while being quiet on the need to “[develop] common strategies with other governments in the region.” The result is a regional policy asymmetry, which Raymond Onuoha from the Regional Academic Network on IT Policy (RANITP) at Research ICT Africa argues is detrimental to the global competitive and developmental needs of regions. Moreover, even if regional policy symmetries exist, countries do not necessarily have similar capacities to implement the policies properly:

[M]any African countries are still dealing with basic issues of sustenance like food and housing etc., so technology and technology policy are not at the front burner of critical issues of concern. [...] A harmonised regional data protection policy regime for the continent might impose enforcement liabilities on member countries that lack the required resources for its implementation.

A key policy problem raised by several authors is the question of legal liability in the event of a “wrong” decision by an algorithm (or, in extreme cases, so-called “killer robots”). If this happens, it is unclear whether, for example, the designer or developer of the AI technology, or the intermediary service provider, or the implementing agent (such as a municipality) should be held liable. One solution proposed is that algorithms should be registered as separate legal entities, much like companies, in this way making liability clearer and actionable (a draft bill to this effect was being debated in Estonia – see the Ukraine country report).

Legislation also needs to have a clear view on when and how AI impacts on the current legal framework and rights of citizens. While in Australia, the country’s automated debt-raising programme “reverses the onus of proof onto vulnerable people (and thus overturns the presumption of innocence),” in Turkey, AI is being used in conjunction with copyright law to censor alternative media. Organisational and institutional culture also needs to be addressed in policy – involving significant effort in change management.

A number of authors are critical of the approach to policy design in their countries (in South Korea, for example, the government implements “policies focused on the utilisation rather than protection

of personal data”). They point out that policies often lack inclusivity and context – both essential to understanding the real-life implications on rights when implementing AI technologies. Policy needs to “centre” those most affected by technological changes. In Pune in India – described as one of the “top smart cities” in that country – the city’s smart sanitation project does not address the caste discriminations against the Dalit community, allowing, in effect, unaccountable private sector service providers to “discipline” already marginalised workers engaged in public services.

A useful methodology for better understanding the specific, contextual implications of AI on vulnerabilities and rights – and which can be built into policy design – is “risk sandboxing”. As Digital Empowerment Foundation explains:

Regulatory and data sandboxing are often recommended tools that create a facilitative environment through relaxed regulations and anonymised data to allow innovations to evolve and emerge. However, there also needs to be a concomitant risk sandboxing that allows emerging innovations to evaluate the unintended consequences of their deployment.

Effective policy advocacy may require significant capacity to be built among civil society organisations. For example, in countries like Poland, algorithmic calculations are part of legal and policy documentation. As Jędrzej Niklas writes, “for civil society organisations to successfully advocate for their interests, they must engage in the technical language of algorithms and mathematical formulas.” Reports such as those on the Seychelles and Malawi also show that some work needs to be done in raising public awareness of AI. Better public information on the practical benefits and human rights costs of AI needs to be made available – as well as more detail of the systems that are in place in countries.

Karisma Foundation offers a useful analysis of media coverage of Prometea in Colombia, showing that most reporting offered little understanding of the system: “[T]here was no explanation about what Prometea was, what it does and how it does

it.” When, as in the Ukraine, there appears to be reasonable public awareness of AI and at least some understanding of how it influences their lives, just less than a quarter of people surveyed said AI caused them “anxiety and fear”.

These reports suggest that this fear is not unfounded. Angela Daly (China) points to a global phenomenon of “ethics washing” – or the “gap between stated ethical principles and on-the-ground applications of AI.” While the city of Xinjiang is described as a “‘frontline laboratory’ for data-driven surveillance” in her report, IP.rec suggests “technological advancement” is as much driven by “desire” as anything else; but, “Does this desire turn people into mere guinea pigs for experimentation with new technologies?” For Maria Korolkova from the University of Greenwich, writing on Ukraine, an AI-embedded future risks “dislocating the axis of power in the citizen-state relationship necessary for democracy to function.”

There are several striking examples of the positive use of AI in these reports, and its potential to enable rights in ways that were not possible before. A number of reports focus on the health sector, but promising – although not problem-free – applications are also discussed in areas such as e-government (see South Africa for a useful discussion on this), in “unmasking” forced labour and human trafficking in Thailand, and in combating femicide (see Italy for an example of one of the country’s most advanced data-driven media research projects).

These reports nevertheless also show that an AI-embedded future poses fresh challenges for civil society advocacy – and that purposive action is required. Compromise might not always be possible. Joy Liddicoat, from the New Zealand Law Foundation Artificial Intelligence and Law Project, questions whether the multistakeholder approach to policy design is failing in the wake of the Christchurch terror attacks in her country. Niklas goes further, pointing to the need for a “radical political advocacy”, one that would “not only engage in changes or improvements to algorithms, but also call for the abolition of specific systems that cause harm.”

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

