



Artificial intelligence in education in India: questioning justice and inclusion

LSE Research Online URL for this paper: <http://eprints.lse.ac.uk/107456/>

Version: Published Version

Book Section:

Nandi, Anulekha ORCID: 0000-0002-1421-1525 (2019) Artificial intelligence in education in India: questioning justice and inclusion. In: Finlay, Alan, (ed.) Artificial Intelligence: Human rights, social justice and development. Global Information Society Watch reports (2019). Global Information Society Watch, 140 - 144. ISBN 9789295113138

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

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/>

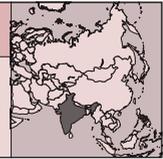
Some rights reserved.

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.



Digital Empowerment Foundation
Anulekha Nandi
www.defindia.org

Introduction

The National Strategy for Artificial Intelligence, released by NITI Aayog (the Indian government's policy think tank) in June 2018, underscores the government's policy intent to mainstream artificial intelligence (AI) in critical social and economic infrastructures.¹ Out of the 10 focus areas identified by the Artificial Intelligence Task Force, constituted by the Ministry of Commerce and Industry,² the education sector has seen the most successful public-private partnerships (PPPs) to deal with some of the institutional gaps plaguing the sector.³

With AI being a data-hungry technology, it becomes increasingly problematic when it is trained on the sensitive personal information of marginalised populations through service delivery in key social infrastructures like education. This is especially concerning given the current lack of a data protection regulation in India and the concomitant carve-outs for state functions in public service delivery. Moreover, the draft data protection bill which is currently tabled before the parliament does not contain explicit provisions on algorithmic decision making, including the right to be informed of its existence and the right to opt out, unlike the European Union's General Data Protection Regulation (GDPR).⁴

The impetus behind the deployment of AI has outstripped legal and regulatory development in the area, leaving a governance vacuum over a general-purpose technology with unquantifiable impact on society and economy. Given the multidimensional and cross-cutting risks and opportunities that this poses, along with complex and dynamic ethical

challenges, it becomes imperative to study and understand use cases to inform and work towards context-sensitive AI governance frameworks.

Institutional AI in the Indian technology and education paradox

The use of technology in education in India traverses the unequal realities of two facets of the country. On one hand there is the segment of the population with access to digital, social and economic resources and on the other there is the vast majority for whom even basic institutions of social infrastructure offer rudimentary support at best. Private investment in education technology – or EdTech – is a burgeoning industry which clocked a valuation of USD 4.5 billion globally in 2015.⁵ As per data from the research firm Tracxn, out of the 300 Indian start-ups that use AI as a core product offering, 11% are based in the education sphere.⁶ India's digital learning market was valued at USD 2 billion in 2016 and is projected to grow at a compound annual growth rate (CAGR) of 30% to reach USD 5.7 billion by 2020.⁷ However, the product offerings that result from these significant investments either aim to offer tutoring services, improve learning outcomes, or provide customised learning, all of which serve to leverage and augment the agency of the first segment of the population. The uptake, adoption and usage of these services proceed through the notice and consent protocols of informed consent due to service requirements of digital distribution platforms like Android's Google Play Store or Apple's Apple Store.

However, institutional applications of AI are based on public data collected by the government through service delivery, especially with regard to the social protection of marginalised, underserved and vulnerable populations, and are not undergirded by the need for adherence to data protection principles. Moreover, a joint reading of the draft data protection bill and judgement of the Supreme

1 NITI Aayog. (2018). *National Strategy for Artificial Intelligence*. New Delhi: NITI Aayog. https://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf

2 <https://www.aif.org.in>

3 NITI Aayog. (2018). Op. cit.

4 Das, S. (2018, 30 July). 8 differences between Indian data protection bill and GDPR. *CIO & Leader*. <https://www.cioandleader.com/article/2018/07/30/8-differences-between-indian-data-protection-bill-and-gdpr>

5 NITI Aayog. (2018). Op. cit.

6 Khera, S. (2019, 21 January). Artificial intelligence in education in India, and how it's impacting Indian students. *The News Minute*. <https://www.thenewsminute.com/article/artificial-intelligence-education-and-how-its-impacting-indian-students-95389>

7 NITI Aayog. (2018). Op. cit.

Court on the use of Aadhaar biometric information⁸ for exercising state functions of public service delivery highlights the exemptions from informed consent or other complementary data protection protocols for state functions aimed at social and economic inclusion. This begs the question as to how the sensitive personal data of citizens, especially of the marginalised, are to be protected within institutional applications of AI through PPP arrangements which lack transparency on the commercial service commitments, data protection protocols and safeguards, data-sharing arrangements and processes of labelling and annotation. These are further compounded by the gaps in explainability, framing, deployment and application.

One of the most pervasive problems within the Indian public education system has been low retention rates beyond primary education, with even lower rates for girls.⁹ In one of the first institutional applications of AI in the social sector in the country, the Government of Andhra Pradesh,¹⁰ in partnership with Microsoft, implemented machine learning and analytics through its Azure cloud platform to predict and prevent public school drop-outs. This report aims to use this partnership as a case study to throw into sharp relief the contextual parameters and questions that must be taken into account when evaluating institutional applications of AI in society and developing ethical governance frameworks that can answer to contextual nuances of the application taking cognisance of the actual incidence of its impact. It might also help highlight issues that would be helpful for future deployments in the sector to address, given that the NITI Aayog

plans to scale up this project with Microsoft on the basis of the Andhra Pradesh experience.¹¹

Identifying parameters of algorithmic decision making and its implications for justice and inclusion

Literacy levels in Andhra Pradesh have been the second lowest in the country, with one of the highest percentages of school drop-outs, most of whom come from farming families or those involved in agriculture.¹² It also topped the list of the highest number of female school drop-outs, with seven out of 10 girls dropping out of school before they reach the 10th standard.¹³ In a partnership with the Government of Andhra Pradesh, Microsoft offered its Azure cloud computing platform with machine-learning and analytics capabilities as a part of the overall Cortana Analytics Suite (CAS)¹⁴ to develop a predictive model for identifying school drop-outs in the state. The project commenced with a pilot of a little over 1,000 schools and 50,000 students and has now been rolled out to all 13 districts covering 10,000 schools and five million children. The aim of the project was for the information gathered and analysed to be made available to district education officers and school principals who could then deploy targeted interventions and customised counselling.¹⁵

Data was triangulated from three databases in order to build the data pipeline for the project. This included the Unified District Information System for Education (U-DISE), containing school infrastructure information and the data on teachers and their work experience, education assessment data from multiple sources, and socioeconomic data from the UIDAI¹⁶ Aadhaar system.¹⁷ By aggregating these multiple data points from different sources, the aim of the project is to track the students' journey through

8 The Aadhaar is a 12-digit unique identification system based on biometric information and demographics issued to an Indian resident. It is governed by the Aadhaar (Targeted Delivery of Financial and Other Subsidies, Benefits and Services) Act, 2016. It became controversial when mobile phone service providers and banks started asking for the card as a condition for using their services. More problematically, it became conditional for the delivery of critical social protection schemes like midday meals to underserved students, availing rationed food items, pension schemes, etc., in some cases with people denied these services dying. The card was the subject of cases filed before the Supreme Court of India which challenged its constitutional validity due to its privacy infringing features and that it was being required to access private sector services. Though the Supreme Court upheld the fundamental right to privacy, in September 2018 it also upheld the constitutional validity of the identification system in that it allowed Aadhaar-based authentication for establishing the identity of an individual for receipt of a subsidy, benefit or service provided by the government by retaining section 7 of the Aadhaar Act that allows for welfare to be made contingent on the production of Aadhaar.

9 Taneja, A. (2018, 31 January). The high drop out rates of girls in India. *Live Mint*. <https://www.livemint.com/Opinion/iXWvKng7uU4L8v05XbDn9l/The-high-dropout-rate-of-girls-in-India.html>

10 Andhra Pradesh is a state in southern India.

11 Agha, E., & Gunjan, R. K. (2018, 28 April). NITI Aayog, Microsoft Partner Up to Predict School Dropouts Using Artificial Intelligence. *News18*. <https://www.news18.com/news/india/niti-aayog-microsoft-partner-up-to-predict-school-dropouts-via-artificial-intelligence-1732251.html>

12 India Today. (2016, 23 April). Education survey shows the poor state of Telengana, Andhra Pradesh. *India Today*. <https://www.indiatoday.in/education-today/news/story/andhra-pradesh-education-319526-2016-04-23>

13 Baseerat, B. (2013, 10 October). Andhra tops in girl school dropouts: Activists. *Times of India*. <https://timesofindia.indiatimes.com/city/hyderabad/Andhra-tops-in-girl-school-dropouts-Activists/articleshow/23937897.cms>

14 Cortana Analytics Suite is the fully managed big data and advanced analytics suite.

15 Srivas, A. (2016, 10 May). Aadhaar in Andhra: Chandrababu Naidu, Microsoft have a plan for curbing school dropouts. *The Wire*. <https://thewire.in/politics/aadhaar-in-andhra-chandrababu-naidu-microsoft-have-a-plan-for-curbing-school-dropouts>

16 The Unique Identification Authority of India is the entity mandated to issue the 12-digit Aadhaar number and manage the Aadhaar database.

17 Srivas, A. (2016, 10 May). Op. cit.

the education system by providing a 360-degree view of students after mapping close to 100 variables. Initial results from the project reaffirmed longstanding notions behind school drop-outs. These include girls being more likely to drop out in the absence of adequate toilet facilities, higher drop-out rates among students failing to score well in key subjects like English and mathematics, which reduces their faith in formal education, along with the role of the socio-economic status of the family and the wider community to which the student belongs.¹⁸ In a study based on the National Family and Health Survey-3, it was found that drop-outs tended to be higher among children belonging to minority Muslim families, scheduled castes and scheduled tribes.¹⁹ Further, children belonging to illiterate parents were four times more likely to drop out than those belonging to literate parents. The possibility of children of non-working parents dropping out is also relatively high.²⁰

Anil Bhansali, the managing director of Microsoft Research and Development, had told the online news outlet *The Wire* in 2016 that the CAS suite deployed in the project “can provide a lot of useful insight as long as you pump in the data and the right modelling,”²¹ with “right modelling” being the operational phrase. With algorithmic decision making coming to play an increasingly significant role in institutionalising individual and systemic bias and discrimination within social systems, it becomes important to evaluate the processes through which these are pervasively deployed.

Data choices: The pilot project was restricted to students of the 10th standard. This is because, according to Bhansali, the 10th standard represents one of the few inflexion points when one takes their first standardised tests and after which a reasonable number of students drop out on their way to 11th standard. Another likely reason is that 10th standard results are already online and the education department has access to gender and subject grading data through examination hall tickets.²²

Educational assessment information for lower classes entails the herculean task of having to be digitised in order to be of use in a machine-learning system.²³

However, it is also the case that the drop-out rates are the highest in secondary education (standards 9 and 10),²⁴ coinciding with the completion of standard 8 after which midday meals are no longer provided, which are a major factor driving school attendance.²⁵ Those who continue beyond standard 8 to reach standard 10 show a comparative degree of resilience to the non-provision of these sorts of interventions aimed at ameliorating the disadvantageous socioeconomic conditions behind school drop-outs. Therefore, using such data as a training model for the system misplaces the inflexion point and thereby undermines other structural and intersectional socioeconomic issues driving high rates of school drop-outs during the transition to secondary education from standard 8 to standard 9. This leads to elision of the structural socioeconomic parameters that constrain equitable access to resources.

In addition, the U-DISE database containing information about teachers’ work experience does not necessarily map the effectiveness or efficacy of a given teacher and their contribution to better learning outcomes well.

Modelling and inferences: Data choices are not the only criteria determining the questions on inclusion and justice. Decision making regarding the input processes that develop the statistical models and inferences made are equally significant in determining the incidences of impact that a given machine-learning project is likely to have in the areas of its intervention.²⁶ Given the lack of transparency on the decision-making process, the insights gained from news reports on the subject show that the input process in developing the model involved a combination of existing knowledge, beliefs, and findings about the factors driving school drop-outs, coupled with the convenience of digitised data.²⁷ Since the extent to which these data were interpreted with bias during the input

18 Ibid.

19 Scheduled Castes and Scheduled Tribes are officially designated historically marginalised groups in India recognised in the Constitution of India.

20 M., Sateesh, & Sekher, T. V. (2014). Factors Leading to School Dropouts in India: An Analysis of National Family Health Survey-3 Data. *International Journal of Research & Method in Education*, 4(6), 75-83. https://www.researchgate.net/publication/269932850_Factors_Leading_to_School_Dropouts_in_India_An_Analysis_of_National_Family_Health_Survey-3_Data

21 Srivas, A. (2016, 10 May). Op. cit.

22 Examination hall tickets offer rights of admission to a test taker during state or national-level examinations. They contain details of the student like the identity number assigned to the student for the examination, a photograph and a signature along with details of the examination such as location and room (where applicable). Sometimes they also contain the student’s name and date of birth. There is no standard format for examination hall tickets and they differ from examination to examination.

23 Srivas, A. (2016, 10 May). Op. cit.

24 PRS Legislative Research. (2017, 2 October). Trends in school enrolment and drop-out levels. *Live Mint*. <https://www.livemint.com/Education/k1ANVHwheaCFWCupY3jkFP/Trends-in-school-enrolment-and-dropout-levels.html>

25 Jayaraman, R., Simroth, D., & de Vericourt, F. (n.d.). The Impact of School Lunches on Primary School Enrollment: Evidence from India’s Midday Meal Scheme. *Indian Statistical Institute, Delhi Centre*. www.isid.ac.in/~pu/conference/dec_10_conf/Papers/RajilJayaraman.pdf

26 Algorithm Watch. (2019, 6 February). ‘Trustworthy AI’ is not an appropriate framework. <https://algorithmwatch.org/en/trustworthy-ai-is-not-an-appropriate-framework>

27 Srivas, A. (2016, 10 May). Op. cit.

process is unclear, the extent to which biased socioeconomic profiles based on caste, gender and religion played a role in determining the drop-out rate, versus structural and institutional barriers, is also unclear. Moreover, it is not unlikely that such models can then influence seat allocations within higher education and government services based on such profiles, undermining India's constitutionally guaranteed affirmative action protections for marginalised and vulnerable groups.

This highlights the problem of using existing knowledge and statistics in an ahistorical and acontextual manner without duly quantifying the structural and institutional indicators that produce such inequalities in the first place. For example, if the model shows that a Scheduled Tribe girl from Jharkhand is more likely to drop out of school in the absence of a targeted intervention, could this lead to fewer seats allocated in higher education, and reservations in government services for women from the community? Moreover, coupled with data choices, it is unclear to what extent the data trained on standard 10 would be effective in predicting drop-out rates in the transition phase from upper primary (standard 8) to secondary school (standards 9 and 10) where arguably the driving factors are more structural and institutional as compared to performance in a given set of subjects.

Service agreements and data protection: Data sharing within PPPs is unclear due to a lack of transparency, especially in a country like India, which is yet to have its own data protection law but harbours high aspirations of becoming the world leader in AI adoption, deployment and innovation.²⁸ Bhansali has said that the data is stored in data centres located in India and is tied to the Andhra Pradesh government's account, and that Microsoft cannot own it or repurpose it.²⁹ While Microsoft might not own or repurpose the data, it is unclear whether it does or does not have the same rights over the insights generated out of processing of such data. It is also not clear what bespoke – or customised – data protection safeguards were incorporated, if at all, within the public-private agreements. An evaluation study of Microsoft's cloud computing shows that irrespective of the geographical location that a customer selects to locate their data, Microsoft warns that customers' data, including personal data, may be backed up in the United States (US) by default. Moreover, if any beta or pre-release Microsoft software was used or there was back-up of

web or worker role software³⁰ in any of its cloud services, data would be stored or replicated in the US.³¹ The movement and replication of data increases the attack surface. These fears are not allayed given that Microsoft is the second most targeted entity after the Pentagon,³² and Andhra Pradesh leads in the leakages of sensitive personal data of its constituents.³³

Conclusion

Given that AI systems are increasingly aiding state institutions in the allocation of resources, it becomes imperative that they align with principles of non-discrimination rather than perpetuate existing misallocations by creating pervasive systems of privilege being trained on unrepresentative data sets and models.

Significant international multilateral and multistakeholder attention has been diverted towards developing ethical governance frameworks for AI. This includes the OECD Principles on AI,³⁴ the European Commission Ethical Guidelines for Trustworthy AI,³⁵ as well as the Toronto Declaration: Protecting the Right to Equality and Non-Discrimination in Machine Learning Systems,³⁶ along with attention in other digital policy global initiatives like the United Nations High-Level Panel on Digital Cooperation.

However, these provide broad-based principles without adequately applied examples, which delimit

30 "Web Role is a Cloud Service role in Azure that is configured and customized to run web applications developed on programming languages/technologies that are supported by Internet Information Services (IIS), such as ASP.NET, PHP, Windows Communication Foundation and Fast CGI. Worker Role is any role in Azure that runs applications and services level tasks, which generally do not require IIS. In Worker Roles, IIS is not installed by default. They are mainly used to perform supporting background processes along with Web Roles and do tasks such as automatically compressing uploaded images, run scripts when something changes in the database, get new messages from queue and process and more." Source: <https://cloudmonix.com/blog/what-is-web-and-worker-role-in-microsoft-azure>

31 Calligo. (n.d.). *Microsoft Azure and Data Privacy*. https://calligo.cloud/wp-content/uploads/Azure-Data-Privacy-Stack.pdf?utm_campaign=Hybrid%20Azure&utm_source=hs_automation&utm_medium=email&utm_content=69182393&_hsenc=p2ANqtz-8e2YnyrNnNjouMWxrv80aYaLLlso_v18apwlbq3HTVVRqgl2WY94jmAKBStWuDTwC-U1F_NPubB4StezcA43mZl1YTw&_hsmi=69182393

32 Ibid.

33 See, for example: MediaNama. (2019, 29 May). Andhra Pradesh exposes Aadhaar of farmers - once again. *MediaNama*. <https://www.medianama.com/2019/05/223-andhra-pradesh-exposes-aadhaar-of-farmers-once-again/>; Jalan, T. (2018, 27 August). CCE-Andhra Pradesh leaks students' gender, caste, quota, Aadhaar data on website. *MediaNama*. <https://www.medianama.com/2018/08/223-apce-students-aadhaar-exposed/>; Tutika, K. (2018, 20 March). Aadhaar data leak of Andhra Pradesh women raises security concerns. *The New Indian Express*. www.newindianexpress.com/states/andhra-pradesh/2018/mar/20/aadhaar-data-leak-of-andhra-pradesh-women-raises-security-concerns-1789648.html

34 <https://www.oecd.org/going-digital/ai/principles>

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

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

28 NITI Aayog. (2018). Op. cit.

29 Srivas, A. (2016, 10 May). Op. cit.

their uptake and applicability and serve to act as an “alternative or preamble to regulation”, thereby diluting “state accountability and rights-based obligations.”³⁷ These also serve to act as light-touch non-discrimination norms that provide the leeway for businesses to not actually engage with non-discrimination principles within data choices, modelling, design and application, thereby ending up entrenching discrimination by making inequalities institutionally pervasive.

A second approach, which is a technical approach, aims to ensure fairness, accountability and transparency (FAT) in AI systems. However, the FAT approach fails to identify structural socioeconomic indicators to contextualise the principles of non-discrimination within systems design.³⁸

It has been argued that multilateral commitments to universally agreed human rights principles with regard to AI would serve to strengthen the intended application of both these approaches.³⁹ However, all approaches must be accompanied with evidence-based case studies to develop principled processes like algorithm impact assessments, explainability, transparency of commercial contracts, etc. with a clear understanding of learnings from use cases, and the role of different stakeholders within the process, rather than principled outcomes like trustworthy AI and fair and ethical machine-learning systems.

Action steps

The following advocacy steps are suggested for India:

- **Risk sandboxing:** Regulatory sandboxing⁴⁰ and data sandboxing⁴¹ 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. Risk sandboxing is envisaged as a natural progression from regulatory sandboxing in which the product is tested for its decision-making impact on vulnerable and marginalised populations on the basis of non-discrimination principles.

- **First stage process-based transparency:** While there has been much discussion about the need for explainable AI to counter the black-boxing phenomenon underlying AI’s opaque decision-making process, there needs to be a first-level transparency with respect to the inputs into the model development process, data choices, and platform capabilities and jurisdictions.
- **Disclosure of service agreements:** There need to be disclosure of service agreements within PPPs deploying AI technologies to understand the data protection commitments and data-sharing practices.
- **Mapping contextual parameters of knowledge used in modelling:** Studies that constitute knowledge about a given subject area are the result of divergent research objectives which should be evaluated for their relevance and bearing to the machine-learning system being deployed before they are factored into predictive modelling. Moreover, socioeconomic and structural indicators – such as caste, gender and family income in conjunction with how that caste group fares overall in the economy – must be identified and mapped into the model along with transparency on the decision making that maps these indicators to train the machine learning system.
- **Representative data choices:** The data on which a machine-learning system is trained must be representative of the population in which it is to be deployed.
- **Recommendations structured on non-discrimination principles:** Recommendations must be structured on non-discrimination principles. This should be done, for example, to avoid instances when an AI system recommends fewer STEM (science, technology, engineering and mathematics) courses for women because the data shows women are less likely to take up STEM subjects.

37 ARTICLE 19. (2019). *Governance with teeth: How human rights can strengthen FAT and ethics initiatives on artificial intelligence*. London: ARTICLE 19. https://www.article19.org/wp-content/uploads/2019/04/Governance-with-teeth_A19_April_2019.pdf

38 Ibid.

39 Ibid.

40 Regulatory sandboxing allows for a controlled environment with relaxed regulations that allows a product or innovation to be thoroughly tested out before being released for public use. It involves a set of rules that allow innovators to test their products within a limited legal environment subject to pre-defined restrictions like limitation on exposure, time-limited testing, pre-defined exemptions, and testing under regulatory supervision. Source: <https://cis-india.org/internet-governance/files/ai-in-india-a-policy-agenda/view>; pubdocs.worldbank.org/en/770171476811898530/Session-4-Pavel-Shoust-Regulatory-Sandboxes-21-09-2016.pdf

41 Data sandboxes allow companies to access large anonymised data sets under controlled circumstances to enable them to test their products and innovations while keeping in mind privacy and security compliance requirements.

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

