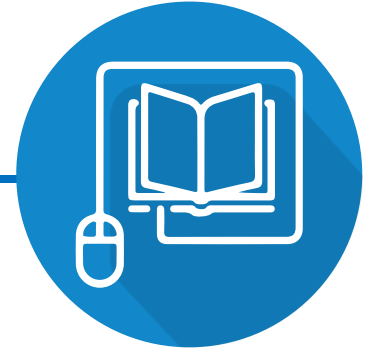


Unite Brief 2023(8)



Towards a More Comprehensive AI Ethics: How Global South Perspectives Can Enrich Current Approaches to AI Governance

Author(s): Dr. Scott Timcke, Research ICT Africa and Theresa Schültken, University Potsdam

Email: stimcke@researchictafrica.net; schueltken@uni-potsdam.de

Publishing Date: October 2023

Peer reviewed by: Karen Bett, Global Partnership for Sustainable Development and Angeline Wairegi, Strathmore University

Contents

ABSTRACT	2
1. INTRODUCTION	3
2. CURRENT APPROACHES TO AI ETHICS IN THE GLOBAL NORTH	4
3. THE SLOW VIOLENCE OF ARTIFICIAL INTELLIGENCE	5
4. GOING BEYOND AI ETHICS GUIDELINES	7
BIOGRAPHICAL NOTE	8
ACKNOWLEDGEMENTS	8
REFERENCES	8

Abstract

Artificial intelligence (AI) ethics aim to help technicians and policy-makers reduce potential negative impacts. However, the discourse is missing important contributions from the Global South. Specifically, efforts to understand how and why AI is embedded in broader power structures and is reproducing existing inequalities are incomplete. This brief addresses the issue of AI Ethics with special attention of how AI impacts inequality and vulnerable communities globally. The brief discusses the prevailing conceptualization of AI Ethics from the Global North and juxtaposes them against contributions from the Global South. It is clear that the discourse can be enriched by merging those conceptions to adopt a broader perspective on AI Ethics. In conclusion, this project, by closely examining power structures and their implications for ethical AI, sheds new light on how and where the field of AI Ethics needs to broaden its scope to fulfill its great promise.

Keywords:

Artificial Intelligence, Inequality, AI Ethics, International perspectives, Impact of AI

Disclaimer: Unless expressly stated otherwise, the findings, interpretations and conclusions expressed in this material are those of the various United Nations staff members, consultants, and advisers of the United Nations Secretariat or external contributors who prepared the works and do not necessarily represent the views of the United Nations or its Member States.

1. Introduction

This brief covers several emerging concerns about Artificial Intelligence (AI) systems in the international order while surveying approaches required to govern issues related to these technologies in a way that produces more just outcomes. Given the recent history of colonial occupation and neocolonial subordination, assessing the impact and uses of AI from the perspective of the Global South is a prerequisite for any meaningful AI ethics paradigm. With the rise of ‘big data for development’ (Hilbert, 2015) development agencies and international governance institutions must continually question whether new AI and data infrastructures enable asymmetric relations of power, visibility, and traceability, or how these technologies can be harnessed to advance more emancipatory projects.

Depending on context, the definition of AI can be broad and varied. For example, ethicist Vincent Müller states that AI refers to “any kind of artificial computational system that shows intelligent behaviour” (Müller, 2020, 2). The computer scientists Stuart Russell and Peter Norvig provide a classic definition of AI that considers computer programs intelligent when they follow “general principles of rational agents” (Russell and Norvig, 2010, 5). These ‘general principles’ are products of human programmers in an organizational setting, meaning that AI can subtly (or glaringly) reflect our human assumptions, bias, oversights and beliefs (Baulawini and Gebru 2018, Gebru, et al 2018, Hao 2019, Crawford and Paglen 2019). The result is that these technologies cannot be politically neutral (Timcke, 2021a).

Discourses about AI in the Global North tend to consistently orbit issues like privacy or designing trusted systems for healthcare and smart cities. While there are general risks around espionage as well as cybercrime, typically in the private sector positive connotations are attached to AI systems and the benefits they bring to society (see Manyika and Bughin 2018). Political representatives are wary about surveillance, automation, the future of work, and prospects of mass unemployment. These kinds of concerns can be seen in legislative initiatives like the European Union’s Digital Services Act and the Digital Markets Act.

Discussions in the Global South about AI are more circumspect. There is relatively more attention given to the social costs resulting from AI systems by academia and civil society organizations. Governments and the private sector tend to overlook conversations about social costs. For example, researchers from the Global South are very attentive to issues of colonialism and imperialism (Kwet 2019, Birhane 2020, Timcke 2021b). Similarly, while AI is said to be on the precipice of revolutionizing precision agricultural production in the Global South (Ampatzidis, Partel and Costa 2020), there are risks that these systems may damage land systems (Ryan 2019).

The reproduction of global inequalities or how unequal technological exchange hinders the Global South from designing AI systems are two examples of economic stratification *between* the Global North and Global South (see Alonso, Kothari and Rehman 2020). Then there is also the issue of how there may be marginalization and effective disenfranchisement *within* countries in the Global South as the economically disadvantaged are unlikely to have the same kind of influence over AI as the wealthy (see Timcke 2021a, chapter 1; Timcke 2023). In short, there are reasons to feel anxious about whether and how AI may continue the exploitation of (the economically disadvantaged in) the Global South thereby perpetuating economic stratifications.

This brief shows how a refocus on perspectives of the Global South enriches the understanding of the impact of AI in our societies. This helps us to refrain from taking up a limited perspective on single applications and instead guides us to core questions of power and inequalities and how to reduce those when talking about AI.

2. Current Approaches to AI Ethics in the Global North

Researchers in the Global North do give attention to the possible negative effects of AI. Numerous guidelines have been drafted to address the problems with the use of AI and outline the responsibility of the technology sector (Jobin, Ienca and Vayena, 2019). Key areas of focus are privacy, data-bias and their harmful outcomes as well as opacity in AI systems (Müller, 2020).

Privacy concerns relate to the question how personal data is used to train AI systems (Müller, 2020). The discussion of bias considers that data-based decisions might lead to racism or discrimination of marginalized people. These concerns relate to questions of opacity, which describes the problem that people affected by an AI based decision are unable to understand how the system arrives at that specific conclusion (Müller, 2020). Currently, these questions are mainly addressed through technical solutions and guidelines provided by technology companies for their AI services.

The principles regularly addressed in those AI Ethics guidelines are accountability, privacy, and fairness (Hagendorff, 2020) sometimes extended by non-maleficence and responsibility (Jobin, Ienca and Vayena, 2019). These abstract principles are intended to guide the technical work in the development and training of AI models. In general, the AI Ethics principles focus on problems for which a technical solution already exists or could be developed (Hagendorff, 2020). The stated goal is, that, with the implementation of those principles, AI systems have less negative impact on societies and marginalized people. But current research points out that these principles address superficial concerns while stalling legislation (Hagendorff, 2020).

There are currently two main trends in AI Ethics visible in the Global North:

- First, there is a heavy actor-focused narrative that prioritizes the application of specific AI-Systems and the responsibility to create those applications in an ethical way. This narrow conceptualization allows companies to address AI systems with a narrow set of ethical principles, prioritizing those principles where a technical solution is available.
- Second, only the development stage of an AI lifecycle is addressed by these guidelines, limiting the scope of AI Ethics. Instead, Crawford suggests taking up a holistic view of an “AI system from birth to death” (Crawford, 2021) to assess the full AI lifecycle. It should encompass more than just the actual development of AI, but instead start early and include, for example, the extraction of the necessary raw materials as well as data sets and, at a later stage in the lifecycle, analyze the impact of AI on society through surveillance and classification.

These narrow conceptualizations of AI Ethics do not actively engage with the various concerns raised by scholars from the Global South about the impacts of AI, as well as the connection between human-machine interaction.

3. The Slow Violence of Artificial Intelligence

More critical perspectives on the problems AI poses to societies are brought up by researchers from the Global South. These researchers challenge the existing paradigms by incorporating questions of power, context-variation and the notion of extractive practices causing harm in societies. For example, as most countries in the Global South import AI technology and export their data, there are concerns about regulatory reach when it comes to enforcing domestic determinations about privacy right. These issues are by no means only relevant to the Global South but do help frame the understanding of how AI shapes societies and the long-term implications these technologies might have.

The concept of 'slow violence' (Nixon, 2011) is useful to better conceptualize the impact of AI, and relatedly how and why the Global South is excluded from setting the agenda around AI Ethics. Slow violence takes place over an extended period meaning that the processes are often hidden from our attention because of a lack of spectacle. Still, these processes pose a significant threat to societies in the Global South given their experience of colonial occupation and the difficulty in overtly recognising the harms AI poses. Just as slow violence is invisible, so are the communities most affected (Nixon, 2011): It is apparent that the current global dialogue on AI, including the dialogue on AI ethics, does not adequately consider the voices of the Global South (Ricaurte, 2022). But attention to these hidden impacts and the voices of the Global South is needed to make AI ethics more comprehensive.

It needs to be acknowledged that there are a variety of different approaches to AI Ethics in the Global South. To summarize here some of the implications at a level of generality and relevance to this brief, the following trends can be identified:

- There is a strong focus on the working of AI as an extension of colonial power and the negative influence of this on the continent and how these forms of knowledge production are situated in oppressive structures (Kadiri, 2021).
- There is a heightened awareness about the educational inequalities and the problems that could arise from this lack of (technological) education when AI is rolled out on the continent (Mancilla-Caceres and Estrada-Villalta, 2022).
- There is a concern with how AI shapes the economic and especially labor situation in the countries due to possible negative consequences of automation through job losses (Noor and Manatan, 2022).

These context-sensitive issues raised on various continents can best be understood as a combination of social and technical systems. As AI systems make decisions about the present and future through classifying information and developing models from historical data, one of the main critiques of AI has been that these technologies reproduce or heighten existing inequalities involving gender, race, coloniality, class and citizenship (Davis et al., 2021, Scheuerman et al., 2021, Benjamin, 2019, Couldry and Mejias, 2021, Eubanks, 2018, Timcke, 2021a). A reproduction of outcome through AI inevitably means a reproduction of inequality. Because of global power structures AI currently leads to slow violence against already vulnerable and marginalized groups (Ricaurte, 2022). Slow violence can take the following forms:

A. The Slow violence of Data Colonialism

If railways were the 'open veins' of colonial extraction (Galeano, 1997), then AI and digital infrastructure plays a similar role in data colonialism. The use of technologies created in the Global North that are

deployed to other countries might lead to harm for local people. For example, through their global reach (and lack of adherence to local regulations) AI firms can extract value in many places while concentrating their wealth in the Global North. To cite an instance, AI datasets are often produced and trained by underpaid workers in the Global South and AI models are often tested in unregulated contexts in the Global South exposing people to potential risk (Gestoso, 2022). Moreover, due to intellectual property protections, companies in the Global South cannot create similar competing products. Monopoly market power converts into monopolies of knowledge.

B. Slow violence in the humanitarian sector

The entrance of technology companies in the humanitarian sector raises concerns about datafication and AI decision-making technologies (Madianou, 2019). Increasingly the outsourcing of refugee registration (and therefore the collection of biometrical data) to private actors provides these organizations with the means to track refugees (Madianou, 2019). These developments raise concerns about the use of these data sets for various other purposes and how this securitization is affecting the most vulnerable populations. In a comparable manner the discourse about data for development and the focus on positive influences might have disguised how technology companies are increasingly trying to reconfigure the social domain into a commercial revenue (Magalhães and Couldry, 2021).

More generally, a paradigm shift in the humanitarian sector can be observed through the uptake of AI: Organizations are trying to move from reacting to crises to trying to predict them. AI could be able to predict different situations and enable the humanitarian sector to act before the predicted situation occurs, be it regarding a natural disaster (Van den Homberg, Gevaert, & Georgiadou, 2020) or the occurrence of migration flows (Beduschi, 2021).

C. Slow violence in the welfare state

The administration of development aid involves 'welfare surveillance' through the creation of verifiable identities and detailed individual records (Webster, 2012). These records can be used to force a distinction between the deserving and undeserving poor or be used to detect fraud over relatively small amounts of money (Dobson, 2019, Alston, 2019). Programs to govern the poor subject these populations to high levels of surveillance in ways that reproduce their marginalization.

It is useful to recognise that AI shares features with its bureaucratic predecessors when assessing conditions of poverty. For example, when government systems distinguish between different groups of people, this social sorting can reinforce their marginalization (Lyon, 2007). If this trend continues, it might have implications for the allocation and dispersion of international aid to countries or agencies in the Global South. AI systems may be designed to excessively audit programs at the expense of allowing locals to use discretion to respond to matters arising.

D. Slow violence through datafication

Datafication describes two current trends in our societies. Firstly, it describes that data is seen as given while failing to recognize that it is socially constructed and can only quantify a certain 'rational' type of

knowledge (Magalhães and Couldry, 2021). This must be seen in light of colonial practices of devaluing knowledge from the Global South which is continued by the current focus on data for AI (Milan & Tréré, 2019). Secondly, it describes an extractive process that tries to commodify social relations, public procedures and people (Ricaurte, 2022). Together this leads to a different understanding of people's self and their independent decision-making. In general, AI systems - because they are perceived as rational - have a source of legitimacy and it has been shown that people tend to adjust their own reasoning to the AI decision (Rik, 2020) rather than questioning it. Moreover, the commodification of society might "lead to the isolation and dehumanization of individuals" (Mhlambi, 2020).

Situating the various AI Ethic considerations brought up by Global South scholars in the framework of slow violence provides us with the means to understand how the hidden long-term effects must be taken into consideration while discussing the future of AI Ethics. This is important given how research trends in the Global North can by-pass the direct experience of data-colonialism. Negative impacts of the long-term effects of AI take place both in Global South and in Global North, exacerbating inequalities and harming people.

4. Going Beyond AI Ethics Guidelines

In general, the attention in the Global North provides a too narrow understanding of AI Ethics with many blindspots (Hagendorff, 2022). Current AI Ethics guidelines are limiting the scope of AI Ethics to the working of single applications. But this is incomplete as the current approach is blind to the long-term impacts of AI on societies. Additionally, most of the notable Ethical AI frameworks in circulation were developed in the Global North without any meaningful and substantive input from stakeholders from the Global South.

By contrast, research from the Global South provides an understanding of AI that is grounded in the impact of technology on communities, keeping in mind the role of power. This is an enriched perspective which focuses on the 'slow violence of AI' which may threaten marginalized societies especially because digitisation processes unfold over a long period and are often hidden.

AI Ethicists must be aware of the processes of slow violences, using this perspective to identify potential threats that AI may pose and actively incorporate perspectives of the Global South. Without taking history into account, current AI Ethics guidelines like FAIR (Findable, Accessible, Interoperable and Reusable) are less suitable for preventing global inequalities and stratification. Additionally, the belief that AI risks could be audited and removed during production comes from the institutions invested in using AI (Vecchione et al., 2021). There are entrenched economic interests in reformism, which in turn produce a set of more ethical, fair, or inclusive precepts, but which result in normalizing existing power structures and perpetuating 'data violence' (Hoffman, 2021, Stark et al., 2021).

The Global South research agenda provides policymakers with an understanding that is based on the overall impact of AI, considers the impact of power structures, and aims to prevent exploitative relationships. This research agenda asks policy makers to consider what conditions are required to harness and channel democratic will in science, technology, and innovation policy thus building an extended AI Ethics that positively impacts societies and actively reduces inequalities. The remaining task is finding ways and means for this research agenda – and the findings it produces – to become integrated with the ethical guidelines being developed by international institutions, then having countries in the Global North adopt those same guidelines in their regulatory frameworks.

Biographical Note

Scott Timcke (PhD, Simon Fraser) is a political economist whose primary area of expertise is in democratic development policy and the role of the state in promoting egalitarian development. He leads the AI & Democracy project at Research ICT Africa. He can be contacted at stimcke@researchictafrica.net.

Theresa Schültken (BA, Heinrich-Heine-University) is a political scientist interested in the intersection of geopolitics and digital politics and the international governance of technologies such as AI or digital data. She worked on this brief during her time as an AI Research Intern at Research ICT Africa in 2023. She is currently a student of International Relations at the University Potsdam. She can be contacted at schueltken@uni-potsdam.de.

Acknowledgements

Scott Timcke and Theresa Schültken wish to thank the Research ICT Africa team for the valuable insights, especially Mark Gaffley for his comments. They also wish to thank the reviewers for their time and attention. Theresa Schültken wishes to thank Scott Timcke for his continued support and guidance.

References

- Alonso, C., Kothari, S., and Rehman, S. (2020) How Artificial Intelligence Could Widen the Gap Between Rich and Poor Nations, IMF, 2 December 2020, <https://www.imf.org/en/Blogs/Articles/2020/12/02/blog-how-artificial-intelligence-could-widen-the-gap-between-rich-and-poor-nations>.
- Alston, P., (2019) "Digital technology, social protection and human rights". *Office of the United Nations High Commissioner for Human Rights*, Available at: <https://www.ohchr.org/en/calls-for-input/digital-technology-social-protection-and-human-rights-report> (Accessed: 17 February 2023)
- Ampatzidis, Y., Partel, V., & Costa, L. (2020). Agroviz: Cloud-based application to process, analyze and visualize UAV-collected data for precision agriculture applications utilizing artificial intelligence. *Computers and Electronics in Agriculture*, 174, <https://www.sciencedirect.com/science/article/pii/S0168169920304695>.
- Beduschi, A. (2021). International migration management in the age of artificial intelligence. *Migration Studies*, 9(3), 576–596.
- Benjamin, R. (2019) *Race After Technology: Abolitionist Tools for the New Jim Code*. Cambridge: Polity Press.
- Birhane, A. (2020) Algorithmic Colonization of Africa, *SCRIPTed* 17(2) available at <https://scripted.org/article/algorithmic-colonization-of-africa/>
- Buolamwini, J., Gebu, T. (2018) "Gender shades: Intersectional accuracy disparities in commercial gender classification" *Proceedings of Machine Learning Research* 81 (2018), 1--15; <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>.

- Couldry, N., and Mejias, U. A. (2021) "The decolonial turn in data and technology research: What is at stake and where is it heading?". *Information, Communication & Society*, 1–17.
- Crawford, K. (2021) *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.
- Crawford, K., Paglen, T., (2019) Excavating AI: The politics of images in machine learning training sets. The AI Now Institute, New York University, 2019; <https://www.excavating.ai>
- Davis, J. L., Williams, A., & Yang, M. W. (2021) "Algorithmic Reparation". *Big Data & Society*, 8(2), 1–12.
- Dobson, K. (2019) "Welfare Fraud 2.0? Using Big Data to Surveil, Stigmatize, and Criminalize the Poor". *Canadian Journal of Communication*, 44(3), 331–342.
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.
- Galeano, E. H. (1997) *Open Veins Of Latin America: Five Centuries Of The Pillage Of A Continent*, New York: Monthly Review Press.
- Gebu, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., Crawford, K. (2018) Datasheets for datasets. In Proceedings of the 5th Workshop on Fairness, Accountability, and Transparency in Machine Learning, 2018; <https://www.microsoft.com/en-us/research/uploads/prod/2019/01/1803.09010.pdf>
- Gestoso, P. (2022) *How artificial intelligence is recolonising the Global South*, Available at: <https://www.themintmagazine.com/how-artificial-intelligence-is-recolonising-the-global-south/> (Accessed on: 17. February 2023).
- Hagendorff, T. (2020) "The Ethics of AI Ethics: An Evaluation of Guidelines". *Minds & Machines*, 30 (1), 99–120.
- Hagendorff, T. (2022) "Blind spots in AI ethics". *AI and Ethics*, 2(4), 851–67.
- Hao, K., (2019) This is how AI bias really happens---and why it's so hard to fix. MIT Technology Review; <https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/>.
- Hilbert, M. (2015) "Big Data for Development: A Review of Promises and Challenges". *Development Policy Review*, 34(1), 135-74.
- Hoffmann, A. L. (2021). "Terms of inclusion: Data, discourse, violence". *New Media & Society*, 23(12), 3539–3556.
- Jobin, A., Ienca, M. and Vayena, E. (2019) "The global landscape of AI ethics guidelines". *Nature Machine Intelligence* 1(9), 389–99.
- Kadiri, A. (2021) "Data and Afrofuturism: an emancipated subject?" *Internet Policy Review*, 10(4), 1-26.
- Kwet, M. (2019). Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class*, 60(4), 3–26.
- Lyon, D. (2007). *Surveillance Studies: An Overview*. Cambridge: Polity Press.
- Madianou, Mirca (2019) "Technocolonialism: Digital Innovation and Data Practices in the Humanitarian Response to Refugee Crises". *Social Media + Society*, 5(3), 1-13.

- Magalhães, J. C. and Couldry, N. (2021) "Giving by Taking Away: Big Tech, Data Colonialism, and the Reconfiguration of Social Good". *International Journal of Communication*, 15, 343-62.
- Mancilla-Caceres, J. F. and Estrada-Villalta, S. (2022) "The Ethical Considerations of AI in Latin America". *Digital Society*, 1(2), 1–6.
- Manyika, J., and Bughin, J., (2018) The promise and challenge of the age of artificial intelligence, McKinsey Global Institute, 15 October 2018, <https://www.mckinsey.com/featured-insights/artificial-intelligence/the-promise-and-challenge-of-the-age-of-artificial-intelligence>.
- Mhlambi, S. (2020) "From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance". *Carr Center Discussion Paper Series*.
- Milan, S., & Tréré, E. (2019). Big Data from the South(s): Beyond Data Universalism. *Television & New Media*, 20(4), 319–335.
- Müller, V. C. (2020) "Ethics of Artificial Intelligence and Robotics". Available at: <https://plato.stanford.edu/entries/ethics-ai/> (Accessed on 14. February 2023).
- Nixon, R. (2011) *Slow violence and the environmentalism of the poor*. Cambridge, Mass.: Harvard University Press.
- Noor, E. and Manatan, M. B. (2022) "Raising Standards. Data and Artificial Intelligence in South East Asia". Available at: <https://asiasociety.org/policy-institute/raising-standards-data-ai-southeast-asia> (Accessed on 14. February 2023).
- Ricourte, P. (2022) "Ethics for the majority world: AI and the question of violence at scale". *Media, Culture & Society*, 44(4), 726–45.
- Rik, P. (2020) "The agency of algorithms: Understanding human-algorithm interaction in administrative decision-making". *Information Polity*, 25(4), 507-522.
- Russell, S. and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*, 3rd edition. New Jersey: Prentice Hall.
- Ryan, M. (2019). Ethics of using AI and big data in agriculture: The case of a large agriculture multinational. *ORBIT Journal*, 2(2) pp1-27, <https://www.sciencedirect.com/science/article/pii/S2515856220300110>.
- Scheuerman, M. K., Pape, M. and Hanna, A. (2021) "Auto-essentialization: Gender in automated facial analysis as extended colonial project". *Big Data & Society*, 8(2), 1–15.
- Stark, L., Greene, D. and Hoffmann, A. L. (2021) "Critical Perspectives on Governance Mechanisms for AI/ML Systems", in Roberge, J. and Castelle, M. (eds.) *The Cultural Life of Machine Learning: An Incursion into Critical AI Studies*. Cham: Palgrave Macmillan, 257–80.
- Timcke, S. (2021a) *Algorithms and the End of Politics: How Technology Shapes 21st-Century American Life*. Bristol: Bristol University Press.
- Timcke, S. (2021b) "Kwame Nkrumah and Imperialist Finance in Africa Today", *Review of African Political Economy*, 6 July 2021, <https://roape.net/2021/07/06/kwame-nkrumah-and-imperialist-finance-in-africa-today/>.
- Timcke, S. (2023) *The Political Economy of Fortune and Misfortune*, Bristol: Bristol University Press.
- Van den Homberg, M. J. C., Gevaert, C. M., & Georgiadou, Y. (2020). The Changing Face of Accountability in Humanitarianism: Using Artificial Intelligence for Anticipatory Action. *Politics and Governance*, 8(4), 456–467.

Vecchione, B., Levy, K. and Barocas, S. (2021) “Algorithmic Auditing and Social Justice: Lessons from the History of Audit Studies”. *Equity and Access in Algorithms, Mechanisms, and Optimization*, 1-9.

Webster, C. W. R. (2012) “Public administration as surveillance”, in Ball, K., Haggerty, K. and Lyon, D. (eds.), *Routledge handbook of surveillance studies*. London: Routledge, 313–20.