COMMENT

<https://doi.org/10.1057/s41599-021-00910-x>

OPEN

Big data for whose sake? Governing migration through artificial intelligence

Tuba Bircan ¹✉ & Emre Eren Korkmaz ²

Although human activity constantly generates massive amounts of data, these data can only be analysed by mainly the private sector and governmental institutes due to data accessibility restrictions. However, neither migrants (as the producers of this data) nor migration scholars (as scientific experts on the topic) are in a position to monitor or control how governments and corporations use such data. Big Data analytics and Artificial Intelligence (AI) technologies are promoted as cutting-edge solutions to ongoing and emerging social, economic and governance challenges. Meanwhile, states increasingly rely on digital and frontier technologies to manage borders and control migratory movements, and the defence industry and military-intelligence sectors provide high-tech tools to support these efforts. Worryingly, during the design and testing of algorithmic tools, migrants are often portrayed as a security threat instead of human beings with fundamental rights and liberties. Thus, privacy, data protection, and confidentiality issues continue to pose risks and challenges to migrant communities and raise important questions for the public and decision-makers alike. This comment seeks to shed light on the lack of effective regulation of AI and Big Data as they are applied in migration 'management'. Additionally, from the perspective of privacy issues and immigrant rights (seeking asylum as a human right, it aims at advocating improved access to Big Data for scientific research which might act as a social control function for the smart border and existing/ongoing migration governance practices of countries. We argue that the use of Big Data and AI for migration governance requires much better collaboration between migrants (including the civil society and grassroots organisations solidarity that represent them), data scientists, migration scholars and policymakers if the potential of these technologies is to be reached in a way that is reasonable and ethical. Numerous critical privacy questions arise are regarding the legal requirements, confidentiality, and rules of engagement as well as the ethical concerns of (mis)use of new technologies. When the secretive nature of the ongoing exploitation of migrant data by states and corporations is considered raising such questions is essential for progress.

¹Vrije Universiteit Brussel, Brussels, Belgium. ²Oxford University, Oxford, UK. ✉email: tuba.bircan@vub.be

Hope-mongering AI: The promise and perils

Although human activity constantly generates massive amounts of data, these data can only be analysed by mainly the private sector and governmental institutes due to data accessibility restrictions. Big Data analytics and Artificial Intelligence (AI) technologies are promoted as cutting-edge solutions to ongoing and emerging social, economic and governance challenges. Such analysis serves several practical purposes, including marketing goods and services, ensuring the success of political campaigns, and identifying criminal activities. New methods of harnessing Big Data to serve the public good are promising but as yet not widespread.

While recent developments in frontier technologies have brought many opportunities, they also raise significant concerns and challenges for societies. Giant steps in AI, blockchain, robotics and Big Data have had tremendous impacts on the socio-economic development of countries (Palvia et al., 2018) as well as on the daily lives of people. However, they have also destabilised traditional approaches to work and life, not to mention individuals' so-called 'algorithmic identities' (Cheney-Lippold, 2011) and their sense of self and the future (Floridi, 2007).

Governments are investing vast sums in developing new technologies, and tech companies are at the forefront of commercialising various tech-based solutions to meet consumer demand. However, it is not difficult to observe growing concerns about the challenges arising from these new frontier technologies. The general optimism in the first decade of the twenty-first century has turned into a more pessimistic outlook due to the actual and potential outcomes that arise when corporations and governments deploy these new technologies. Today, 'Big Tech' increasingly serves as a synonym for 'Big Brother', and these firms stand accused of being monopolistic predators hungry for ever more consumer data. The result is 'surveillance capitalism', in which consumers divulge their most intimate selves in exchange for convenience and the tantalising prospect of previously undreamed-of goods and services (Zuboff, 2019). Therefore, such a critical change in the public mood around these technologies has shifted the frame from a utopian vision of human progress to a picture of an emerging dystopian tomorrow in which humankind is subsumed into the digital realm. Technology offers even more ways to manipulate and control the citizen-as-consumer with the help of non-human 'smart products' (Meier, 2011). Thus, it has moved to the centre of the debate over the growing power of the dominant forces in society, supplementing and indeed catalysing their existing tools of hegemony, control, and oppression.

Given the apparent power imbalances, technology remains an arena of political struggle. The same technologies can be used to promote peace and spark conflict. For this reason, the political struggle to hold corporations and governments accountable can still change the course of technological development. For instance, anxiety over fake news and manipulative campaigns undermining democracy and the trustworthiness of elections is spurring many NGOs and political movements to demand that digital campaigns observe democratic principles and to hold political parties' feet to the fire for their social media advertising.

In sum, various means to challenge the dystopian vision of the future of humanity remain. Many organisations in democratic societies retain in their grasp the tools to hold governments and technology companies to account for the direction that technological change takes. Even so, migrants and refugees lack the same access citizens have to the arena of political contestation to protect and advance their interests. For anyone wishing to observe the likely shape of a dystopian future of a tech-oriented society in which the people have ceded their political autonomy, a close look at the daily experiences of migrants and refugees vis-à-vis the large tech providers offers sufficient clues.

AI at the border

The study of human migration has recently taken a giant step forward through the use of Big Data and AI. Numerous states and international organisations involved in migration management have turned to these new technologies to enhance their ability to govern border spaces and control movement across them. The reasoning has centred on the apparent payoffs of increased efficiency, improved preparedness, and enhanced border security. Specific applications that have been touted include the ability to predict population movements, process administration, and border checks and prevent unauthorised migration. Moreover, automated decision-making in visa and asylum applications and AI lie detectors for immigrants and refugees encroach on human rights owing to limited or no international regulation. These aspects spearhead the controversial discussions.

Big Data analysis is also providing social scientists with new insights in the field of migration research. Such analysis enables the study of temporary or circular migration patterns and real-time monitoring of public opinion and media discourse on migration. It also sheds much-needed light on aspects of migration we currently have limited knowledge about, such as the integration prospects of recently arrived migrants and the possible future migration patterns. Moreover, given the new knowledge generated from migrants' data has the potential to be exploited by states and corporations alike, migration scholars must not lose sight of the principles of ethical research and should retain a critical approach to the practices of these powerful actors.

Big Data projects in migration/border management can be divided into two main categories. In the first category, the main aim is to analyse Big Data to provide faster and more efficient services. Projects in the second category are piloting Big Data analysis for speculative targets, such as AI lie detection or automating asylum or visa decision-making processes. Distinguishing these two categories is significant as they pose different kinds of problems and require different regulatory tools.

One of the first AI applications in the first category was initiated in 2007 by the Hong Kong Immigration Department as a part of the eBrain project. The agency employed a suite of AI technologies to streamline administrative tasks like visa applications and processing travel documents, identity cards, and work permits. In mid-2017, the United Nations initiated the Unite Ideas Internal Displacement Event Tagging and Extraction Clustering Tool challenge. The winner was the Data for Democracy team, which built a tool capable of tracking and analysing refugees and other people forced to flee or evacuate their homes.

The European Asylum Support Office (EASO) uses machine learning to predict pressures on the asylum administrations of member states of the European Union and associated countries (i.e., Norway and Switzerland). To do so, EASO builds on three types of data collected on past events—namely, data from traditional countries of origin and transit (including social media monitoring), data on pressures at the EU's external borders, and data on the outcomes of previous asylum applications in the EU. EASO's algorithm predicts pressures up to four weeks in advance and suggests possible future medium-term scenarios using historical and current data.¹ The European Data Protection Supervisor (EDPS) offered a sharp critique of the EASO's social media monitoring of migrants and refugees:² 'Social media users monitoring is a personal data processing activity that puts individuals' rights and freedoms at significant risk. It involves uses of personal data that go against or beyond individuals' reasonable expectations. Such uses often result in personal data being used beyond their initial purpose, their initial context and in ways the individual could not reasonably anticipate'. The EDPS underscored the important principles of 'purpose limitation' and 'data

minimisation', whereby personal data should be collected only for 'specified, explicit and legitimate purposes. The EDPS' warning also dovetails with the aims of this comment because as the public discusses the data mining and surveillance techniques to predict, manage and stop migratory movements, such legal and social pressure over official institutions and corporations will likely increase.

Such Big Data monitoring projects are not only deployed by EU agencies. Similar approaches have been observed in border security and migration management projects in the US and Canada. For instance, US Customs and Border Protection (CBP) and Immigration and Customs Enforcement (ICE) buy commercial databases to track mobile phones to identify undocumented immigrants. It is also reported that from 2010 to 2014, CBP spent about \$2.5 million to purchase cell-site simulator technology to develop fake towers that can detect and intercept mobile phone text and voice messages and pull the location and other information from mobile devices that are trying to connect to it (Ghaffary, 2020; Riotta, 2020). Also, as Akhmetova (2020) notes, 'In 2018, the Canada Border Services Agency used private third-party DNA services such as Ancestry.com to establish the nationality of individuals subject to potential deportation. This is deeply concerning because one's DNA is not related to immigration journey, 'legality', nationality and should bear no impact on one's immigration/asylum applications. Another issue is the coercive nature of privacy invasion—individuals who gave their DNA samples to these companies might not have given consent nor knew that their data could be used by governments to assess immigration applications'.

Issues also arise with the second category, speculative Big Data pilot projects. For example, in 2018, Canada's Immigration Department ran a pilot project (Artificial Intelligence Solution) to assess AI-supported decision-making for immigration and asylum applications. Agencies within the US Department of Homeland Security also use new technologies to automate migration-related and asylum-related decisions. However, human rights experts have criticised this approach for using vulnerable migrants and asylum-seekers as experimental subjects to train AI algorithms (Molnar, 2019).

The European Commission funded projects using AI to develop a lie-detection system to ramp up security at European borders (iBorderCtrl from 2016–2019)³ and create autonomous border surveillance systems (ROBORDER from 2017 to 2021).⁴ Moreover, in 2018, the mandate and operation areas of the European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice (eu-LISA) were expanded⁵ with a particular focus on the implementation of the EU's asylum, border management and migration policies. The values of eu-LISA's investments are identified as accountability, transparency, excellence, continuity, teamwork, and customer focus, which privilege the 'customer' (i.e., states and/or agencies) rather than migrants and fall well short on privacy and human rights aspects.

UN agencies and NGOs also collect data from migrants and refugees. The data is necessary to deliver services and make long-term plans. However, concerns arise about how such data is collected, where it is stored, whether and how people in vital need of assistance give their consent, and what measures are taken to prevent people's data from being used against them. Moreover, the challenges go beyond discussions of consent since user privacy and secondary use are central issues as well.

To illustrate, migrants and refugees often find themselves providing 'informed consent' to primary data collection even though they often have insufficient understanding of the implications or potential threats. Visa and asylum applications and getting access to aid at refugee shelters where biometric

information is requested are examples of such 'consent'. Moreover, migrants and refugees may find themselves at the whim of politics through mechanisms like 'function creep', where the use of technology and information extends beyond the stated initial purpose.

Biometric surveillance and processing in public spaces (and advanced analyses in social media data) are other examples where individuals may be at risk but are given no information or insight into the collection or analysis of personal data. The current Afghan crisis offers a clear illustration of the risks and challenges here. On the one hand, we see evidence of the expanded deployment of high-tech surveillance systems by the EU and national authorities at external borders in anticipation of new migration waves, raising questions about the ethical implications of these technologies. On the other, the Taliban's potential access to the biometric data of Afghan refugees registered by humanitarian or military agencies calls urgent attention to the risks of 'function creep' in the uncertain field of migration governance.

So far, no satisfactory response has been offered to these serious political, ethical, and social questions. That this is the case reflects the minimal representation and bargaining power—and thus political muscle—of people on the move. Although citizens can be victimised by the development of new technologies—for instance, through job losses to automation or the diminution of the public sphere—they still have ways to raise their concerns and apply pressure to governments and corporations. Refugees and undocumented migrants lack these opportunities. As commentators (Bigo, 2002; Maxmen, 2019) have pointedly asked: Who will defend migrants/refugees against the power of governments and corporations, whose capabilities for surveillance and oppression have expanded significantly through the new technologies?

Big tech and migrants

In recent years, global technology companies have also begun to pay more attention to migratory movements. Various tech corporations offer services and products to UN agencies, NGOs, states and migrants/refugees on a commercial basis (i.e., not as part of corporate social responsibility projects). Financial institutions, telecom companies, mobile phone operators and tech firms seek to profit from their investments. Thus, serving refugees and displaced populations has increasingly become a profitable business. As a result of their technical superiority and financial power, commercial actors act as suppliers to UN agencies and NGOs and introduce themselves as critical stakeholders, actively working with states and the UN agencies to manage and control migratory movements. However, neither migrants (as the producers of this data) nor migration scholars (as scientific experts on the topic) are in a position to monitor or control how governments and corporations use such data.

Corporate involvement in migration, displacement and refugee issues is not limited to humanitarian emergencies, and corporations' keen interest in border management is a very concerning development. States increasingly rely on digital and frontier technologies to manage borders, and the defence industry and military-intelligence sector provide high-tech tools for this purpose. For instance, Europe's largest arms sellers also market 'smart' border management tools (Akkerman, 2016). The 'smart border' constructed by the US along the southern frontier with Mexico and the EU's digitalised border management systems (including Eurosur⁶ and Eurodac⁷) are examples of such technology in action. However, during the design and testing of algorithmic tools, migrants are often portrayed as a security threat instead of human beings bearing fundamental rights and liberties. Thus, issues surrounding privacy, data protection and

confidentiality continue to pose risks and challenges to migrant communities.

An over-reliance on AI for migration governance

Implementing technological developments such as AI algorithms for migration management rests on large volumes of data from various sources; machine learning is one of the fundamental steps in improving these algorithms. This is why AI applications are constantly in need of ever more data. Beduschi (2020) identifies three major challenges regarding the quality of the data used to train algorithms: (1) migrants' data privacy, (2) algorithmic accountability and (3) fairness. The challenges that attend the rise of AI should not be overlooked. Moreover, researchers and policymakers must avoid being distracted by the hype surrounding AI and focus on ensuring that comprehensive regulations are developed to protect the common good.

Molnar (2019) argues that states deliberately lean into the lack of regulation in this space because it renders migrants more trackable and intelligible. Against this backdrop, the use of advanced technology for surveillance and data collection are justified on national security grounds or the apparent 'objectivity' of data-driven policymaking, or even—somewhat improbably—on humanitarian or development grounds. Concerns arise as to whether these 'solutions' are being tested on refugees and migrants because they can scarcely object and often lack even basic knowledge about what is involved. Nevertheless, as mentioned, improving algorithms requires vast amounts of data from a variety of sources and governments and corporations have increasingly turned to conflict zones and refugee camps as experimental fields. As a result, technology firms are seeking to market their platforms to the EU and national governments as 'migration prediction systems' in order to boost sales and profits (Taylor and Meissner, 2020). Against this backdrop, informed consent and vulnerable groups' 'right to refuse' intrusive data-gathering techniques are increasingly overlooked in the race for ever-more efficient monitoring and 'service delivery' (as the infamous case of iris scanning of refugees in camps in Jordan bears out).

Issues surrounding privacy, data protection, and confidentiality continue to pose risks and challenges to migrants. There is no effective auditing mechanism to ensure governments and corporations are accountable for using migrants' data. In addition, several critical ethical questions arise about the legal requirements, confidentiality, and rules of engagement. These questions mostly refer to the possible (mis)use of new technologies for more conservative and preventive migration policies that violate human rights. An additional concern is the role of tech corporations as contractors for states and institutions, where data gathering for commercial purposes is combined with contracted analysis on these datasets. In other words, given the secretive way in which states and corporations exploit migrant data, which points out to a growing problem related to the exploitation of data for preventive and restrictive migration governance. The use of Big Data and AI for migration governance requires much better collaboration between migrants (including civil society and grassroots organisations that represent them), data scientists, migration scholars and policymakers if the potential of these technologies is to be reached in a way that is reasonable and ethical.

Most discussions on politics, power and AI are in-depth but not widely spread. The applications in migration governance are often built on the challenging trade-offs for societal benefits. Responsibility and accountability of the actors in AI applications linked to decision-making mechanisms are obscurely defined. Hence, who can challenge the automated decisions for migration management is shrouded in mystery. Bearing in mind that the notion of 'managing' migration is typically a euphemism for

'preventing' it, who benefits the most from the AI and high technology utilisation for managing migration remains an open question. A central issue concerns who is absent from the decision-making table—namely, scholars and data scientists but also civil society and migrants themselves. To be able to overcome the existing challenges and be prepared for future complications in global migration governance, the discussions should rely on 'human rights' and there is a need for clear identification for the stakeholders who will closely be involved with decisions and act as a controlling body.

In sum, the 'influx' of Big Data and AI applications to address societal challenges raises profound questions about the fair distribution of benefits arising from these approaches. Digital technologies and AI have the power to influence and shape democracy, as the recent misinformation campaigns in the US and many other countries lay bare. Considering the sensitivity of migration in many Western democracies, the risks of misinformation and disinformation make it all the more important that the scientific, political, and public discourse is as transparent as possible, especially since the stakeholder groups involved are often very opaque.

Received: 29 December 2020; Accepted: 27 September 2021;

Published online: 20 October 2021

Notes

- 1 See European Migration Network, 2021, "Accurate, timely, interoperable? Data management in the asylum procedure". https://ec.europa.eu/home-affairs/sites/homeaffairs/files/00_eu_emn_study_data_management_final_common_template_final_en.pdf
- 2 https://edps.europa.eu/sites/default/files/publication/19-11-12_reply_easo_ssm_final_reply_en.pdf
- 3 <https://cordis.europa.eu/project/id/700626>.
- 4 <https://roborder.eu/the-project/aims-objectives/>.
- 5 Regulation (EU) 2018/1726.
- 6 https://ec.europa.eu/home-affairs/policies/schengen-borders-and-visa/border-crossing/eurosur_en
- 7 <https://www.elulisa.europa.eu/Activities/Large-Scale-It-Systems/Eurodac>

References

- Akhmetova R (2020) Efficient discrimination: on how governments use artificial intelligence in the immigration sphere to create and fortify 'invisible border walls'. University of Oxford, COMPAS Working Paper 20–149
- Akerman M (2016) Border wars: the arms dealers profiting from Europe's refugee tragedy. Transnational Institute and Stop Wapenhandel
- Beduschi A (2020) International migration management in the age of artificial intelligence. *Migr Stud.* <https://doi.org/10.1093/migration/mnaa003>
- Bigo D (2002) Security and immigration: toward a critique of the governmentality of unease. *Alternatives* 27(1):63–92
- Cheney-Lippold J (2011) A new algorithmic identity: soft biopolitics and the modulation of control. *Theory Cult Soc* 28(6):164–81
- Floridi L (2007) A look into the future impact of ICT on our lives. *Inf Soc* 23(1):59–64
- Ghaffary S (2020). The "smarter" wall: how drones, sensors, and AI are patrolling the border. *Vox Media.* <https://www.vox.com/recode/2019/5/16/18511583/smart-border-wall-drones-sensors-ai>
- Maxmen A (2019) Can tracking people through phone-call data improve lives? *Nature* 569(7758):614–7
- Meier P (2011) Do "Liberation technologies" change the balance of power between repressive states and civil society? Unpublished PhD thesis, Faculty of the Fletcher School of Law and Diplomacy
- Molnar P (2019) Technology on the margins: AI and global migration management from a human rights perspective. *Camb Int Law J* 8(2):305–30
- Palvia P, Baqir N, Nemati H (2018) ICT for socio-economic development: a citizens' perspective. *Inf Manag* 55(2):160–76
- Riotta C (2020) Trump administration bought access to cell phone database to target immigrants, report says. Independent. <https://www.independent.co.uk/news/world/americas/trump-ice-immigration-cellphone-location-data-report-surveillance-tracking-a9324221.html>
- Taylor L, Meissner F (2020) A crisis of opportunity: market-making, big data, and the consolidation of migration as risk. *Antipode* 52(1):270–90

Zuboff S (2019) *The Age of surveillance capitalism: the fight for a human future at the new frontier of power*. Hatchett, New York

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Tuba Bircan.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021