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
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 Comninos (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 Milanes (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)

Sida

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
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.
Introduction

If the Cambridge Analytica scandal served as the public’s “great privacy awakening,”
1 for public policy experts it affirmed several troubling messages about human vulnerability given the current state of the governance of big data and artificial intelligence (AI) systems. What started as a legitimate academic research project quickly became scandalous when a British political consulting firm used the data collected – from up to 87 million people’s Facebook profiles – for a different purpose: to influence the 2016 United States (US) election through targeted political advertisements.

The data transfer occurred without consent from Facebook’s users or, arguably, even Facebook itself, reinforcing the idea that big data and AI systems pose significant threats not only to the right to privacy, but to the enjoyment of human rights and the integrity of democratic institutions.2 As the scandal unfolded, and the European Union’s General Data Protection Regulation and the Council of Europe’s modernised Convention 108+ entered into force, experts cautioned that in the absence of new approaches to data governance, even a new a “bill of data rights” could not check the power imbalances between data controllers and data subjects.3

Facebook’s refusal to attend hearings before the International Grand Committee on Big Data, Privacy and Democracy, even under subpoena, points to the urgency of finding new ways of dealing with the platform’s power.

Current approaches to data governance suffer from a lack of transparency and accountability, in part because big data – in combination with AI – continues to make end runs around consent and privacy self-management.4 AI-enabled methods of analysis can be used by companies to generate, infer and collect sensitive information about people that they have neither provided nor confirmed.5 Companies have access to an array of data collection methods, some of which circumvent consent without detection: data is now, as the Cambridge Analytica case demonstrated, extracted through online profiling, purchased from third-party brokers, or derived from aggregated data sets. The complexity and opacity of information flows make it virtually impossible for individuals to discern, much less self-manage, the risks or rights they engage when consenting to the use of their personal data.6

Another problem is that current approaches to data governance tend to concentrate data in the hands of powerful digital platforms, preventing the public from sharing in its value.7 In today’s digital society, individuals serve as the inputs to AI systems and yet they wield little control over its outputs. While exceptional, scandals like Cambridge Analytica prove the following rule: not only do current approaches to governing data exclude (most) individuals from sharing in its value, but they expose them to human rights abuses, too.

References

Accordingly, an increasing proportion of policy workshops, discussions and research over the last year have focused on designing inclusive data governance models that facilitate public accountability and promote a more equitable distribution of data’s economic value. While a number of novel approaches have been considered, proposals based on fiduciary models of data governance have garnered significant attention. This report provides a brief overview of current research and policy discussions related to two such proposals – “information fiduciaries”, which aim to improve the accountability of online platforms, and “data trusts”, a flexible governance tool that is being considered for a range of different purposes – while offering an assessment of their unique value propositions and implementation challenges.

Information fiduciaries

The nature of fiduciary relationships and the precise duties they create is a contested subject. Yet as Richard Whitt notes, “all definitions of fiduciaries share three main elements: (1) the entrustment of property or power; (2) the entrustors’ trust of fiduciaries; and (3) the risk to entrustors emanating from the entrustment.” As such, the information fiduciary proposal recommends raising the intensity of obligations owed by data controllers (i.e. companies) to data subjects (i.e. individuals) through the imposition of fiduciary duties of care, confidentiality and loyalty, which would transform large data controllers like digital platforms into “information fiduciaries”.

The intent is to correct the power imbalance between companies and individuals by giving companies duties similar to those held by doctors, lawyers and accountants towards their patients or clients.

While scholars such as Lillian Edwards and Jack M. Balkin may be credited for first considering the imposition of fiduciary obligations, albeit through different methods, their approach has faced some criticism. Whereas Edwards has suggested that fiduciary obligations are “implied” whenever a data subject shares personal data with a data controller, on account of the risk the former exposes her- or himself to, Sylvie Delacroix and Neil Lawrence argue that taking up a duty of loyalty to manage data subjects’ rights in their best interests would place data controllers in a conflict of interest with their competing duty to maximise shareholder value.

To resolve this tension, Balkin has proposed that special immunities or financial incentives could help induce data controllers to take up a limited fiduciary obligation that could be defined by statute. Several problems have been identified with this “grand bargain”. Special incentives for platforms to behave as fiduciaries may not be enough to nullify the conflict of interest between a platform’s duty of loyalty to manage data subjects’ rights and its duty of loyalty to shareholders. As Delacroix and Lawrence have put it:

[The “information fiduciary” proposed by Balkin would be placed in a position that is comparable to that of a doctor who gains a commission on particular drug prescriptions or a lawyer who uses a company to provide medical reports for his clients while owning shares in that company.]

Mike Godwin has argued that a “professional framework of fiduciary obligations for tech companies” supported by “professional codes of ethical conduct that bind the tech companies that have fiduciary duties to us” could be one way of ensuring platforms take their positions as information fiduciaries seriously.

Balkin’s recommendation to restrict digital platforms’ fiduciary duties unfairly discounts the intangible vulnerabilities of living in an online world, where privacy and human rights violations routinely go unnoticed or unchallenged. Facebook may not be a doctor or YouTube an accountant,

and yet each handles sensitive personal data that has been entrusted to them, and which can expose individuals, as entrusters, to abuse.\(^23\) Balkin’s position may be further undermined in his characterisation of data controllers’ limited duty of loyalty as a prohibition from “act[ing] like con men”\(^22\) or creating “an unreasonable risk of harm to their end user”:\(^23\) corporations already bear a duty to do no harm under tort law duty of care. This aspect of Balkin’s proposal has led Lina Khan and David Pozen to question whether Balkin’s proposal “is a fiduciary approach in any meaningful sense at all,” and, ultimately, to recommend that competition and antitrust policy may be the more productive channel to explore.\(^26\)

While the imposition of fiduciary obligations on data controllers may have been realistic in the past – when business models were less entrenched and operations less complex – today they are unlikely to broker meaningful change. For example, would a fiduciary duty model have prevented the misuse of personal data in the Cambridge Analytica scandal? Moreover, as Whitt observes, the forcible imposition of fiduciary obligations tends to produce suboptimal results, creating relationships of “grudging” care or loyalty.\(^25\) To this end, Whitt posits that the market may one day privilege companies who voluntarily compete around the self-imposition of a positive duty of loyalty.\(^26\) Yet such a prospect would likely depend on a clear signal that the market value of assuming a positive duty of loyalty outweighs the status quo, and on the emergence of alternative service providers. Absent important changes in the competitive landscape surrounding data controllers, neither of these developments is likely to occur, and the information fiduciary proposal will have limited impact on the power asymmetries embedded into current approaches to data governance or problems related to privacy self-management and data concentration. Structural problems such as these may require more than a light touch approach.

**Data trusts**

Data trusts result from the application of the common law trust to the governance of data or data rights. Trusts begin with an asset, that a “settlor” places into a trust.\(^27\) A trust charter stipulates the purpose and terms of the trust, which exists to benefit a group of people, known as the “beneficiary.” In more basic terms, a data trust creates a legal way to manage data rights for a purpose that is valuable to a beneficiary.\(^28\)

In a data trust, data subjects would be empowered to pool the rights they hold over their personal data into the legal framework of a trust.\(^29\) A trustee is appointed with a fiduciary obligation to manage the trust’s assets in accordance with the trust charter and the interests of its beneficiaries. The trustee is accountable to the beneficiaries for the management of the trust, and has a responsibility to take legal action to protect their rights.

While there is currently no common definition for data trusts, they are often described in reference to the particular problem their proposer is aiming to solve.\(^30\) As outlined below, governments have focused on the potential to use data trusts to promote data sharing and “responsible” innovation. The “civic data trust”\(^31\) has been theorised as a way to protect the public interest in data governance decision-making processes. Perhaps the most expansive vision for data trusts at scale is the concept of the “bottom-up data trust”,\(^32\) which has been proposed as a way to return the power that stems from aggregated data to individuals.

To be sure, data trusts are not a governance model in and of themselves, and their effectiveness will depend on the complementary use of other tools that constitute good practice in corporate governance. Rather, data trusts provide a flexible framework that is capable of balancing a constellation of different interests or rights associated with a range of different stakeholders and use cases.

**Data trusts as data-sharing vehicles**

Governments’ interest in data trusts has primarily focused on their potential to facilitate responsible data sharing and innovation in the AI sector. The following is a list of such proposals.

- In 2017, an independent review commissioned by the United Kingdom (UK) recommended “data

---


\(^{23}\) Ibid.


\(^{26}\) Ibid.


trusts” as a way to “share data in a fair, safe and equitable way,” adding that they would likely play an important role in growing the AI sector.33

• In 2018, the Open Data Institute (ODI) announced a partnership with the UK Office for Artificial Intelligence and Innovate UK to run three data trust pilots focusing on tackling illegal wildlife trade, reducing food waste and improving municipal public services.34

• In May 2019, the Canadian government announced a new Digital Charter35 that referenced data trusts as a possible way to facilitate data sharing in a privacy- and security-enhancing manner for research and development purposes in areas such as health, clean technology or agribusiness. The Canadian government also included several recommendations related to data trusts in a discussion paper36 that accompanied the Digital Charter, outlining proposals for the reform of Canada’s federal privacy legislation.

• In May 2019, the Organisation for Economic Co-operation and Development adopted the OECD Principles on Artificial Intelligence (the “OECD Principles”), which recommend data trusts as a way to support the safe, fair, legal and ethical sharing of data.37

• In June 2019, the G20 Digital Economy Ministers incorporated the OECD’s recommendation on data trusts into their “human-centred AI Principles”.38

• Finally, that same month, the European Union High-Level Expert Group on AI (AI HLEG) published a series of policy and investment recommendations,39 which acknowledged the need to foster the creation of “trusted data spaces” for data sharing, referencing the data trust proposals in Canada and the United Kingdom as examples.

Of particular interest in the Canadian proposal is the idea that data trusts could be used to alleviate the burden of consensual exhaustion and privacy self-management for transactions involving de-identified data. Specifically, the proposal states that “de-identified information could be processed without consent when managed by a data trust,”40 before adding that other protections such as a “prohibition against intentional re-identification or targeting of individuals in data, or re-identification as the result of negligence or recklessness” would need to be put in place.41 The proposal further stresses that clear linkages between statutory enforcement provisions and the oversight of a data trust would also be necessary. The proposal reflects the belief that a trustee’s purpose-driven mandate, fiduciary duties and built-in accountability to beneficiaries could incentivise a level of proactive risk management that could remove the need to seek consent in subsequent data transactions with other trusts.

Civic data trusts

Civic data trusts move beyond the appointment of single trustees to build fiduciary governance structures that manage the use and sharing of rights to data on behalf of beneficiaries.42 The purpose of a civic data trust is to embed civic values and participation processes into the governance and use of digital technologies.43 By incorporating civic participation into the trustee organisation, civic trusts could ensure that decisions regarding the governance of data take into account evolving concepts of digital rights and the public good.44

Sean McDonald and Keith Porcaro identify at least three ways that a civic data trust is unique:

[T]heir mission is to define and support the implementation of systems of public participation in decisions about data rights; the trustee organization itself must develop public participation models for its core governance decisions; and [they] can be designed to create reciprocal

34 Open Data Institute. (2018). Data trusts: lessons from three pilots. https://theodi.org/article/odi-data-trusts-report. Interestingly, though the ODI concluded that trust law was not necessary to advance these pilot projects, it has chosen to continue using the term “data trust”, risking popular confusion as to whether or not a data trust should always imply the application of trust law, or whether the word “trust” is merely being used as a “marketing tool”. See Delacroix, S., & Lawrence, N. (2019). Op. cit.
37 https://www.oecd.org/going-digital/ai-principles
41 Ibid.
43 Ibid.
relationships between the public (the trust), technology companies (the licensee), and technology stakeholders.45

In 2018, Sidewalk Labs, a subsidiary of Alphabet, proposed to establish an “independent urban data trust”46 to help manage the data collected as part of its planned smart city development project in the city of Toronto. While Sidewalk Labs drew a lot of attention to data trusts as a concept, their proposal was criticised for its lack of detail, failure to incorporate feedback from community organisations and residents, and for not including fiduciary obligations for the proposed trustee organisation.47

Bottom-up data trusts

Delacroix and Lawrence propose a bottom-up48 approach to data trusts as a way to return the power that stems from aggregated data to individuals. Data subjects would be empowered to pool their data into a trust that would champion a social or economic benefit of their choosing.49 Professional data trustees would exercise the data rights of beneficiaries on their behalf.50 The data trustees would act as an independent intermediary that negotiates the terms of data collection and use between data subjects and data collectors.

As more people join a data trust, the trustee’s negotiating power over the data controller would grow. In similar fashion, the pooling of data rights could act as a powerful collective action mechanism against abuse by a data controller, as trustees could exercise the right of portability on behalf of all the trust’s beneficiaries and withdraw the sum of the trust’s data rights en masse.

Delacroix and Lawrence envision an ecosystem of data trusts in which data subjects could choose a trust that reflects their aspirations, and be able to switch trusts when needed. Delacroix and Lawrence explore the application of bottom-up data trusts in several domains, including health care, social media, genetics, financial services and loyalty programmes.51

Implementation challenges

Like information fiduciaries, data trusts face their own implementation challenges. First, clarity is needed regarding the legal foundation – whether in property or contract law – that would enable data subjects to pool any rights they may have to the personal data they participate in generating into a data trust of their choosing. Without changes to the current conception of data ownership, this may represent a barrier to the availability of data trusts as a viable model in the context of online platforms.

Second, data trusts would likely require a new class of professional data trustees52 capable of balancing competing and complex interests related to data access and use. Given the increasing scale of data transactions and potential risks, however, some question whether a single trustee or even a trustee organisation would be able to discharge the trustee’s duties, or if technological solutions, such as an ecosystem of “personal AI”53 trustees, may be necessary.

Core features of trusts, including the nature and scope of fiduciary obligations (but also their governance structures and technical architectures), will need to achieve a level of standardisation for data trusts to be deployed at scale. The Hague Convention on the Law Applicable to Trusts and on their Recognition,54 which uses a harmonised definition of a trust, and sets conflict rules for resolving problems in the choice of the applicable law, could be a natural starting point for this conversation.

Conversely, civil law jurisdictions – where the reception of trust law is more recent and its features more fluid – may be particularly well suited to the task of adapting fiduciary models of governance to the evolving field of data rights management. The Quebec Civil Code, for instance, conceives of the trust as a universality of rights affected to a particular purpose,55 which a trustee has positive legal powers to administer on behalf of a trustee or trustee organisation. The fact that neither the settlor, trustee or beneficiary retains rights of ownership in

45 Ibid.
51 Ibid.
52 Ibid.
54 https://assets.hcch.net/docs/8618ed48-e52f-4d5c-93c1-56d58a610cf5.pdf
the asset, moreover, may remove legal barriers to the sharing or pooling of data rights.

Other important issues that require further research and testing include the accountability and liability procedures that will need to be developed in the context of data misuse, the integrity of licensing\textsuperscript{56} in data supply chains to ensure organisations seeking to use data from a data trust do so in accordance with its terms, and the role of public institutions in defining and conducting oversight of high-level requirements for data trusts in particular sectors, for instance, to ensure data trusts themselves are not manipulated to form new oligopolies of power. Public awareness regarding the opportunity but also the risks associated with the management of their data rights, is another.

**Way forward?**

It is important to recall that this is not the first time society has successfully devised checks and balances capable of addressing problematic concentrations of power. Good governance in democratic political regimes – the separation of powers, for instance – has helped safeguard individual rights and advance the public good while providing the certainty needed for innovation and economic growth.

The information fiduciary model represents an important first step in recognising that digital platforms should hold obligations towards internet users that are proportionate to the risk of harm they may potentially cause. Nevertheless, practical limitations related to the ability of fiduciaries to manage competing duties to both data subjects and their shareholders may impact their ability to build public trust.

If the data trust agenda appears more ambitious, this is as much an indication of data trusts’ promising features as it is a reflection of the public’s aspirations for data governance in the digital age: representation, shared rights, accountability and remedy.\textsuperscript{57} Not only are these just demands, but meeting them may help create an environment for the digital economy that is sustainable in the long term.\textsuperscript{58} So while data trusts may face a number of concrete implementation challenges related to their legal, governance and architectural foundations, they remain an indisputably promising innovation that merits greater investment – narrowly, as tools that could facilitate fair and ethical data sharing to alleviate burdens related to consent and privacy self-management; and broadly, as ways to empower the public to participate in decisions regarding the use of their personal data, and to collectively seek redress in cases of harm.


\textsuperscript{58} Ibid.
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.