



Mapping (in)visibility and structural injustice in the digital space

Kebene Wodajo*

Institute for Business Ethics, University of St.Gallen, Girtannerstrasse 8, CH-9010, St.Gallen, Switzerland

ARTICLE INFO

Keywords:

Structural injustice
Digital space
Privacy
Human rights

ABSTRACT

This study aims to map digitally mediated injustice and to understand how judicial versus non-judicial bodies contextualize and translate such harm into human rights violations. This study surveys judicial and quasi-judicial cases and case reports by non-judicial bodies, mainly civil society organizations, international organizations, and media. It divides digitally mediated harms identified through the survey into three categories: direct, structural, and hybrid harm. It then examines how these three forms of harm are represented and articulated in judicial judgments and case reports. To differentiate between the three forms of digitally mediated harm, the study uses Iris Young's political philosophy of structural injustice and Johan Galtung's account of structural violence in peace studies. The focus of this study is understanding the forms of injustices that are present but rendered invisible because of how they are contextualized. Therefore, the epistemology of absence is applied as the theoretical approach, that is, interpretation of absence and invisibility. The epistemology of absence facilitates the identification of structural and intersectional injustices that are not articulated in the same way they are experienced in society. The assessment reveals four observations. (1) Structural injustice is rarely examined through a conventional adjudicatory process. (2) Harms of structural quality examined by courts are narrowly interpreted when translated into rights violations. (3) The right to privacy, often presented as a gateway right, addresses structural injustice only partially, as this right has a subject-centric narrow interpretation currently. (4) There are limitations to the mainstream way of seeing and representing risks and injustices in the digital space, and such a view yields metonymic reasoning when framing digitally produced harms. As a result, the conventional way of contextualization is blind to unconventional experiences of vulnerability, which renders structural and intersectional injustices experienced by marginalized communities invisible.

1. Introduction

Risks from emerging technologies,¹ such as artificial intelligence (AI), are multifaceted and impact a plethora of fundamental rights, including civil, political, socioeconomic, and cultural rights. The impact affects both individual and group well-being. Several UN human rights reports have shown that, alongside their benefits, emerging technologies threaten equal enjoyment of the rights to privacy, freedom of opinion and expression, access to information, and of effective protection from hate speech, discrimination, and surveillance (UNA/73/312, UNA/HRC/44/57, UNA/75/590). Likewise, scholars have identified risks from data-driven technologies and predictive models, from fields ranging from employment to national security (Eubanks, 2018; Noble, 2018; Zuboff, 2019).

Data-driven technologies mediate almost every aspect of human activities, sometimes with adverse impacts on the socioeconomic and private lives of people. Such impacts may be continuations of existing inequalities and embedded injustice within the political economy and social structure (Benjamin, 2019). They range from systemic exclusion from basic public services to discrimination based on, for example, race, sex, or sexual orientation, and manifest existing power relationships and positionality in society (Eubanks, 2018). Some risks arising from data-driven technologies disproportionately impact and directly relate to the vulnerable position of individuals or groups in a community. People in regions of humanitarian crises, such as asylum seekers and refugees, are exposed to invasive humanitarian surveillance by private and public agencies because of their vulnerable positionality resulting from cumulative factors, such as authoritarian regimes, war, national

* Corresponding author.

E-mail address: kebene.wodajo@unisg.ch.

¹ For the purpose of this paper, "emerging technologies" refers to data-driven technologies including but not limited to AI/machine learning, the Internet of Things, biometric, cyber and 5G, and digital assistants. For this categorization, the study uses the UN Library & Archives, Science & Emerging Technology at <https://library.unog.ch/c.php?g=673842&p=4792828>

<https://doi.org/10.1016/j.jrt.2022.100024>

Received 24 August 2021; Received in revised form 7 December 2021; Accepted 25 January 2022

Available online 26 January 2022

2666-6596/© 2022 The Author. Published by Elsevier Ltd on behalf of ORBIT. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

security policies, and the tech industry’s surveillance-based business models. Historically marginalized groups have become targets of continued surveillance by public agencies, such as police departments and welfare services, and are excluded from basic services partly because of existing stereotypes attached to their weak positions in society (Komarraju, Arora & Raman, 2021). Algorithmically assisted management and workplace control disproportionately impact historically marginalized communities that make up a larger portion of the low-wage workforce. These groups have historically faced higher levels of scrutiny by the state and employers grounded on stereotypes about their trustworthiness, honesty, and work ethic (Nguyen, 2021). In these conditions, digital technologies deepen, expand, and amplify existing injustice and power asymmetries.

A significant portion of risks from emerging technologies have structural and intersectional characteristics, but this risk has received little attention, if at all. Structural and intersectional characteristics refer to the path-dependent historically rooted positionality that exposes people to perpetual vulnerability, produced by multiple actors in the socio-political and economic structure (Young, 2011) and further amplified by the deployment of emerging technologies. These are harms not always recognizable through conventional causation and intent parameters of legal liability. Moreover, not every risk and injustice in the digital space is equally represented, articulated, and translated into rights violations by judicial and quasi-judicial bodies due to the nature of the risks and how the risks are interpreted. How and from whose perspective digital risks are contextualized and framed is rarely questioned (Arora, 2016), and there is scant empirical evidence.

Citron & Solove (2022) analyzed why privacy harms caused by technologies merit rethinking. They argued that the aggregative nature of the harm, the unknowable and future risk, and individualistic focus of judicial remedies were among challenges in addressing digital privacy harms through litigation. However, their work did not specifically interrogate the unique character of structural harms. Benjamin (2019) studied racial structural injustice embedded in and enabled by new technologies. She called for a justice-oriented emancipatory approach to everyday production, deployment and interpretation of data, data-driven technologies, and their real-life implications. Eubanks (2018) unpacked the disproportionate impacts of automated decision-making on the poor, and Birhane (2021) discussed challenges of algorithmic injustice, such as replication of historical bias. While this literature excellently present structural harms experienced in the digital space, it ignores the challenge of adjudicating and translating injustice into the language of rights, does not differentiate between types of harms, and questions how harms are interpreted and translated into the language of rights during an adjudication process. Benjamin’s (2019) solution of a justice-oriented approach while being a good start, needs to be narrowed down and applied more precisely. This study takes a step toward filling this research gap. Specifically, this study identifies and maps types of harms or injustices experienced in the digital space to understand how they are translated into rights violations by adjudicative bodies, such as judicial and quasi-judicial bodies, and non-judicial agencies, such as civil society organizations (CSOs), international organizations (IOs), and media.

To understand forms of injustices that are present but rendered invisible because of how they are contextualized, this study applies the epistemology of absence as the theoretical approach (Santos, 2014). The study focuses on the interpretation of absence and invisibility using a survey of 107 judicial and quasi-judicial judgments and case reports by CSOs, IOs, and media. This enables to locate structural and intersectional injustices that are not articulated in the same way they are experienced in society and translated into rights violations. The study classifies digitally mediated harms identified through the survey into three broad categories of harm: direct, structural, and hybrid. It then examines how these forms of harm are represented and articulated in judicial judgments and case reports. To differentiate between the direct and structural forms of digitally mediated harm, this study uses Young’s

(2011) definition of structural injustice and Galtung’s (1969) account of structural violence.

2. Direct and structural harms in the digital space

While the risks of rights violations and injustices replicated by emerging technologies are articulated in a growing body of literature in the language of a specific, isolated right violation. However, not enough attention has been given to the structural and intersectional nature of injustices replicated by the digital ecosystem, partly because of a lack of scrutiny of the differences and interplay between “conventional” rights violations, which I call “direct harm,” and “structural harm” produced by structural causes. Direct harm refers to harms conveniently recognizable through proof of causation, damage and when necessary, intent. In this study, the reference to “structure” denotes social, economic, and political system and forces that shape the design, experience, and use of emerging technologies. In this context, violence and harm do not necessarily denote physical abuse, but are to be understood as a harm and constraint on individuals’ capability to realize their potential (Galtung, 1969; Nussbaum, 2003).

Drawing on legal and moral philosophy and peace studies research, this section presents the differences between and interplay among direct harm and structural violence in the digital space (see Table 1). Direct harm (a conventional rights violation, under civil or criminal liability) materializes when at least three conditions are fulfilled: there is a perpetrator to whom wrong is attributable; damage/harm results from the violation of a legally protected right; and there is intention to harm (Matsuda, 2000). Mainstream moral and legal reasoning of harm that entails responsibility includes identification of an agent whose action is causally connected to the harm (Beale, 1895; Hall, 1943). The harm may arise from voluntary and intentional acts or as an unintended consequence. Furthermore, the harm must be a result of violating a legally protected right or certain moral standard. These elements are highlighted in Galtung’s (1969) peace study, in which he defined direct violence as a case in which the actors (perpetrators) and objects (victims) of violence are readily identifiable. Some harms that occur in cyberspace fall under this form of direct violence.

For instance, cases of surveillance and profiling by private and public actors have the character of direct harm. In the case of *Catt v. the United Kingdom*, a peace activist found out that police had collected and retained data about him in a police database. The European Court of Human Rights (ECtHR) found that the UK government violated Article 8 of the European Convention on Human Rights (ECHR), the right to privacy of the victim. In this case, harm was directly attributed to the UK government, and the harm sustained was identified as a privacy invasion resulting from violation of a legally protected right to privacy. A non-judicial case in Jordan involved a rights groups claiming that the licensing requirements of the Jordan Ministry of Transport would obligate ride-calling apps to share users’ personal data, trip details, and geolocation, with the law permitting judicial and security agencies to

Table 1
Differences between direct and structural harm.

		Direct harm	Structural violence
Parameters	Causation	Attributable to a particular agent’s action (public or private)	Not attributable to a particular agent’s action. Produced by structural causes.
	Harm	Noticeable and identifiable	Might not be noticed, partially identifiable (as rights violation).
	Perpetrator of the harm	Identifiable	Not fully identifiable. Produced by multiple agents.
	Intent	Necessary (except in cases such as strict liability)	Not necessary/relevant.

have direct access to companies' servers and databases, thereby facilitating mass surveillance. The harm was readily identifiable—invasion of privacy and violation of personal data protection. The harm was attributable to an identifiable agent, the Jordan Ministry of Transport, and the violated right was identifiable as a violation of privacy. The example shows the following features of digitally mediated harms: the harm was identified as a violation of specific legally protected rights, and the perpetrator and cause of the harm were detectable.

However, this is not always the case. In 2016, a report by ProPublica disclosed that, in the US, Facebook was used by housing advertisers to target Facebook members who were house-hunting and systematically excluded anyone with an affinity for African-American, Asian-American, or Hispanic people. In addition, data-driven technologies rolled out in humanitarian crisis situations to improve inter-agency and inter-state cooperation and monitor fraud tend to increase migrants' and refugees' vulnerability to different forms of human rights violations (Madianou, 2021). Function creep results in collected data being shared and repurposed; for example, it may result in people being targeted as potential terror suspects. A predictive AI model deployed to forecast the crime rate of people with no ties to criminal activity creates and reproduces racially discriminatory feedback owing to the discriminatory dataset on which the model relies (Benjamin, 2019). In these examples, even though harm can be identified partially, the source of the harm is not clearly detectable or attributable to a particular private or public agency. Rather, the harm is path dependent and imitates the vulnerable socioeconomic and political position of victims. The harm emanates from historical inequality and marginalization embedded within the economic, political, and social system (Browne, 2015). Even though the targeted exclusion of people of African origin from access to housing or disproportionate impact of predictive policing on this group is a direct harm, its root cause is traceable to the systemic violence and marginalization of African-Americans since the era of slavery and Jim Crow segregation and remains relevant in today's #BlackLivesMatter movement (Muhammad, 2011).

Structural harm is a state of deprivation and vulnerability that constrains people from fully realizing their potential. Galtung defined structural harm as indirect violence in which "there may not be any person who directly harms another person in the structure. The violence is built into the structure and shows up as unequal power and consequently as unequal life chances" (Galtung, 1969: 171). According to Young (2011), structural injustice is systemic deprivation and unjust results are traceable to structural influence rather than an individual agent. The same process that deprives some groups enables others to dominate by providing opportunities for developing their capacities. These features are often an indirect and cumulative result of different actors' decisions in an institution (Young, 2011). Structural violence is not always translated into the language of rights and may be identified as a human rights violation. Scheper-Hughes and Bourgois argued that structural violence is "generally invisible because it is part of the routine grounds of everyday life" (Scheper-Hughes & Bourgois, 2004:4). For example, the heightened vulnerability of migrants and refugees to different forms of tech-mediated risks can be attributed to their historical and contemporary weak positions in the international community (Lester & Dussart, 2014) and the position of their home country as a failed state. The recipient state often regards them as stateless, voiceless, and sometimes a national security threat (Latonero, Hiatt, Napolitano, Clericetti & Penagos, 2019; Molnar, 2019). This is purported to justify invasive surveillance by governments, the private sector, and humanitarian agencies through rolling out weakly regulated and uncritical AI technology (Madianou, 2021). In this situation, technology deployed for efficiency and often with good intentions results in unintended adverse impacts of structured violence. This violence is produced collectively by multiple well-intentioned private and public agents operating in the socioeconomic political structure.

Today's discourse of human rights in the digital ecosystem conflates direct harm and structural violence or attributes all other risks in the

digital space to direct harm. This approach seems to manifest metonymic reasoning and has two characteristics: (1) it claims to be exclusive, complete, and universal, even though it is merely one way of representation and interpretation; and (2) alternative ways of being, seeing, and interpretation cannot be conceived outside their relationship with mainstream representation (Barreto, 2021; Santos, 2014). This theoretical approach helps show context blindness and uncritical interpretation of experiences in the digital space and their translation into rights violations. This study elucidates this gray area and the invisibility of structural injustice based on empirical evidence.

3. Methodology

To understand the distribution and representation of digitally mediated harms and how these are translated into human rights violations, a total of 107 cases from judicial, quasi-judicial, and non-judicial sources are surveyed.

The cases are sourced and identified from state-based and non-state-based judicial and non-judicial sources. Subject to availability of database and language limitation, national or state-based judicial cases and complaints to regulatory bodies (quasi-judicial) are sourced from national courts' electronic databases. For cases sourced from regional judicial and quasi-judicial bodies (or regulatory bodies), this study uses electronic databases: HUDOC for cases sourced from the ECtHR, Organisation for Economic Co-operation and Development (OECD) Watch Complaints and Database of specific instances for cases sourced from the OECD, and the Economic Community of West African States (ECOWAS) court database. Company-based regulatory bodies and cases filed at the Facebook Oversight Board are also reviewed. As main sources of non-judicial CSO sources of cases, this study uses CSOs specializing in human rights and digital rights, such as Privacy International, Access Now, Business & Human Rights Resource Center, Citizen Lab, Paradigm Initiative, Open Rights Group, Amnesty International, Jordan Open Society Association, and ProPublica. These CSOs are also chosen for their relatively wide coverage of cases in terms of geography and case diversity, and to overcome informational and language barriers from different national jurisdictions (or countries). In addition, this study uses UNHRC archives under the Special Rapporteur on contemporary forms of racism. Four cases or investigative reports (by DW Documentary, BBC Africa, Al Jazeera, and The New York Times) are included. A diversity of cases based on state-based and non-state-based judicial versus non-judicial sources is chosen to understand (1) which fora propagate what types of digitally mediated harm and why; and (2) how these harms are contextualized, represented, and translated into rights violations by different bodies.

Cases are selected based on three criteria: (1) whether the case involves harm produced by emerging technologies, particularly data-driven technologies often reported to produce high human rights risks; (2) whether the alleged harm is a result of a human rights violation; and (3) availability of adequate information for the analysis. The survey, while global, is subject to information availability. Hence, from 36 countries and two regions (the EU and Middle East and North Africa countries), 23, 15, and 11 cases are linked to three jurisdictions, namely, the UK, US, and EU, respectively. All remaining jurisdictions contribute five or less than five cases each. Cases include countries or regions where the harm occurred as well as litigation in judicial and quasi-judicial cases. As shown in Fig. 1, from a total of 107 cases, 64 (63%) are judicial and quasi-judicial, while 34 (33%) are CSO and IO reports. A larger sample of cases is taken from judicial and quasi-judicial sources because this study seeks to observe how structural injustice (harms) are interpreted and examined through a conventional adjudicatory process. Then, other sources, CSOs, IOs and media, are used to compare which venue accommodates structural injustice and how these harms are contextualized, represented, and translated into rights violations by venues other than adjudication.

This study excludes five judicial cases indirectly related to digitally

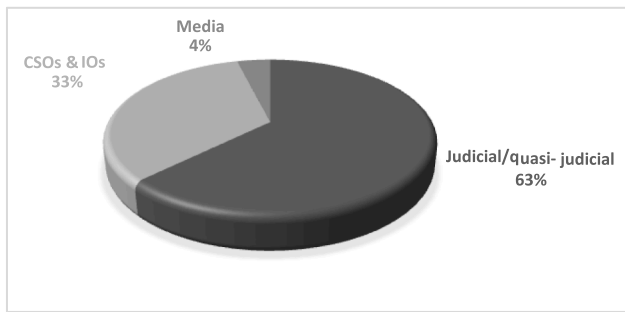


Fig. 1. Total cases surveyed and sources.

produced harm, because the disputed issues (jurisdiction, competition, and procurement) fall outside the margin of direct human rights harm.

For the analysis, two case categorizations are applied: (1) by type of harm and (2) based on rights reportedly violated. First, the cases are classified into one of direct, structural, or hybrid harms. This categorization is preferred because the study aims to make sense of the pattern and visibility of harms with structural quality in the broader body of cases brought to judicial and non-judicial venues by differentiating between different types of harms caused by technologies. While other means of categorization (e.g., based on causes of action or types of legal action) are possible, they cannot identify harms of structural nature that are not identifiable through proof of causation and actual harm. Therefore, the survey manually and systematically assesses the nature of the alleged harm in each case based on the definition drawn under Section 2 (Table 1) and how such harm is contextualized by each source. This is done through close reading of case reports, claims brought and judgments by courts, CSOs, UN archives, and media.

Cases are categorized as direct harm when the alleged damage is (1) attributable and causally linked to a specific action or inaction by individual or institutional agents, and (2) when violation of legally protected rights and sometimes intent to harm are provable. Cases are identified as structural violence when the alleged damage results from structural causes, that is, collectively and systematically produced harm that cannot linearly be attributable to an action or inaction by a single individual or institutional agent. Structural harm is a position of deprivation and vulnerability that exposes the victim to different forms of violence and is produced through the collective action of multiple actors over time (Galtung, 1969). In the hybrid category, a direct harm is identifiable but produced by underlying structural inequalities and historical injustice. To understand and identify the structural nature of harms, the positionality of victims (in each case) in the broader socio-economic and political structure is considered. This observation is inferred from the text of the case reports, claims, judgments, and established facts in the literature. In addition, differences in the types of cases often brought to judicial bodies and those reported by non-judicial bodies are carefully assessed. This analysis is used first to differentiate between different types of harms produced by technology and then to study the pattern of types of harms brought to judicial/quasi-judicial bodies compared to those reported by non-judicial organs.

Once the nature of the harm is categorized, how these harms are translated into the language of human rights violations is assessed. Which right is correlated with what type of harm and why? This analysis shows how technologically caused harms are interpreted in the language of human rights. Which rights are identified as most violated and why? How are structural harms translated into the language of human rights and does this translation reflect victims' experience? The survey directly identifies the main human rights violations in each category (structural, direct, and hybrid) as identified by claimants, judicial and non-judicial sources, and the justifications given for such identification are closely reviewed. Depending on the nature and number of harms in each case, more than one right may be violated. In some cases, multiple intersecting rights may be violated at a time; therefore, the number of rights

violated exceeds the number of cases reviewed.

4. Results and Discussions

The survey reveals four interrelated observations that help clarify the gray area in the discourse on technology and human rights, discussed in the following subsections.

4.1. Structural injustices from emerging technologies and the courts

As shown in Fig. 2, most harms that are taken to judicial or quasi-judicial bodies are direct harms. Of the 67 judicial and quasi-judicial cases, 48 are direct. Of the remainder, 13 are hybrid, and 6 structural. Structural harms constitute a larger proportion of cases reported by non-judicial bodies. The survey from non-judicial sources shows that 18 of 39 cases are hybrid and 14 structural harms. Of the 39 non-judicial case reports reviewed, only 7 are purely direct harms.

This breakdown supports the challenges in addressing structural injustice through the conventional adjudicatory process (Young, 2011). The conventional liability parameters to establish harm and liability require proof of causation, violation of legally protected rights or a moral standard, and often proof of intent to harm (Matsuda, 2000). It is almost impossible to identify these parameters, especially causation, for harms of a structural nature (Ho, 2007; Ripstein, 2018; Wettstein, 2005; Young, 2011). It is often challenging to identify harm and attribute responsibility (Jugov & Ypi, 2019). For example, in *NGO KHAM DEL-CHEVO and others v. North Macedonia*, the claimants showed how ethnic/racial profiling through retained biometric data was used for systemic discrimination and inconsistencies in the assessment of systemic racial profiling and divergent reasoning by courts on the question of identifying discrimination in such cases (European Roma Rights Centre, 2020).

Unlike observations from the survey of court cases, most cases reported by non-judicial bodies show a structural and a hybrid nature. Among the non-judicial cases surveyed, those that disproportionately impact marginalized communities and amplify existing social stereotypes have a structural and a hybrid nature and include algorithmic decision-making in healthcare services, employment and workplace management, humanitarian crisis management and border enforcement, and biometric identification. In healthcare services, an algorithm intended to help and enroll high-risk patients in care management programs unintentionally encoded racial bias by using previous healthcare cost records as a proxy for healthcare needs to predict risk level (Obermeyer, Powers, Vogeli & Mullainathan, 2019). An algorithmically assisted workplace management and employment process intended to improve managerial efficiency and objectivity was found to systematically impact and exclude people from marginalized groups and low-paying sectors owing to unrepresentative historical data that the algorithm model was trained on (AI Now Institute, 2020; Nguyen, 2021). Furthermore, the use of AI technology during humanitarian crises exposes refugees and asylum seekers to risks, such as exacerbation of existing biases and technological experimentation, especially for digital border enforcement (Johns, 2017; Latonero et al., 2019; Molnar, 2019; UNA/75/590). The harms materialized are deeply rooted in historical and structural inequalities in the social structure rather than an isolated wrongdoing or policy decision. They are path-dependent inequalities and injustices produced by multiple actors, including private and public actors, in the structure over time and manifested in digitally mediated everyday life (Eubanks, 2018). As the harm cannot be attributed to a specific agent in such cases, it is complex to adjudicate causation.

Structural inequality determines the degree of leverage one has over the structure. From the survey, this shows up in the collection, processing, use, and retention of biometric data in forced migration and humanitarian aid settings. According to research, beneficiaries have little or no meaningful control over the decision to provide their biometric data (Johns, 2017; UNHCR Innovation Service, 2020). As shown

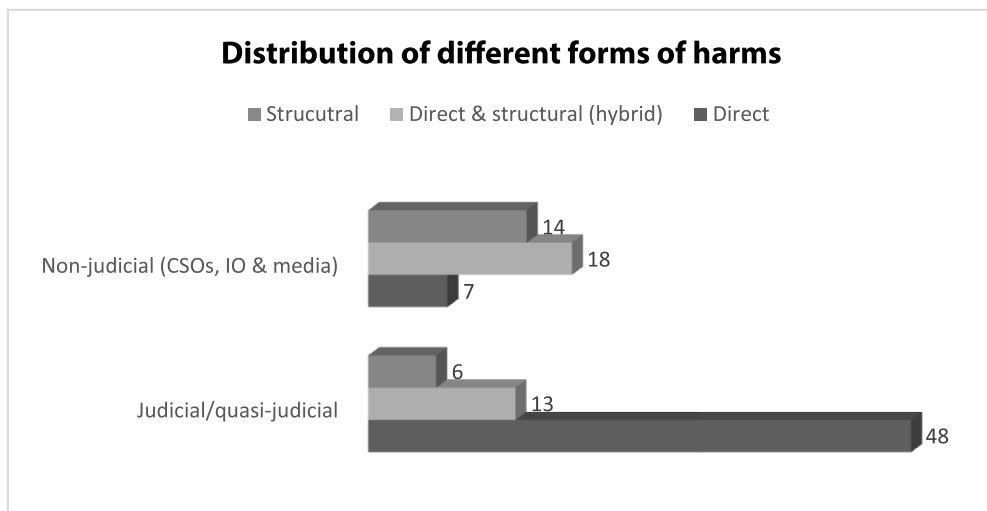


Fig. 2. Distribution of different forms of harms.

by archives of the UN Special Rapporteur, consent in this situation loses value when its alternative means starvation or a return to the crises from which they have fled. Thus, even in the absence of ill motive or wrongdoing, the weak socioeconomic and political position of some groups exposes them to intersecting vulnerabilities and risks of rights violations that are not neatly attributable to specific wrongdoers. Therefore, harms of a structural nature do not often go through the regular adjudicative process in a court of law.

4.2. Structural harms from emerging technologies as human rights violations

There are differences in the way structural and hybrid harms are interpreted and translated into rights violations by judicial and non-judicial bodies. This is derived from the case review on how rights violated are correlated to the harms, and whether the harm is interpreted to cover a broad array of rights or is narrowly focused on individual liberty by each venue. Building on the finding that structural harms are infrequently taken to court, the survey shows that structural harms that are examined by courts are narrowly interpreted and translated into human rights violations, regardless of the verdict. Often, judicial and quasi-judicial interpretations fall in the margins of civil and political rights. This trend differs to that of case reports from non-judicial sources (see Fig. 3).

From the survey, out of 15 structural and hybrid cases examined by judicial and quasi-judicial bodies, 12 were interpreted as violations of the right to privacy, 9 as violations of the right to equality (freedom from discrimination), and 4 as violations of personal integrity. The reason for this pattern is likely to relate to the contemporary design of regulatory regimes and the availability of statutory causes of action relating to privacy rights, particularly such instruments as the General Data Protection Regulation. The right to privacy as the most commonly violated right is often paired with other rights, such as the rights to equality, personal identity, nationality, access to justice and information, and personal development. Meanwhile, cases reported by non-judicial sources interpret structural and hybrid digital harms mainly as a violation of the right to equality (24 cases). Of the 31 non-judicial cases with a structural or hybrid nature, 17 cases are interpreted as violating the right to privacy. Other types of rights violations identified by non-judicial sources for cases of a structural and a hybrid nature include the rights of refugees and asylum seekers; the right to human/personal integrity; freedom of expression; the right to work; the right to food; the right to housing; the right to health; and the right to liberty, security, and personal development. Similar to the trend for judicial cases, but

with more frequency, for non-judicial cases, the right to equality and the right to privacy are paired with other types of rights.

The core differences between judicial and non-judicial bodies adopt a broader interpretation of the right to equality encompassing civil and political rights as well as socioeconomic rights. However, judicial bodies focus on the right to privacy, which in most cases is narrowly interpreted as civil and political rights. (2) While both sources use the right to equality fairly frequently to interpret and translate structural and hybrid harms into human rights violations, judicial bodies almost always restrict the interpretation to individualistic civil and political rights. Non-judicial bodies, meanwhile, apply a relatively broad interpretation with a societal lens.

An interesting contrast between non-judicial and judicial interpretations of digitally mediated structural harms is noticeable in the Indian Supreme Court's decision on the constitutionality of *Aadhaar*, India's biometric identification system. The court ruled that *Aadhaar* is valid and constitutional on the grounds that there are sufficient measures in place to protect data privacy, and that it is difficult to undertake surveillance based on *Aadhaar*. However, the court demanded that the government introduce a robust data protection law and stipulated that *Aadhaar* was not required for some services, such as applying for a mobile phone SIM card. However, these measures are partial solutions primarily focused on a narrow interpretation of privacy protection. As extensively argued by CSOs, this decision does not adequately address the risk that all-purpose biometric identification replicate and deepen structural inequality owing to technical exclusion and possibly function creep (Privacy International, 2021). According to rights groups, making access to welfare schemes conditional registration in a national database and production of an *Aadhaar* identification would exclude the most vulnerable (Privacy International, 2019). This amounts to context-blind legitimization of emerging technologies with an adverse impact based on narrowly defined protection of the right to privacy without robust consideration of its effects on other socioeconomic rights.

There are a limited number of judicial cases in which privacy and equality rights are interpreted relatively broadly but an exception is *NJCM, Platform Bescherming Burgerrechten, and others v. The Netherlands*. The court ruled that the Dutch government's use of an automated surveillance system to detect possible benefits and tax fraud was unlawful. The court's interpretation of harm covers a broad range of human rights violations. The verdict mainly interprets privacy in two interrelated settings: (1) from the perspective of the right to personal development and identity and (2) through the lens of freedom from discrimination. According to the court, the right to respect for private life encompasses

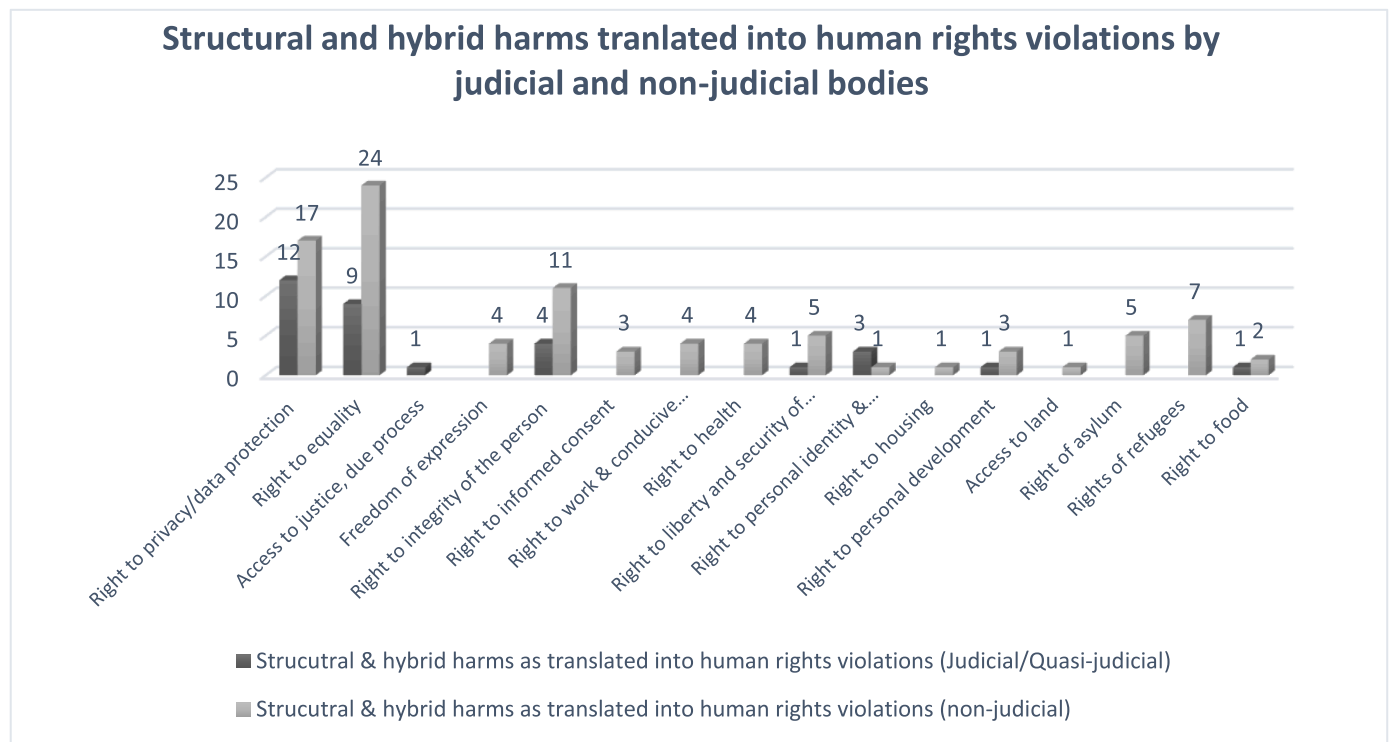


Fig. 3. Translation of structural and hybrid harms into human rights violations by judicial and non-judicial sources.

the right to personal development and personal identity, which are related to and impacted by how well personal data are protected. However, the right to private life in the context of data processing concerns the right to equal treatment and the right to protection against any form of discrimination, stereotyping, or stigmatization.² The right to personal development is an important addition that explains the enabling and disabling effect of technology and the paramount importance of protection to privacy for the sake of protecting not only liberty rights but also socioeconomic rights. However, such broad interpretation happens only rarely.

In most cases, courts stick to the narrow interpretation of both privacy and equality rights, although the tendency to pair the right to privacy with other rights supports the view that privacy is a “gateway.” However, as discussed in the following subsection, the trend toward a narrow interpretation of privacy means that structural harms have a higher chance of being ignored through the lens of privacy.

4.3. Privacy as a gateway to other rights and structural harm

The survey to some extent supports the claim that the right to privacy is a gateway to safeguard other rights in the digital space. However, this comes at the cost of overshadowing multifaceted harms of structural quality that cannot fit into and be fully addressed through individual focused consent-based privacy rights.

There is no exhaustive, conclusive definition of the right to privacy. Lately, there has been a push toward broad interpretation of privacy rights although interpretation with a social dimension is not novel (Palmer, 2009). In *Howard v. UK*, the right to privacy was interpreted by ECtHR as a right that obliges states to facilitate the right to live in and access one’s home as well as the right to protect against interference as an extension of property rights.³ European case law and scholars agree

that the right to privacy and the concept of “private life” cannot be exhaustively defined (Mowbray, 2004),⁴ but they include individuals’ rights to physical and moral integrity; rights to physical, social, and sexual identity; the right to personal development; and the right to have relationships with other humans and the outside world (Akandji-Kombe, 2007: 37; Mowbray, 2004). Meanwhile, other regional human rights documents, such as the African Charter on Human and Peoples’ Rights, do not expressly recognize the right to privacy. However, the state’s obligation to respect, protect, and fulfill elements of the right to privacy are inferred from other rights that are expressly guaranteed by the charter, such as the right to life and human dignity (African Commission on Human & Peoples’ Rights, 2015: 36).

Nonetheless, scholars advise that the scope of privacy is dynamic and context specific—it depends on society’s socio-cultural values and political philosophy (Cohen, 2019; Westin, 2003). Cohen (2019) argued that this view requires detaching privacy from the liberal legal theory built around a subject-centered protection of the autonomous self through the notion of consent. Decisions and experiences in the digital space cannot simply be theorized in terms of individual choice and consent. Individual choices and consent are socially shaped by networks of relationships, practices, beliefs, and embodied experiences (Kim, 2019). Experiences in the digital space are shaped by background socioeconomic and other factors through which existing power asymmetry invisibly play out (Nissenbaum, 2009). These invisible factors are referred to as structural inequality in this study, which locates the notion of privacy beyond individual choice and consent. The current narrow, liberal, individual-centric interpretation of the right to privacy lacks this perspective. Only a broad, dynamic, context-sensitive, and specific interpretation of privacy would enable the right to privacy to be a gateway to other rights.

Observations from a survey of judicial and quasi-judicial cases of structural and hybrid nature highlight the limitation that a narrow interpretation of privacy overshadows structural injustice. For example,

² NJCM, *Platform Bescherming Burgerrechten and others v. The Netherlands* (para 6.24).

³ *Howard v. UK*, Application No. 10825/84, 52 DR 198.

⁴ *Van Küick v. Germany*, 12 September 2003, para. 69.

in most surveillance and data extraction challenges against private and public agencies, data subjects' consent and notice are used as justifications by disputing parties and courts. However, this narrow subject-centered translation overlooks the power imbalances between individuals and the public on the one hand and private agencies on the other, as well as the historical, socio-cultural factors that shape and constrain experiences in the digital space (Cohen, 2007, 2012; Nakamura, 2002).

The limits of consent and notice as universal legitimizing conditions for satisfying privacy and data protection in the survey are illustrated in HaMoked's challenge to a mobile application against the Coordinator of Government Activities in the Territories (COGAT) in the High Court of Justice, Israel. The application was said to facilitate access to basic and necessary information concerning residency, including stay and entry permits following administrative disruption from the COVID-19 pandemic. However, it was discovered that the terms and conditions of use upon downloading the application required users to provide consent for COGAT and third parties to access all personal data. This raised questions about the sustainability of individual consent and notice-based safeguards to privacy. What happens when the data subject has no option but to accept the terms and conditions? A non-judicial case report on technological experimentation conducted by state and non-state actors on refugees, migrants, and stateless people is a case in point (Latonero et al., 2019). As archives from the UN Special Rapporteur show, such experimentation involves testing various technological products under circumstances in which targeted groups have limited or no means of providing informed consent, and where the human rights consequences of the testing and experimentation are negative or unknown (UN A/75/590, 2021: paras 38–40). In such conditions, people in a vulnerable position have no or very limited recourse to challenge the technological requirement. This is a kind of power relationship channeled through the digital space that is often overlooked in the conventional narrow interpretation of privacy. Therefore, for privacy to act as a gateway to safeguard other rights, the interpretation of privacy must go beyond a subject-centered narrow protection of the liberal "self" and reflect socioeconomic and political factors that shape the experiences and decisions of individuals in cyberspace.

4.4. Invisibility of structural harms and metonymic reasoning adopted by courts

Does judicial and quasi-judicial translation of digitally produced structural harms reflect victims' experience? The survey shows that some representations and interpretations of digital harms by judicial bodies demonstrate metonymic reasoning, which is closely linked to the limitation of a subject-centric narrow interpretation of the right to privacy and Galtung's account of invisibility of structural violence that is built into the structure and shows up as unequal power relationship (Galtung, 1969). To illustrate this observation, I use the Kenyan High Court decision on the constitutionality of *Huduma Namba*, the Kenyan biometric data system. A critical lens shows the limits of mainstream representation of human experience by considering the human consequences of alternative ways of seeing, knowing, and interpreting social phenomena (Barreto, 2021). Metonymic reasoning (1) claims to be exclusive, complete, and universal, even though it is merely one representation and interpretation; and (2) cannot conceive alternative ways of being, seeing, and interpretation (Barreto, 2021; Santos, 2014). In doing so, metonymic reasoning imposes itself through productivity and legislative thoughts and justifications. Some court verdicts observed in a survey of digitally produced structural and hybrid harms show reasoning that makes no space for non-mainstream ways of seeing and interpreting harms experienced in digitally mediated everyday life.

In *Nubian Rights Forum et al. v. the Honourable Attorney General of Kenya et al.*, the High Court of Kenya found that collecting DNA and GPS for the purpose of identification was intrusive and unnecessary, and hence, was a risk to privacy; however, the ruling did not adequately

consider the discriminatory and exclusionary effects of the biometric ID.⁵ As argued by rights groups, even though the verdict clarified the need for a robust regulatory framework addressing the possibility of exclusion, it failed to recognize the structural and historical inequalities that shape the outcome from such digitalization (Muriungi, 2019). Evidence shows that enrolment for digital identification requires proof of identity to which many marginalized sections of the community lack access (Open Society Justice Initiative, 2020). This systematically deepens existing discrimination in effect, as access to various public services is conditional on having such ID.

As correctly argued by rights groups, the court did not find the risk of discriminatory effect to be sufficient grounds to support the unconstitutionality of the biometric ID. It stated that the National Integrated Identity Management System's function was "to create, manage, maintain and operate a national population register as a single source of personal information of all Kenyan citizens and registered foreigners resident in Kenya," thereby applying to all people resident in Kenya; "[t]o that extent, the law as enacted does not differentiate between nationals or foreigners, and thus everyone... in Kenya, are entitled to register. ... [though] one has to present themselves with an identity document."⁶

While it is commendable that the Kenyan High Court upheld the pressing need to protect privacy, the argument shows a metonymic line of reasoning and blindness to underlying structural inequalities in at least two ways. First, it is problematic to limit risks that emanate from invasive biometric data collection to narrowly defined subject-centered privacy violations and hence, to ignore its discriminatory effect. Doing so would suggest that a subject-centered understanding and interpretation of privacy is a complete and universal way of seeing people's experiences in the digital space, and overlooks the historical, socioeconomic, and political factors that shape experiences and decisions in cyberspace. Second, there is resistance to an alternative way of seeing and attending to structural inequalities in the superficial treatment of the discriminatory effect of such technology on the basis that the law as enacted treats everyone equally and that the biometric identification functions to manage, maintain, and operate as a single source of personal information. In this way, the court imposed a narrowly defined representation and interpretation of harm in the digital space through productivity and legislative justification. By referring to the function of the biometric data as a single source of personal information for everyone, its reasoning was based on a productivity and efficiency argument. However, this overlooked the unquantifiable adverse human consequences of digitalization triggered by underlying structural causes (Galtung, 1969). Second, the court's claim that the law treats all equally and that nobody can be excluded from obtaining an identity document does not consider the embedded structural and historical factors that constrain and shape people's digital experiences. These factors are often disregarded by the lens of the law and do not violate legal principles in a noticeable or conventional sense of wrongdoing. As Young (2011) and other scholars stated, structural injustice is a form of injustice that materializes without breach of a law. As reported by rights groups, the discriminatory impact of *Huduma Namba* emanates from cumulative factors (Eken, 2019; Mahmoud, 2019), including policies and regulations that subject some groups to special vetting when accessing identity documents, the colonial legacy that left unequal relationships between different groups in the country; and regional security issues that created a sense of prejudice against some groups. The accumulation of historical and structural factors shapes outcomes in the digital space. Attributing these factors when interpreting and translating structural harms produced digitally to a mere violation of subject-centered privacy amounts to metonymic reasoning and, hence, acts as resistance against

⁵ *Nubian Rights Forum et al. v. the Honourable Attorney General of Kenya et al.* (paras 771; 990–991).

⁶ *Nubian Rights Forum et al. v. the Honourable Attorney General of Kenya et al.* (paras 771).

alternative ways of seeing people's experiences.

5. Conclusions

This study aimed to identify and map different types of harms channeled through cyberspace and to assess how these harms are translated and interpreted into rights violations by different agencies. The survey of cases reveals four important findings. First, judicial bodies do not often adjudicate structural harms, supporting the hypothesis that the unique character of structural injustice sets barriers for victims of digitally replicated structural harms to pursue remedies through adjudicatory processes. It can be argued that the alternative remedy for structural injustice can be sought elsewhere. While this is partly possible through such mechanisms as company-based grievance mechanisms (e. g., Facebook Oversight Board), Human Rights Due Diligence (HRDD) as required under the UN Guiding Principles on Business and Human Rights, other soft laws, and ethical codes of conduct, none of these replace the adjudicative process. With the exception of mandatory HRDD, these are mostly self-enforcing non-binding ethical standards with no third-party enforcer, leaving victims at the mercy of technology companies and states with no alternative for meaningful and enforceable remedy. Hence, alongside other efforts, it is necessary to rethink the adjudicative process for cases of digitally replicated structural injustices. This has far-reaching regulatory and policy implications, as judicial and quasi-judicial rulings set precedents and serve as a source of law, especially in common law legal systems.

Second, the limited number of structural and hybrid harms taken to court are narrowly interpreted and translated into rights violations. The predominant interpretation tends to involve a subject-centric right to privacy with minimal consideration of socioeconomic and structural factors affecting the decisions and experiences of data subjects. Third, the survey supports the emerging claim that protecting the right to privacy serves as a gateway to safeguard other rights in the digital space. However, this holds only for a broad, context-specific, and evolving concept of privacy. The current narrow interpretation of privacy means that structural harm remains only partially addressed or unaddressed in some cases. Fourth, sometimes, the narrow interpretation of harm as a mere privacy violation conditioned on notice and consent seems to demonstrate metonymic reasoning. This implies that the canonical way of regarding experience in the digital space is universal and complete, and hence, ignores the structural causes that shape human experiences and decisions in this space.

Subject to informational and language limitations, the survey in this study looked only at cases reported by judicial and non-judicial sources. Hence, not all structural injustices are covered. Moreover, structural and intersectional injustices often go unnoticed even by victims (Jugov & Ypi, 2019), making it unrealistic to capture every form of structured injustice in cyberspace. Nevertheless, the results of this study can inform the current governance gap in the adjudicatory process and policy decisions related to human rights risks of emerging technologies by highlighting the need for rethinking how human experience in cyberspace should be seen and interpreted.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study is conducted as part of the author's GFF-IPF project. The author thanks participants and organizers of the Brown Bag seminar hosted at the Institute for Business Ethics for their feedback and comments on the earlier draft of the paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jrt.2022.100024.

References

- African Commission on Human and Peoples' Rights (2015). Principles and guidelines on human and peoples' rights while countering terrorism in Africa. May.
- AI Now Institute (2018). AI Now 2018 Report.Pdf. https://ec.europa.eu/futurium/en/system/files/ged/ai_now_2018_report.pdf (accessed 11 July 2021).
- Akandji-Kombe, J.-F. (2007). Positive obligations under the European convention on human rights: A guide to the implementation of the European convention on human rights.
- Arora, P. (2016). The bottom of the data pyramid: Big data and the global south. *International Journal of Communication*, 10, 1681–1699. <https://ijoc.org/index.php/ijoc/article/view/4297> accessed 11 July 2021.
- Barreto, J.-M. (2021). Epistemologies of the South and human rights: Santos and the quest for global and cognitive justice. *Indiana Journal of Global Legal Studies*, 29, 395–422.
- Beale, J. H. (1895). Recovery for consequences of an act. *Harvard Law Review*, 9(1), 80–89. <https://doi.org/10.2307/1322317>
- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new jim code* (1st ed). Medford, MA: Polity.
- Browne, S. (2015). *Dark matters: On the surveillance of blackness*. Durham: Duke University Press.
- Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, 2(2), 1–9. <https://doi.org/10.1016/j.patter.2021.100205>
- Cohen, J.E. (2007). Cyberspace as/and space. Georgetown law faculty publications and other works, January. <https://scholarship.law.georgetown.edu/facpub/807>.
- Cohen, J. E. (2012). *Configuring the Networked Citizen*. Georgetown Law Faculty Publications and Other Works, 803, 1–19. <https://scholarship.law.georgetown.edu/facpub/803>.
- Cohen, J. E. (2019). Turning privacy inside out. *Theoretical Inquiries in Law*, 20(1), 1–31. <https://doi.org/10.1515/til-2019-0002>
- Eken, M. (2019). Rethinking Kenya's digital ID system. In: Open society justice initiative. 19 December. Available at: <https://www.justiceinitiative.org/voices/kenyas-control-versal-digital-id-scheme-faces-push-back>.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York, NY: St. Martin's Press.
- European Roma Rights Centre (2020). NGO kham delchevo and others v. north macedonia (Para. 19 & 29). Available at: http://www.errc.org/uploads/upload_en/file/5209_file1_third-party-intervention-kham-delchevo-and-others-v-north-macedonia-5-february-2020.pdf (accessed 11 July 2021).
- Galtung, J. (1969). Violence, peace, and peace research. *Journal of Peace Research*, 6(3), 167–191.
- Hall, J. (1943). Interrelations of criminal law and torts: II. *Columbia Law Review*, 43(7), 967–1001. <https://doi.org/10.2307/1117710>
- Ho, K. (2007). Structural violence as a human rights violation. Available at: /paper/Structural-Violence-as-a-Human-Rights-Violation-Ho/4b11aaf94196a56f73722d0b25383e262c3b00b.
- Johns, F. (2017). Data, detection, and the redistribution of the sensible in international law. *American Journal of International Law*, 111(1), 57–103. <https://doi.org/10.1017/ajil.2016.4>
- Jugov, T., & Ypi, L. (2019). Structural injustice, epistemic opacity, and the responsibilities of the oppressed. *Journal of Social Philosophy*, 50(1), 7–27. <https://doi.org/10.1111/josp.12268>
- Kim, N. S. (2019). *Consentability: Consent and its limits*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316691311> accessed 11 July 2021.
- Komaraju, S.A., Arora, P., & Raman, U. (2021). Agency and servitude in platform labour: A feminist analysis of blended cultures. Media, culture & society July. doi: 10.1177/01634437211029890.
- Latonero, M., Hiatt, K., Napolitano, A., Clericetti, G., & Penagos, M. (2019). Digital identity in the migration & refugee context. Italy case study. Data & society. April 15.
- Lester, A., & Dussart, F. (2014). *Colonization and the origins of humanitarian governance*. Cambridge: Cambridge University Press.
- Madianou, M. (2021). Nonhuman humanitarianism: When “AI for good” can be harmful. *Information, Communication & Society*, 24(6), 850–868. <https://doi.org/10.1080/1369118X.2021.1909100>
- Mahmoud, M. (2019). Stopping the digital ID register in Kenya—A stand against discrimination. In: Namati. 25 april. Available at: <https://namati.org/news-stories/stopping-the-digital-id-register-in-kenya-a-stand-against-discrimination/> (accessed 12 July 2021).
- Matsuda, M. (2000). On causation. Georgetown law faculty publications and other works, January. Available at: <https://scholarship.law.georgetown.edu/facpub/319>.
- Molnar, P. (2019). Technology on the margins: AI and global migration management from a human rights perspective. *Cambridge International Law Journal*, 8(2), 305–330. <https://doi.org/10.4337/cilj.2019.02.07>
- Mowbray, A. (2004). The development of positive obligations under the european convention on human rights by the european court of human rights. Bloomsbury Publishing, Oxford. Available at: <https://www.bloomsburyprofessional.com/uk/the-development-of-positive-obligations-under-the-european-convention-on-human-rights-by-the-european-court-of-human-rights-9781841132617/> (accessed 6 July 2021).

- Muhammad, K. G. (2011). *The condemnation of blackness: Race, crime, and the making of modern urban America* (31909th ed). Cambridge, Mass: Harvard University Press.
- Muriungi, P. (2019). Case study: Kenya's biometric ID system. Catalysts for collaboration. Available at: <https://catalystsforcollaboration.org/case-study-kenyas-biometric-id-system/> (accessed 12 July 2021).
- Nakamura, L. (2002). *Cybertypes: Race, ethnicity, and identity on the internet*. New York: Routledge.
- Nguyen, A. (2021). The constant boss: Work under digital surveillance. Available at: https://datasociety.net/wp-content/uploads/2021/05/The_Constant_Boss.pdf (accessed 12 July 2021).
- Nissenbaum, H. (2009). *Privacy in context: Technology, policy and the integrity of social life*. Stanford, CA: Stanford University Press.
- Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: NYU Press.
- Nussbaum, M. (2003). Capabilities as fundamental entitlements: Sen and social justice. *Feminist Economics*, 9(2–3), 33–59. <https://doi.org/10.1080/1354570022000077926>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science (New York, N.Y.)*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- Open Society Justice Initiative (2020). Nubian Rights Forum et al. v. the Honourable Attorney General of Kenya et al. (NIIMS Case). Available at: <https://www.justiceinitiative.org/litigation/nubian-rights-forum-et-al-v-the-honourable-attorney-general-of-kenya-et-al-niims-case> (accessed 12 July 2021).
- Palmer, E. (2009). Protecting socio-economic rights through the European Convention on Human Rights: Trends and developments in the European Court of Human Rights. *Erasmus Law Review*, 2, 397.
- Privacy International (2019). UNSR Submission on digital technology, social protection and human rights, p.2. May. Available at: https://privacyinternational.org/sites/default/files/2019-07/P1%20submissions%20to%20UNSR%20Extreme%20Poverty_May%202019.pdf (accessed 11 July 2021).
- Privacy International (2021). Exclusion by design: How national ID systems make social protection inaccessible to vulnerable populations. Available at: <https://privacyinternational.org/long-read/4472/exclusion-design-how-national-id-systems-make-social-protection-inaccessible> (accessed 29 March 2021).
- Ripstein, A. (2018). Reply: Relations of right and private wrongs. *Jurisprudence*, 9(3), 614–625.
- Santos, B. S. (2014). *Epistemologies of the south: Justice against epistemicide* (1st ed). London, New York: Routledge.
- Scheper-Hughes, N., & Bourgois, P. I. (2004). *Violence in war and peace: An anthology* (1st ed). Malden, Mass: Blackwell Publishing.
- Solove, D.J., & Keats Citron, D. (2021). Privacy Harms GW Law faculty publications & other works. 1534. https://scholarship.law.gwu.edu/faculty_publications/1534.
- UNHCR Innovation Service (2020). Managing risk to innovate in UNHCR. Medium, 14 October. Available at: <https://medium.com/unhcr-innovation-service/managing-risk-to-innovate-in-unhcr-91fe9294755b>.
- Westin, A. F. (2003). Social and political dimensions of privacy. *Journal of Social Issues*, 59(2), 431–453. <https://doi.org/10.1111/1540-4560.00072>
- Wettstein, F. (2005). From causality to capability. *Journal of Corporate Citizenship*, 2005 (19), 105–117. <https://doi.org/10.9774/GLEAF.4700.2005.au.00014>
- Young, I. M. (2011). *Responsibility for Justice*. Oxford political philosophy. Oxford, New York: Oxford University Press.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. London: Profile Books.