AI and Ethics (2023) 3:1049–1061
https://doi.org/10.1007/s43681-023-00311-7

OPINION PAPER

The sophistry of the neutral tool. Weaponizing artificial intelligence and big data into threats toward social exclusion

Guilherme Giantini[1]

Received: 8 February 2023 / Accepted: 4 June 2023 / Published online: 20 June 2023

Abstract

In late modernity societies, the double meaning of 'monitoring' is not a coincidence insofar, it can be a subject or an object attribution, thus suggesting a phenomenological intersection between surveillance studies and technological deployments in mass media. Emerging applications of Artificial Intelligence (AI) tools corroborate an intensified and optimized collection of personal data, either objective ones granted by individuals themselves or subject ones silently taken by algorithmic learning. From an eventual possibility to an invisible probability, Big Data may be used from devising purchase preference profiles to political bias in election periods and to reinforce bigotry against social minorities, especially transphobia. This paper's objective is to address the use of AI and Big Data as social surveillance systems tools for the establishment of more sophisticated strategies of social control. Before late modernity, disciplinary discursive power was an addressed tool to perform social control in Western societies by institutions such the Roman Catholic Church. Currently, AI technologies are tooled to perform a security-based society regulation, potentially deploying gathered data as threats against social categories that deviate from moral-based norms. Incapable of broadly embracing all cultural and social developments throughout history, such norms refer to social regulation and standardization that turn out to be exclusionary for the existence of distinctive individuals whose identities don't conform to such moral standards. The issue of ethical AI regulation is therefore grounded in questioning to what extent Western culture values and practices are still consistent in the standardized and global deployment of social and ethical policies addressed to cultures that may hold distinctive cultural perceptions and values. Theoretical reflections on post-modern panoptic frameworks, such as synoptic and banoptic devices, were carried out to assess the impact of emerging surveillance technologies as social control strategies for the reinforced marginalization of categories of exclusion. Instances of recent technology-based violence discriminations, such as misogyny, religious intolerance, racism, xenophobia, and transphobia, are provided and seen through the lens of current AI development and transphobia. The efforts of global, universal, and unilateral influence of Western culture's values on AI ethical regulation is counteracted with a reflection on decentralized bottom-up approaches to culture by means of applied ethnographic research to bring the potential of local culture into AI policy making. It is expected to corroborate future research on local-based ethical AI approaches designed within a specific culture's values to mitigate and avoid social vulnerability and violence.

Keywords Artificial intelligence · Big data · Social control · Surveillance · Ethical policies · Transphobia

1 Introduction

Out of all the natural communication systems, human language is understood to have been built on multiple combinatorial and compositional settings, which enables rearranging open-ended sound-based elements into close-ended linguistic structures, such as morphemes and words [1]. However, word polysemy is a semantic phenomenon that trespasses linguistic barriers, as verified in several human language-related instances [2]. Considering the constraining of the semiotic dimension by the symbolic [3] nuances in the formulation of languages, which sets one's world perception and social communication practices, and the current state of social order and organization in Western societies, especially in late modernity, the realm of meanings can be ambiguous. However, it can also be resourceful to understand contemporary phenomena undertaken by the convergence of topics,

✉ Guilherme Giantini
giantinigui@gmail.com

[1] Faculty of Fine Arts, University of Porto, Porto, Portugal

such as social surveillance, AI, and the oppression toward individuals placed in social categories of exclusion.

The meaning of *Monitor*, both in English and Brazilian Portuguese, has a parallel double connotation grounded on the phenomenology of perception [4], namely the essential and universal relation between subject and object [5, p. 44]. In English, the monitor subject is a "person who has the job of watching or noticing particular things", with another object meaning as a 'computer screen or a device with a screen on which words or pictures can be shown" [6]. Similarly, In Brazilian Portuguese, the *monitor* subject is a "person in charge of teaching and guiding sports or certain subjects", while its literal object is a "physical or logical device that observes, supervises, controls, or verifies operations of an electronic, computational, or similar system" [7]. From a global semantic perspective, these definitions may suggest how, in the context of post-modern society, the link between social control and emerging surveillance technologies occurs.

On the one hand, social interactions have been changing upwards since the Third Industrial Revolution at the end of the 1980s [8], which attributed a notably digital and technological character to the means of production and communication [9], making it increasingly immaterial and dependent on electronic devices. On the other hand, although terminologically defined only in 1894 [10], social control is a phenomenon as old as the organization of national societies and as mutable and adaptable to social transformations as the magnitude of imperceptibility of its action strategies. Therefore, the intersection between these two phenomena allows us to refine the definition of *monitor*, by considering that such surveillance and control are *also* aimed at the individual who uses the device, and not just on the opposite. The concept of 'watching and being watched' still and strongly applies in current post-modern times to the same extent that the surveillance plays on social scale is neither unilateral nor centralized, as was the case in modern Western era, in the image of Foucault's Panopticon [11].

Having in mind the ubiquitous presence and daily dependence on electronic devices, such as monitor screens from personal computers and mobiles, and the maintenance of invisible social control strategies, the objective of this paper is to evidence the use of AI as social surveillance systems for the establishment of new, softer, and more pervasive forms of social control. Based on the maintenance of status quo structures and aiming at security-based society regulation, such AI tools may be used as threats toward social categories that deviate from the moral-based norms if not regulated by ethical policies. Incapable of broadly embracing all cultural and social developments throughout history, such norms refer to social regulation and standardization that turn out to be excluding for the existence of distinctive individuals whose identities don't conform to such moral standards.

This paper was carried out by a theoretical reflection on the impact of the intersectional phenomenon of social control and emerging technologies for the reinforcement of categories of exclusion. Social minorities framed in such categories are consistently under discriminations like racism, misogyny, xenophobia, religious intolerance, and especially transphobia, a key focus discussed at the image of a recent event. Thereby, this topic was approached with post-modern surveillance theories grounded on the panoptic framework [11], namely the Banopticon [12] and the Synopticon [13], as well as recent studies on the intersection of AI and social impacts.

This paper also builds on questions regarding social issues and technology, as follows: If AI technology is neutral, how is its use biased by human deployments by power institutions, such as in social control and surveillance systems? If not neutral, how have such state-of-the-art technologies been used ethically? Which should be the protocols for an ethical use of AI, and by whom, and how should they be elaborated for guaranteeing its safe and democratic deployment for society, therefore, avoiding abusive uses that perpetuate inequalities?

It is expected to contribute to a social issues perspective of AI's current discussion, addressing que maintenance, and reinforcement of bigotry toward social minorities as recurrent evidence. A brief genealogy of the influence of power institutions in surveillance toward social control is also expected to corroborate the understanding of how AI tools may be manipulated to sustain historical inequalities and social hierarchy. The discussion of a local-based AI ethical regulation instead of moral-based policy arrangements embraces the expected contributions to the current discussion on AI and social impacts.

2 Artificial intelligence and surveillance: an up-to-date governing technology for a settled social control strategy

In late modernity societies, the same power institutions' discursive control strategies that shaped the preceding period are further developed and extended, evolving culture characterized by the weakness of tradition and the rising of people's reflexive ability regarding their role in society [14]. In this way, in which the discontinuities of modernity concomitantly become the themes and problems of society [15], individuals of the digital age become, to a certain extent, both the subject and the object to whom social control is addressed [13]. As we are constantly being watched by devices that are less and less physically tangible and visible, whether on an individual scale, by the cameras of our 'monitors', or on a collective scale, by CCTV equipment inside commercial establishments and institutions, we are docilely coerced to corroborate such ubiquitous surveillance [16] mainly by granting personal data.

These data become the potential targets of social power relations, which operative system builds on information manipulation to substantiate coercive actions of social interactions through microphysics of power, as in Foucault's thought [11], or macrophysics of power in the case where they are accumulated in online databases. To the extent that the data allow the elaboration of 'knowledge' about the individual, these are, therefore, 'known' and 'classified' according to parameters of social control, such as facial and corporeal features.

In the current context of mass digital media, personal data is either transferred by individuals themselves through cookie acceptance, newsletter subscriptions or the inevitable filling out of registrations, or is silently removed without the individual's knowledge to serve commercial or political purposes [17]. Therefore, data collection moves from an eventual possibility to an invisible probability, especially when considering the emerging applications of AI tools for information processing.

There is still no consensus regarding a consistent and precise definition of AI, blurred by misleading terminological swops with 'Machine Learning' [18] and 'Big Data' in public imagination of socio-technical concepts [19], or by AI's intentional marketing deployments by global networks as a smoke-screen or a distraction to environmental damaging and climate change [20]. Even though, the term is often used to refer to information technology systems that perform functions usually performed by humans [21]. In general terms, AI is employed in the automation of processes by asking questions, discovering, and testing hypotheses, and automatically making decisions based on advanced analysis, operating in extensive databases [22]. It also contributes to the generation and management of Big Data, understood as the "high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision-making, and process automation" [23, 24].

In the social sphere, AI systems are increasingly present, from synchronous translation tools, experimental applications of self-driving cars, digital assistants, and facial recognition services [25], anchoring themselves on huge amounts of personal data [26] to perform analytic and decision-making functions. Throughout history, ruling classes have developed systems for maintaining social order, the sustainable and democratic management of which is beneficial to society. However, social control has been, and clearly still is, frequently used to consolidate ruling classes' power and prevent social change.

2.1 Post-Panoptic theoretical and technological developments in late modern Western societies

Especially during the Modern period, social control took place through disciplinary power, as theorized by Foucault [11]. This model of power made use of such systems as instruments of social domination for the normalization of all society individuals by classifying and controlling deviant behaviors through the creation of prone-to-crime classes. However, this logic was applied to societies of the Modern period, in which social control strategies operated in a centralized way, as represented in the image of Jeremy Bentham's Panopticon [27]. For it is a theory that reflects on social control in a historical period that is no longer in force, the theoretical framework presented in *Discipline and Punish: The Birth of the Prison* (1979) does not perfectly apply to current times. However, it can be taken, according to Lyon [28], as a theoretical foundation for analyzing and understanding surveillance and social control in the period of Late Modernity. In this regard and considering the different operating logics that characterize the actual historical period, AI is a phenomenon of great impact in fields, such as technology and post-modern socio-political science, thus promoting fundamental changes in terms of social control and biopolitics [29].

2.1.1 The Panopticon

According to Foucault [11], the switch in the character of punishment between the XVIII and XIX centuries had both a viewer and a content shift, thus leaving the social spectacle and physical aspects of torture of the body to an enclosed, monitored, corrective, and institution-centered disciplinary control of the soul. These changes in social order accompanied the emergence of the prison as an architectural apparatus in the image of Bentham's concept, which inner organization enabled the surveillance of a few ones toward many surveyed others. This gave rise to the acknowledgment of panopticism as a Modern operation of social control.

2.1.2 The Synopticon

Nonetheless, in *The Viewer Society: Michel Foucault's 'Panopticon' revisited* (1997), Thomas Mathiesen wondered how absolute was the Foucauldian statement on the surveillance evolution "*from* a situation where the many see the few *to* a situation where the few see the many" [13, p. 219]. By leaving out of equation, the progressive development of modern media since the birth of mass press between 1750 and 1830, the same period of modern prison's rise, Foucault did not consider the occurring opposite surveillance phenomenon then enabled by evolving mass media. Over the past 150 to 200 years, mass media have evolved across printed press, film, radio, television, video, and now digital technologies that allow millions of people to see a few on stage by the interface of cameras and monitor screens. It has been characterizing Western societies at the light of Synopticism as well as while panoptic devices evolved to sharpen the eye of a few in the surveillance of many.

Therefore, one can notice that both surveillance phenomena occurred, and have been occurring simultaneously, in both a panoptical and a synoptical manner, underpinned by surveillance systems as a set of tools for power performed by institutions. The Roman Catholic Church and the military are examples of such once meaningful powers institutions that reciprocally operated the Panopticism and Synopticism in intimate interaction and by means of joint technologies aiming at social control [13]. The early beginning of the XXI century, precisely the year of 2001, is a key period to illustrate the Synopticon as well as to understand its effects and correlational shifts among security studies, surveillance systems and social control.

2.1.3 The banopticon

9/11's terrorist attacks are taken as a clear example of Synopticism as a mass media surveillance event [30] that trespassed the mediatic boundaries of the United States (US), being on worldwide news. This opened precedents for the US and their allies, such as the United Kingdom, Australia, and some countries of the European Union, to put in place a state of unease underpinned by an idea of global insecurity addressed to menaces by terrorism and criminal organizations at the Western level. Such a plan endorsed a kind of governmentality characterized by a state of exception that would legitimate the political discourse of war against terrorism attributed to foreigners, ethnic and religious groups. Under the statement of protection of citizens and collective survival, the obsession on security was operationalized by the dissemination of innovative technologies for intelligence services. They focused on monitoring and controlling social behavior to artificially differentiate the 'insider' from the 'outsider', that is, to specify the distinction between categories of inclusion and exclusion and sort social individuals accordingly. By circumscribing the notion of security into the boundaries of social norms, thus sharpening the frame of ab-normalization [12], such political turn aided by the emerging and precise technological surveillance systems had the risk of being endangering insofar as they could accentuate exclusion and discrimination toward many vulnerable social categories.

In this regard, Didier Bigo [12] theorized the term 'Banopticon' to refer to generalized state of exception characterized by an intensified and globalized governmental use of surveillance technologies to collect personal data and create numerous databases, thus defining potential profiles of risk based on features and behaviors of individuals. With the objective of guaranteeing the security of the normalized by predicting others' criminal behavior, personal data were then collected, exchanged, and cross-referenced among databases. Therefore,

> "(…) biometrics have become widespread and are linked by transnational databases; iris-scanners have been developed and justified at airports—now installed at Schiphol, Amsterdam, and being implemented elsewhere in Europe and North America as well; CCTV cameras are present in public places, enhanced if possible with facial recognition capacities such as the Mandrake system in Newham, South London; and DNA databanks are used to store genetic information capable of identifying known 'terrorists'" [12].

Although being an effective tool for accurate surveillance in specific environments, such as airports and transiting zones, to promote social order, the use of personal databases did not remain exclusive to such the intent of maintaining social order. In the era of mass communication media, political instability and consumerism, personal data started to draw attention to occurrences in which their instrumentalization by powerful AI processing tools started attending interests from political, economic, religious, and governmental spheres, thus potentially bringing new forms of societal clashes.

2.2 The commodification of personal data and the manipulation of social behavior

Commercial applications of AI destabilize the formerly state-centered monopoly of social control. Devices constantly connected to the internet, particularly through consumer digital technologies, provide private companies with a vertiginous amount of personal data about users. This equips the private sphere with the power to bias individual purchasing behaviors by making use of AI to compute mass data through machine learning applications and predictive analytics software, to produce advertisements and appealing customized discounts. To the extent that such AI applications have led to the commodification of social control [29], purchasing behavior is not the only one to be influenced. Social behavior, economic issues, marketing purposes, political campaigns, provision of services, and governance challenges, in a broader spectrum, also become conditioned to social control initiatives implemented by Big Data analytic and AI [31]. A concrete manifestation of such phenomenon is the influence power that mass communication tools, namely Twitter, had in the 2016 US elections, which resulted in the seizure of power by the Republican Donald Trump in the presidency [32].

An ascending application of AI in social surveillance occurs in facial recognition systems. These systems have already been used for almost 20 years [26, 33] in the US, operating on government-provided images, such as citizen ID and driver's license photos[26, 34]. In present days, they operate in online public databases fed by users (such as Facebook, YouTube, Instagram, and other social media platforms). Just like the mechanism of 'Clearview AI', a US facial recognition company, these tools market the access toward surveillance services [26] to other private companies and state bodies. This is a common practice in China, where the exercise of mass social surveillance with the use of AI facial recognition technologies is already a very well established 'security' strategy [35]. Chinese private companies have been making a continued push for leadership and primacy in AI as they export surveillance technologies to liberal democracies in the Global North and South [26, 36], also targeting authoritarian states such as Saudi Arabia in the building of '*smart cities*' [37]. Therefore,

> "AI technology makes possible social control, whether in China, as an expression of its authoritarian regime, or globally, by allowing Chinese access to these systems and their data, and by facilitating local authorities in their social control of citizens" [26, p. 51].

In a global aspect of digital technology-based surveillance systems and social control by means of commodified personal data, what is mainly at stake are the ethical risks toward social assistance policies, anti-democratic practices, social deviation, and violations of human rights [38, 39]. With no transparency and sensibleness, these systems can identify and detect activists at a protest or to chase someone on the subway, revealing sensitive data on target individuals, such as their names, where they live, what they do and who they know. This means that AI tools can be used to turn images available on the web, from social media to other websites, as potential 'ammunition' [26].

3 Weaponizing data to reinforce the exclusion of social minorities: the hazards of non-consensual information collection

The already usual non-consensual collection of personal data is potentially problematic for human rights issues of individuals who are not considered suitable into social norms, therefore being consistently targets of bigotry, such as misogyny, religious intolerance, xenophobia, racism, and transphobia, to name a few. Surveillance systems operated by AI tools might represent hazard to these social categories when their deployment is analyzed from an ethical view.

An example of both misogyny and religious intolerance may be verified in the severe laws regarding the use of hijab by Iranian women. During an interview with the secretary of headquarters of the Ministry of Good and Prohibition in August 2022 [40], the Iranian government agency suggested that police should use face recognition AI technology to monitor women who fail to comply with the hijab law. By matching video images with national identity databases, their identification becomes more accurate, then justifying the application fines and arrests [41]. In the following month, a 22-year-old woman named Jina "Mahsa" Amini died after being severely beaten by Iran's morality police for her considered inappropriate use of the hijab, which generated a wave of strikes in the country [42].

Non-white people, especially black people, have been submitted to distressing social consequences from events in which facial recognition technology is used uncritically and unethically, such as false arrests and excessive surveillance. The effect of algorithmic failures in accuracy recognition has its origin in institutional racial bias built on historical disparities, which, if further enforced, may strengthen pre-existing inequalities even more [43, 44]. Two scientific studies have analyzed the accuracy performance of facial recognition algorithms and revealed their lower precision toward that females, black people, and individuals between 18 and 30 years old [45, 46]. This may be understood as an evidence of a social category bias in the use of AI facial recognition technology in being convenient to male and white individuals.

Regarding xenophobia and fear toward the foreigner in a society, social media are likely tools for both constructive cultural exchange and hostility against those perceived as 'others' [47], as there is consistent evidence of strong links between social media platforms usage and hate crime, especially against refugees [47, 48]. Having the structural national-states differential attribution of rights and responsibilities based on citizenship status (citizen or non-citizen) [47], migrants and refugees are known for not having the same access to political rights and protections. This is visible in international borders and control movements, whose surveillance data is recently managed by Big Data and AI technologies to predict population flows and control unauthorized migration [31], leaving those seeking asylum in their own vulnerability.

The intersectional frameworks of AI and race or binary gender have been gathering attention in mediatic platforms, possibly aiding the increase of population's consciousness on the issue, but overshadowing other similar issues like xenophobia and sexual and gender identities [47]. For the latter category is frequently unobserved, corroborating missing and unmeasurable data on this regard [49], an event of transphobia will be further discussed in the following section as a comprehensive instance of discriminatory deployment of AI and Big Data into persecutory threats toward this specific category of social exclusion.

3.1 From objective to subjective data: threats to LBGT + community

The ascending technical advancement of AI has been manifested with the highest rigor when extracting precise information from users. Potentially, it may overcome the limits of the objective dimension of facts (what is seen, the materially tangible), thus accessing the users' subjective dimension (the invisible immaterial substance, the behavioral and psychological characteristics that make the subject's identity unique). In the past, capturing someone's sexual identity was a subjective social categorization, whose accuracy of perception was related to the level of familiarity and proximity to individuals who deviated from sexual norms, in this case cisgender homosexual people [50]. Recently, Deep Neural Networks, another AI subcategory, have been trained to perform this function, as in a study based on the analysis of a set of 35,326 facial images obtained from a dating website [51]. From a single image, the Neural Network.

> "could correctly distinguish between gay and heterosexual (cisgender) men in 81%, and in 71% of cases for (cisgender) women. Human judges achieved much lower accuracy: 61% for men and 54% for women. The accuracy of the algorithm increased to 91% and 83%, respectively, given five facial images per person" [51, p. 246], parenthesis added.

Although the paper builds on questionable scientific theories, as Phrenology and biological determinism [52], this study tackles a sensitive matter in AI, personal data and social studies. The question raised goes beyond the promotion of stereotypes of sexual performance and binary gender–social constructs [53] anchored in gender essentialism [54] and in various social conventions that seek to combat social deviance through the fixed and immutable correlation between sex and gender in the phenotypic manifestations of individuals. It refers to the ascending potential of AI applications in predicting psychological traits, such as sexuality, as is a critical social issue at the intersection of surveillance technology, social control strategies, and the instrumental strengthening of categories of exclusion [12] emerging from late modernity. Precisely, it is the possibility of obtaining personal data without consent that deliberately harms the privacy of individuals and may potentially be used to support persecution of non-heterosexual people in countries, such as Iran and Saudi Arabia, where 'deviant sexuality' is punishable by death [52].

3.2 AI and transphobia: the case of 'Gender Mapping'

It therefore becomes even more dangerous to belong to a community classified as deviant from the social norm in a historical period when AI surveillance technology, at the command of companies and authoritarian states interested in maintaining the interests of their elites, is used to create, analyze, and manage online databases. An instance of growing danger for historically excluded groups in the actual landscape of surveillance by AI is transgender and transsexual communities, known to be more marginalized, deprived of medical care [55] and dispossessed of many human rights than other LGBT + individuals. In August and September 2022, a transphobic event involving public databases occurred in the US: several hospitals that provide child and adult medical care related to gender transition for trans-sexual and transgender people faced bomb threats and other types of violence [56]. This is directly related to social surveillance by AI and its consequent formation of databases insofar as:

> "(…) a centralized online list called the "Gender Mapping" project, founded by Alix Aharon of the TERF (Trans Exclusionary Radical Feminists) Women's Liberation Front, is raising concerns that more providers could be targeted next. The map, which is hosted on Google Maps' publicly available software using the "My Maps" feature, documents the locations of thousands of establishments around the world that serve trans people in some way. (…) The Gender Mapping Project's website claims that its goals include "abolishing the gender industry" and holding "those who are harming [children] to account" [56, p. 1].

This is a demonstration of what can happen on a global scale when making the unethical use of advanced technologies for collecting and storing data. This intersectional issue between AI and social control can be explained within the context of late modernity: if, in modern societies, social surveillance took place in the panoptic light over a delimited and distinguished area, then, after the phenomenon of globalization, the zones of indistinction became crucial for the performance of power [57]. Therefore, post-panoptic social surveillance has become de-territorialized and rhizomatic, which corroborates exclusionary classification and control strategies [28], having as an instrument the creation and constant maintenance of databases with individuals' information for crystallizing the operation of what Foucault [58] named as biopower.

Furthermore, when considering that, from the 1980s onwards, there has been a rise of an obsessive idea of 'security' in most of social organization's *modus operandi* [59], at the beginning of the XXI century, this condition became directly linked to social surveillance studies [12, 28]. This occurs to the extent that exceptional practices typical of the state of exception are established to constantly cause uncertainty and social discomfort. Consequently, the oppression over categories of inclusion is reinforced while also strengthening exclusion of the most vulnerable ones with the intention of reducing or preventing the mobility of assigned individuals. The instrumentalization of this form of government in the shape of mechanisms of social surveillance creates and perpetuates the banishment of these social categories, for which the banoptic apparatuses are theorized [12]. Still, by resuming the effects of such a governance format in the context of mass communication and geolocation devices, which are processed and managed by AI tools, individuals are given the responsibility of keeping watch while being watched, which is why synoptic devices are also theorized and relevant [13]. This may explain, from the point of view of social surveillance theory, how organized social groups, in the image of TERFs threatening health centers for trans population in the US [56], assume the self-delegation of security by claiming the fight against the 'risk' [60]. A risk of compromising social achievements and the legitimization of identities built on biological essentialism, however necessarily raising social and cultural norms tied to a binary—and exclusionary—perspective of gender. Thus, the AI weaponized threats to health centers that provide care to those who deviate from gender and sexual norms under the pretext of 'security' is questionable for its exclusionary contents—those against which these same activists fight for in the patriarchal social system built on power dominance and privileges exclusive to men.

4 On ethicizing the tool: a matter of theological morality or of social policy?

It becomes clear that the unethical uses of AI tools are enabled by inconsistent regulation. The reliable and transparent use of AI is still conditioned to a robust formulation of ethical policies, which building protocols need a comprehensive study on democratic equity, for which cultural values and social differences need to be addressed.

Current AI policies-makers struggle to set a globally defined application standard within a unilateral approach to culture, heavily based on Western culture's values [61, 62], like those regarding sexual and gender stereotypical roles in family moral values sustained by the Roman Catholic Church [63]. This becomes socially problematic as innate and diverse cultural differences are neglected or left out of discussion [62], generating violence-prone tensions and misalignments between cultural values for "the culture that develops and shapes the AI differs from the globally diverse cultures of the human–AI interaction contexts" [64].

In this regard, another approach to AI systems and ethical stances is upheld. The deployment of a bottom-up approach to culture into AI ethical legislation is suggested by applied ethnographic research, bringing the potential of local culture to be discussed, analyzed, and embedded into policy making. As an instance of such a scenario toward a more socially equitable implementation of AI systems, the current vulnerability situation of Hijras in India and Bangladesh is discussed within the framework of gender-based violence supported by technology and considering the implications of reclaiming local culture into local AI policy legislation.

4.1 The paradoxical discourse and the catholic power resumption quest via AI

Although historically known to unify social and cultural behavior by a unique set of moral values to be followed, as for the Inquisition example, the Roman Catholic Church has entered the AI and Ethics discussion. Stated in the recent document *Rome Call for AI Ethics* [65], the Vatican.

> "(…) supports an ethical approach to Artificial Intelligence and promote a sense of responsibility among organizations, governments, and institutions with the aim to create a future in which digital innovation and technological progress serve human genius and creativity and not their gradual replacement." [66, p. 1].

Representing an institution that for a long time held the power of social control long before modern societies and that today tries to recover its importance [67] without giving up its doctrine for the exercise of power [68], this statement by Pope Francis initiates a paradoxical discussion. It focuses on the threat to human activity by uncritically emphasizing the technocratic paradigm that, for economic, financial, and political purposes, tends to gain control over the human subject [69] and, therefore, it is necessary.

> "(…) to accept that technological products are not neutral, for they create a framework which ends up conditioning lifestyles and shaping social possibilities along the lines dictated by the interests of certain powerful groups. Decisions which may seem purely instrumental are in reality decisions about the kind of society we want to build" [69, p. 80].

However, speeches that display very coherent contents toward anti-discrimination and human rights demand special attention to warning inconsistencies, especially when rendered at the intersection of ethical policies and new technological tools with mass media impact. When delivered by current and former power institutions, like the very Roman Catholic Church, which held power and social control in older times, such inconsistencies may bring potentially harmful impacts to society and individual freedom. If, on the one side, recent official statements [70] as well as academic studies [71] try to bring 'Christian wisdom' closer to AI ethical development in order to inform it with moral values, turning its use fair and righteous for all, on the other side, a closer look at their arguments might unveil disguised power control purposes. As Pope Francis states himself: "being homosexual isn't a crime (…) but it's a sin" [72], a progressive political discourse screened by dogmatic moral values may yet conceal discriminatory notions to maintain the persistence of categories of exclusion, therefore evidencing a moral gap in the 'fair and righteous for all'. If human rights must also be part of non-heterosexual people's lives, but "Church teaching holds that homosexuals acts are sinful, or 'intrinsically disordered'" [72], there is a risk of perpetuating old stigmatization logics into current social issues, such as the matter of ethics toward AI. Among other duties, this junction may potentially mitigate or end issues such as the moral differentiation discourses on corporeal materiality performed by both religion and medicine [73], which, in turn, stimulated "the belief that certain body types and persons fall outside measures of normality, categorizing 'different' bodies as disabled and undesirable'" [74, p. 77].

In this regard, it is relevant to frame this paradoxical operational discourse in terms of panoptic and synoptic strategies for social control, performed by The Roman Catholic Church. During its apex of power,

> "the confession during which many isolated individuals confide their secrets one by one to the unseen representative of the Church, has functioned panoptically as a setting in which the few—the priests—have seen and surveyed the many—the people of the town. Simultaneously, the Catholic Church has definitely functioned synoptically, with its enormous cathedrals intentionally placed in very visible locations for synoptical admiration, drawing large masses of people to listen to the sermon, and with the Pope speaking from the balcony of St. Peter's on Easter Day" [13, p. 223].

Mainly since the XVI century [75], both panoptical and synoptical surveillance devices for social control had fear as a discursive and strategic token, operated by the intimidating final judgment of God in the case of not compliance with Church's teachings. Now in the early XXI century, fear seems to still have a relevant importance in terms of biopolitics and ethical issues, especially within the banoptical framework. On the one hand, it is commonly associated with AI insofar as its full impact potential for society is yet known, and, on the other hand, civil society tends to behave precautionary and let fear take over decision-making in unknown situations, which may compromise liberty and rights [26], favoring normative social categories at the expense of the ban of non-normative ones [12]. Hence, it is indispensable to be alert to the notion that, in Capitalism, power reinvents [76]. It manifests in the relationships exercised in different cultural practices, devices, technologies, techniques, and strategies, but which still admits resistance [11]. In the current context of late Capitalism and Modernity, in which the greater the dispersed, broad, penetrating [77], mild, and subtle regime of power, the greater the rigor on the docilization of bodies [28], it is expected that the current and subsequent post-panoptic paradigms, operated or not by mass media, the internet and further advancements in communication technologies, still offer the possibility of resistance to the same extent [78].

4.2 Envisioning regional and local-based AI ethical policies

Ultimately, if the aim is formulating ethical AI policies to guarantee fairness and righteousness for all, it is fundamental to understand the role of culture and which range of social aspects are taken into consideration within its technological deployments [79]. Beyond manifesting advancing know-how and technical wisdom, new technologies are a medium of imagining the future from a society's world vision [80], which means that social values are intrinsic to their design [79]. However, as instances of AI-triggered social vulnerability and exclusion demonstrate, not all cultural nuances are considered when writing and building up AI logics.

On the one hand, the Western world culture has historically played a hegemonic cultural dominance role, now working toward imposing their own values and ethical narratives into AI development, thus forging and manipulating behaviors while denying the diversity and peculiarity of other possible ethical framings [61]. According to Hagerty and Rubinov´s literature review, considering more than 800 academic research manuscripts, AI developments "in a global context are biased toward perspectives held in the U.S., and limited by a lack of research, especially outside the U.S. and Western Europe" [79, p. 1]. On the other hand, the worldwide inequalities in terms of broad access to the internet and telecommunication bandwidth capacity attest the digital exclusion and the underrepresentation of people with restricted or no internet access, corroborating marginalization tendencies in a global and technological sphere [79], especially in the Global South [81].

Actively committing to undermined regional and cultural differences from diverse societal contexts is a key action [61, 79, 82–85] to address ethical AI's critical assessment [79], endorsing the withdrawal of ethics from monopolized communication frameworks [61]. Therefore, a comprehensive ethnographic research agenda is essential in scientific [79, 81] professional [85], and social and geopolitical practices to understand how AI may reinforce social inequalities [79]. Out of global hegemonic influences, the set of information gathered from contextual cases, debates, and listening to the diversity of cultural particularisms [61]. It can be implemented in local-based ethical policies to regulate a righteous technological implementation in compliance with different meanings for privacy and fairness, for example, within a society´s inherent set of values [79].

As a matter of fact, the very concept of 'privacy' illustrates how nuanced its meaning can be interpreted and undertaken in different cultures. In 2010, the Indian government displayed a liberal approach to privacy and data protection by approaching governance "without much safeguarding personal data" [86], while Chinese government´s conservative take on privacy is shaped by a reverse proportional relation between individual autonomy and state power [62, 87]. In both situations, the idea of privacy within the relationship of government and society is different not for semantic discrepancies [79], but for distinctive cultural aspects and particularisms in regional and local societies, such as worldviews, belief systems, social practices [64], values, and behaviors.

4.2.1 Embracing a local culture into technological ethics: awareness on Hijras

In this regard, the history of Hijras in India and Bangladesh may illustrate how a specific cultural phenomenon, intersected by categories of gender, sexuality, belief, and politics, became a socially vulnerable group to whom persecution is currently endorsed by mass communication technologies. Also, it offers an opportunity to wonder how current technologies could be regulated by embedded ethnographic findings into building ethical policies to reflect a society's imagination prevented from western's cultural hegemony imposed by colonization.

Hijras are people who do not conform to the binary categories of gender, identifying neither as man nor as woman, but as a third gender [88]. In ancient India, they used to have significant status within the spheres of religion, politics, and public administration. During the Mughal Empire activity between the 1526 and 1720 [89], the Hijras were believed to hold blessing abilities and were taken as.

> "the most faithful authoritative domestic workers. They displayed enormous ability, esteem, and some managed to accumulate a huge amount of wealth. They were generally appointed as wardens of the harem, and some rose to the ranks of army generals, illustrious teachers, and court consultants. With the downturn of the Mughal Empire and the emergence of British rule, their influence was declined" [90, 91].

During the colonial time, the Western ideas and values were enforced over Indian culture in a strike to erase anything considered unclean and dirty by the lens of the Western world. Between 1858 and 2018, the Indian Penal Code has held the Sect. 377, which made illegal any "unnatural offenses" that were deemed "against the order of nature". It counted for events of discrimination, harassment, and persecution toward Hijras and any other person who would not identify as straight and cisgender [92]. Although some advancements have been conquered toward reclaiming human rights and dignity to Hijras, such as the recognition of their gender category in 2013 and 2014 by Bangladesh [93] and India [94], respectively, their social exclusion still exists.

According to Bansal et al., violence has gained a new sphere of action within the increasing use and ease of access to digital technologies worldwide, supporting gender-based violence toward women, children, sexual, religious, and ethnic minorities [95]. This encompasses cyberstalking, cyberbullying, sexual harassment, image-based abuse, doxing, and impersonation [95, 96], all of which easily carried on through social media applications and other mass communication platforms able to store users' information on Big Data sets. This remounts to the issue of inconsistent ethical policies in the context of AI and other technological tools that ends up on maintaining or even reinforcing structural social exclusion. In low- and middle-income Asian countries, the existing regulation policies do not guarantee the safeguard of all citizens. They are also undertaken to allegedly prosecute people, like when the "anti-obscenity and anti-pornography laws have been used in India and Bangladesh to persecute vulnerable populations, such as Hijras and the LGBTQIA + community" [95]. Although The Supreme Court of India has informed, in 2017, that that privacy is a fundamental right guaranteed by the Indian constitution [97], it does not make up for the actual state of policies regarding data privacy and protection.

While the Indian Penal Code still reverbs a range of Western culture values, as the binary gender structures and behaviors, they consistently collide to some of those particular to the Indian culture, as the case of Hijras' non-binary third gender, leading to social injustices. Within this example of cultural values' incongruence undertaken in law making and ethical policies, the comprehension that a unifying cultural set of values may cause negative social consequences if not complying with local culture. Therefore, a regional or local-based ethical legislation, regarding technology use or other human deployments, would address the safeguard, fairness, and righteousness for all in a same cultural ecosystem if local values were actively embedded into such regulation and policies building.

5 Conclusion

This paper discusses current societal implications, as well as ethical and peril thresholds, of AI technology tools performed in discrimination-driven threats toward specific social categories, such as those targeted by misogyny, religious intolerance, xenophobia, racism, and transphobia. Reinforcing historical disparities built throughout time, AI systems may be used in power relations for the maintenance of structural inequalities. The dual semantic level of the word *monitor* is used to introduce the current relation, in late modernity societies, of either *monitoring* or *being monitored*, both mediated by technological surveillance devices. The phenomenological intersection between surveillance studies and technological deployments in mass media becomes even more relevant, serious, and dangerous when understood from the perspective of panoptical and post-panoptical theories.

The posed questions point out to AI's neutrality as a tool, to their ethical agencies into society and to how and by whom should the ethical protocol policies by elaborated to a comprehensive democratic deployment. To a certain extent, AI technologies are as neutral as any other human-made tools but differ in their operational sophistication in data collection and processing. The paradigm of AI's neutrality is not yet broken as its current deployments can be biased and bolster social discrimination [98], for which regulatory frameworks of ethical policies need to be addressed to corroborate a sustainable use of AI tools in terms of human rights guarantees [99]. Historical power structure instances, such as The Roman Catholic Church, make efforts of approaching AI ethical policy protocols with Western cultural and religious bias, preventing this technology's regulation to be broadly deployed considering local culture´s diversity, thus becoming harmful to societies built on different cultural and moral principles.

A potential righteous policy development for ethical AI use may be found on the other way of globalization culture. Rather than devising a universal set of AI policies built on similar cultural values, then standardizing ethical deployment toward a normativity, the elaboration of regional or local-based AI legislations, built on cultural divergence, social diversities, and intersections among race, gender, origin, and belief, may encompass a broader, more dynamic, and efficient ethical tool. To implement such an envision, rigorous ethnographic research can be a valuable tool to inform AI and Big Data policy-makers, rather than undertaking Western cultural values, often established during colonization periods, and followed since then. Significant for future research in the field, the study on the incorporation of diverse and specific cultural aspects into AI ethical regulations could help avoiding incongruence between proposed policies and local culture, thus mitigating social vulnerability of historically marginalized groups regarding technology-facilitated violence. If technology is shaped with a society's cultural projection for the future, then a more inclusive society in the current historical period of AI and personal data collection would benefit from abandoning Western epistemologies in reclaiming their own cultural values into authentic regulations. Therefore, it shall be expected that such approach would secure a more coherent condition of prevailing undefined principles, such as privacy, security, and protection, into the own and proper terms of a distinct cultural identity.

Acknowledgements The author hereby presents sincere acknowledgements to Professor Fátima da Cruz Rodrigues, invited Assistant Professor at the Faculty of Law of the University of Porto and Professor at the Lusíada University—North, for her lectures and instructional discussions that inspired and substantiated this research.

Funding None.

Data availability As per the theoretical approach undertaken, no dataset was generated. All information regarding transparency of collected theoretical data can be found in the References section.

Declarations

Conflict of interest No conflicts of interest to disclose.

References

1. Kirby, S., Tamariz, M., Cornish, H., Smith, K.: Compression and communication in the cultural evolution of linguistic structure. Cognition **141**, 87–102 (2015). https://doi.org/10.1016/j.cognition.2015.03.016
2. A. Vicente and I. L. Falkum, 'Polysemy', in Oxford Research Encyclopedia of Linguistics, Oxford University Press, 2017. [Online]. Available: https://oxfordre.com/linguistics/view/https://doi.org/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-325. Accessed 2 Sep 2022
3. Kristeva, J.: Revolution in poetic language, p. 271. Columbia University Press, New York (1985)
4. Merleau-Ponty, M.: Phénoménologie de la perception [Phenomenology of perception] 3 rd edition, vol. 10. Gallimard, Paris (1945)
5. R. Barbaras, La perception: essai sur le sensible. Vrin, 2009. https://isbnsearch.org/isbn/9782711621637
6. 'Monitor', *Cambridge Dictionary*. https://dictionary.cambridge.org/dictionary/english/monitor (Accessed Jan. 10, 2023).
7. Dicio, 'Monitor', *Dicio—Dicionário Online de Português*. https://www.dicio.com.br/monitor/ (Accessed Dec. 03, 2022).
8. J. Rifkin.: The Third Industrial Revolution: How Lateral Power Is Transforming Energy, the Economy, and the World. St. Martin's Publishing Group, 2011. [Online]. Available: https://books.google.pt/books?id=vbVELATjyEUC. Accessed 15 Sep 2022
9. J. P. dos Reis Velloso, L. Martins, and I. N. de A. Estudos.: A Nova ordem internacional e a terceira revolução industrial. J. Olympio Editora, 1992. [Online]. Available: https://books.google.pt/books?id=druZAAAAIAAJ
10. Hollingshead, A.B.: The concept of social control. Am. Sociol. Rev. **6**(2), 217–224 (1941). https://doi.org/10.2307/2085551
11. Foucault, M.: Discipline and punish: the birth of the prison. Random House, New York (1979)
12. Bigo, D.: 'Security, exception ban and surveillance. In: Lyon, D. (ed.) Theorizing surveillance: the Panopticon and beyond, pp. 46–68. Willan Publishing, London (2006). https://doi.org/10.1080/10714420701715563 . (**Accessed: Nov. 23, 2022. [Online]**)

13. Mathiesen, T.: The Viewer Society: Michel Foucault's "Panopticon" Revisited. Theor. Criminol. **1**(2), 215–234 (1997). https://doi.org/10.1177/1362480697001002003
14. Giddens, A.: Modernity and self-identity: self and society in the late modern age. Polity Press, Cambridge (1991)
15. Beck U, Giddens A, Lash S.: Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order. Stanford: Stanford University Press, 1994. Accessed: Dec. 03, 2022. [Online]. Available: http://www.sup.org/books/title/?id=2440
16. Andrejevic M: Reality TV: The Work of Being Watched, vol. Critical Media Studies: Institutions, Politics, and Culture. United States of America: Rowman & Littlefield Publishers, 2004.
17. Bergoglio, 'Discorso del Santo Padre ai partecipanti alla Plenaria della Pontificia Accademia per la Vita letto da S.E. Mons. Vincenzo Paglia', 2020. https://press.vatican.va/content/salastampa/it/bollettino/pubblico/2020/02/28/0134/00291.html (Accessed Dec. 05, 2022).
18. Brown S.: 'Machine learning, explained', MIT Sloan, 2021. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained (Accessed Apr. 15, 2023).
19. Elish MC, Boyd D.: 'Situating Methods in the Magic of Big Data and Artificial Intelligence'. Rochester, NY, Sep. 20, 2017. Accessed: Apr. 15, 2023. [Online]. Available: https://papers.ssrn.com/abstract=3040201
20. Crawford, K.: The atlas of AI: power, politics, and the planetary costs of artificial intelligence. Yale Univ Press (2021). https://doi.org/10.2307/j.ctv1ghv45t
21. Copeland BJ.: 'Artificial intelligence', Enciclopedia Britannica, 2022. https://www.britannica.com/technology/artificial-intelligence (Accessed Feb. 04, 2023).
22. Delcea R.: 'Artificial intelligence and Big Data', Eiopa—European Commission, Mar. 09, 2022. https://www.eiopa.europa.eu/browse/digitalisation-and-financial-innovation/artificial-intelligence-and-big-data_en (Accessed Dec. 04, 2022).
23. Gartner G.: 'Definition of Big Data—Gartner Information Technology Glossary', Gartner. https://www.gartner.com/en/information-technology/glossary/big-data (Accessed Apr. 15, 2023).
24. Henningsen, J., Cavender, B., Muccio, J., McQuade, J., Herbranson, T., Moore, C.: Big data & big data analytics. Phalanx **47**(1), 38–41 (2014)
25. Martínez-Plumed, F., Gómez, E., Hernández-Orallo, J.: Futures of artificial intelligence through technology readiness levels. Telemat Inf. **58**, 101525 (2021). https://doi.org/10.1016/j.tele.2020.101525
26. Vicini, A.: Artificial intelligence and social control: ethical issues and theological resources. J Moral Theol. **11**(1), 41–69 (2022)
27. Miller, J.-A., Miller, R.: Jeremy Bentham's Panoptic device. October **41**, 3–29 (1987). https://doi.org/10.2307/778327
28. Kenner, A.M.: The search for surveillance theories. In: Lyon, D. (ed.) Theorizing surveillance: the Panopticon and beyond, pp. 3–20. Willan Publishing, London (2006). https://doi.org/10.1080/10714420701715563 . (**Accessed: Nov. 23, 2022. [Online]**)
29. Pasquarelli W, Aijer H.: 'AI & the New Age of Social Control', 2018. http://blog.worldsummit.ai/ai-the-new-age-of-social-control (Accessed Dec. 05, 2022).
30. Lyon, D.: '9/11, Synopticon, and Scopophilia: watching and being watched. In: Haggerty, K., Ericson, R. (eds.) The new politics of surveillance and visibility, pp. 35–54. University of Toronto Press, Toronto (2005). https://doi.org/10.3138/9781442681880-003
31. Bircan, T., Korkmaz, E.E.: Big data for whose sake? Governing migration through artificial intelligence. Humanit Soc Sci Commun. **8**(1), 1 (2021). https://doi.org/10.1057/s41599-021-00910-x
32. Fujiwara, T., Muller, K., Schwarz, C.: The effect of social media on elections: evidence from the United States. NBER **28849**, 89 (2021). https://doi.org/10.3386/w28849
33. Valentino-DeVries J.: 'How the police use facial recognition, and where it falls short', the New York Times, Jan. 12, 2020. Accessed: Dec. 05, 2022. [Online]. Available: https://www.nytimes.com/2020/01/12/technology/facial-recognition-police.html
34. Hill K.: 'The Secretive Company That Might End Privacy as We Know It - The New York Times', The New York Times, 2021. https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html (Accessed Dec. 05, 2022).
35. Weber V, Ververis V.: China's surveillance state: a global project. How American and Chinese companies collaborate in the construction and global distribution of China's information control apparatus. 2021. https://www.top10vpn.com/assets/2021/07/Chinas-Surveillance-State.pdf
36. Wang M.: 'China's dystopian push to revolutionize surveillance', Washington Post, Oct. 28, 2021. Accessed: Dec. 05, 2022. [Online]. Available: https://www.washingtonpost.com/news/democracy-post/wp/2017/08/18/chinas-dystopian-push-to-revolutionize-surveillance/
37. Feldstein. S: 'How Much Is China Driving the Spread of AI Surveillance?' Carnegie Endowment for International Peace, 2019. Accessed: Dec. 05, 2022. [Online]. Available: http://www.jstor.org/stable/resrep20995.7
38. Saheb, T., Saheb, T., Carpenter, D.O.: Mapping research strands of ethics of artificial intelligence in healthcare: a bibliometric and content analysis. Comput Biol Med. **135**, 104660 (2021). https://doi.org/10.1016/j.compbiomed.2021.104660
39. Saheb, T.: Ethically contentious aspects of artificial intelligence surveillance: a social science perspective. AI Ethics (2022). https://doi.org/10.1007/s43681-022-00196-y
40. Those who wear hijabs should wait for a fine / Khabarfori's controversial conversation with the secretary of the headquarters of the Ministry of Good and Prohibition., (2022). Accessed: Jan. 29, 2023. [Online Video]. Available: https://www.youtube.com/watch?v=FUXF5Rp1wt8
41. Johnson K.: 'Iran says face recognition will ID women breaking Hijab Laws', Wired, 2023. Accessed: Jan. 28, 2023. [Online]. Available: https://www.wired.com/story/iran-says-face-recognition-will-id-women-breaking-hijab-laws/
42. McGrath M.: 'Mahsa Amini: the spark that ignited a women-led revolution', Forbes, 2022. https://www.forbes.com/sites/maggiemcgrath/2022/12/06/mahsa-amini-the-spark-that-ignited-a-women-led-revolution/ (Accessed Jan. 29, 2023).
43. Perkowitz, S.: 'the bias in the machine: facial recognition technology and racial disparities. MIT Case Stud Soc Ethical Responsib Comput (2021). https://doi.org/10.21428/2c646de5.62272586. (**No. Winter 2021**)
44. Najibi A.: 'Racial discrimination in face recognition technology'. In: Science policy and social justice, Special Edition.Harvard University, 2020. Accessed: Feb. 04, 2023. [Online]. Available: https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/
45. Khiyari, H.E., Wechsler, H.: Face verification subject to varying (age, ethnicity, and gender) demographics using deep learning. J Biom Biostat. (2016). https://doi.org/10.4172/2155-6180.1000323
46. Klare, B.F., Burge, M.J., Klontz, J.C., Vorder Bruegge, R.W., Jain, A.K.: Face recognition performance: role of demographic information. IEEE Trans. Inform. Forensic Secur. **7**(6), 1789–1801 (2012). https://doi.org/10.1109/TIFS.2012.2214212
47. Tomasev, N., Maynard, J.L., Gabriel, I.: Manifestations of Xenophobia in AI systems. arXiv (2022). https://doi.org/10.48550/arXiv.2212.07877
48. Müller, K., Schwarz, C.: Fanning the flames of hate: social media and hate crime. J. Eur. Econ. Assoc. **19**(4), 2131–2167 (2021). https://doi.org/10.1093/jeea/jvaa045

49. Tomasev N, McKee KR, Kay J, Mohamed S.: ‘Fairness for Unobserved Characteristics: Insights from Technological Impacts on Queer Communities.’ In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society. 2021. pp. 254–265. doi: https://doi.org/10.1145/3461702.3462540.
50. Brambilla, M., Riva, P., Rule, N.O.: Familiarity increases the accuracy of categorizing male sexual orientation. Personal Individ. Differ. **55**(2), 193–195 (2013). https://doi.org/10.1016/j.paid.2013.02.023
51. Wang, Y., Kosinski, M.: Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. J. Pers. Soc. Psychol. **114**(2), 246–257 (2018). https://doi.org/10.1037/pspa0000098
52. Lewis P.: ‘“I was shocked it was so easy”: meet the professor who says facial recognition can tell if you’re gay’, The Guardian, Jul. 07, 2018. Accessed: Apr. 25, 2023. [Online]. Available: https://www.theguardian.com/technology/2018/jul/07/artificial-intelligence-can-tell-your-sexuality-politics-surveillance-paul-lewis
53. Butler, J.: Gender trouble: feminism and the subversion of identity, vol. 3, pp. 171–175. Routledge, London (1989)
54. Kuhnen, T.A.: É possível dizer algo novo sobre essencialismo de gênero? Estudos Feministas **21**(1), 409–412 (2013)
55. Roberts, T.K., Fantz, C.R.: Barriers to quality health care for the transgender population. Clin Biochem **47**(10–11), 983–987 (2014). https://doi.org/10.1016/j.clinbiochem.2014.02.009
56. Factora J.: ‘TERFs Are Using Google Maps to Track and Target Trans Healthcare Providers’, Them, 2022. https://www.them.us/story/terfs-google-maps-hospitals-community-centers (Accessed Dec. 05, 2022).
57. Agamben G.: Homo Sacer: Sovereign Power and Bare Life. Stanford: Stanford University Press, 1998. Accessed: Dec. 06, 2022. [Online]. Available: http://www.sup.org/books/title/?id=2003
58. Foucault M.: ‘História da sexualidade I. A Vontade de Saber’. in Antropos. Relógio d’Água, Lisboa. 1994. https://isbnsearch.org/isbn/9789727082407
59. Garland D.: The culture of control: crime and social order in contemporary society. In: Clarendon studies in criminology. Oxford University Press, 2001. [Online]. Available: https://books.google.pt/books?id=mo-HPwAACAAJ
60. Rose N, O’Malley P, Valverde M.: ‘Governmentality’. Rochester, NY, Sep. 16, 2009. Accessed: Dec. 06, 2022. [Online]. Available: https://papers.ssrn.com/abstract=1474131
61. Goffi, E.: Escaping the Western Cosm-ethical hegemony: the importance of cultural diversity in the ethical assessment of artificial intelligence. AIEJ. (2021). https://doi.org/10.47289/AIEJ20210716-1
62. Wong, P.-H.: Cultural Differences as excuses? human rights and cultural values in global ethics and governance of AI. Philos. Technol. **33**(4), 705–715 (2020). https://doi.org/10.1007/s13347-020-00413-8
63. Seasons M (2005) ‘An examination of modern family communication and moral values in America and Europe’, Senior Honors Theses and Projects, 2005, [Online]. Available: https://commons.emich.edu/honors/67
64. Prabhakaran V, Qadri R, Hutchinson B.: ‘Cultural Incongruencies in Artificial Intelligence’, Workshop on Cultures in AI. AI in Culture, 2022. https://arxiv.org/pdf/2211.13069.pdf
65. Italian Government, ‘Rome Call for AI Ethics’, 2020. Accessed: Dec. 06, 2022. [Online]. Available: https://www.vatican.va/roman_curia/pontifical_academies/acdlife/documents/rc_pont-acd_life_doc_20202228_rome-call-for-ai-ethics_en.pdf
66. Pontifical Adacemy for Life, ‘Artificial Intelligence 2020’, 2020. https://www.academyforlife.va/content/pav/en/news/2020/intelligenza-artificiale-2020.html (Accessed Dec. 06, 2022).
67. Curp, T.D.: “A healthy revolutionary spirit?” The Catholic Church in Wielkopolska, Polish Catholicism and Poland’s Postwar Transformations, 1945–1947. Polish Rev **46**(2), 173–193 (2001)
68. Tatarczyk D.: ‘The Catholic Church and Its Impact on Public Policy in Contemporary Democracies’, Tese de Doutoramento, Western University of Michigan, Estados Unidos, 2018. [Online]. Available: https://scholarworks.wmich.edu/cgi/viewcontent.cgi?article=4305&context=dissertations
69. Bergoglio JM.: ‘Encyclical Letter LAUDATO SI’ Of the Holy Father Francis On Care for Our Common Home’, Vatican Press, Vatican, 2017. Accessed: Dec. 06, 2022. [Online]. Available: https://www.vatican.va/content/dam/francesco/pdf/encyclicals/documents/papa-francesco_20150524_enciclica-laudato-si_en.pdf
70. AP NEWS, ‘The AP Interview takeaways: Pope decries expanding gun use’, *AP NEWS*, 2023. https://apnews.com/article/pope-francis-ap-interview-highlights-8b9ec42afec4e0c0691a54f756b257bc (Accessed Jan. 27, 2023).
71. Huizinga G.: ‘Righteous AI: the Christian voice in the Ethical AI conversation’, PhD Thesis, University of Washington, United States of America, 2022. [Online]. Available: https://aiandfaith.org/wp-content/uploads/2022/09/RIGHTEOUS-AI-THE-CHRISTIAN-VOICE-IN-THE-ETHICAL-AI-CONVERSATION.pdf
72. AP NEWS, ‘The AP Interview: Pope says homosexuality not a crime’, AP NEWS, 2023. https://apnews.com/article/pope-francis-gay-rights-ap-interview-1359756ae22f27f87c1d4d6b9c8ce212 (Accessed Jan. 27, 2023).
73. Goffman, E.: Stigma: notes on the management of spoiled identity. Englewood Cliffs, Prentice-Hall (1963)
74. McLaughlin, J., Coleman-Fountain, E.: The unfinished body: the medical and social reshaping of disabled young bodies. Elsevier **120**, 76–84 (2014). https://doi.org/10.1016/j.socscimed.2014.09.012
75. Coleman, D.: Moral formation and social control in the Catholic reformation: the case of San Juan de Avila. Sixt Century J **26**(1), 17–30 (1995). https://doi.org/10.2307/2541523
76. Mokyr, J.: Capitalism reinvents itself. Curr Hist **112**(757), 291–297 (2013)
77. Cohen S.: Visions of social control: crime, punishment and classification. Wiley, 1985. Accessed: Dec. 06, 2022. [Online]. Available: https://www.wiley.com/en-it/Visions+of+Social+Control%3A+Crime%2C+Punishment+and+Classification-p-9780745600215
78. Doyle, A.: Revisiting the synopticon: Reconsidering Mathiesen’s “The Viewer Society” in the age of Web 2.0. Theor. Criminol. **15**(3), 283–299 (2011). https://doi.org/10.1177/1362480610396645
79. Hagerty A, Rubinov I.: ‘Global AI ethics: a review of the social impacts and ethical implications of artificial intelligence’, ArXiv, 2019, Accessed: Feb. 04, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Global-AI-Ethics%3A-A-Review-of-the-Social-Impacts-of-Hagerty-Rubinov/b5344d0c1281b91c223837334ae20d457666ca20
80. Heidegger, M.: The question concerning technology and other essays. Garland Publishing, New York & London (1977)
81. Blackwell, A.F.: Ethnographic artificial intelligence. Interdisc. Sci. Rev. **46**(1–2), 198–211 (2021). https://doi.org/10.1080/03080188.2020.1840226
82. Hoffmann, A.L.: Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. Inf. Commun. Soc. **22**(7), 900–915 (2019). https://doi.org/10.1080/1369118X.2019.1573912
83. Selbst AD, Boyd D, Friedler SA, Venkatasubramanian S, Vertesi J.: ‘Fairness and abstraction in sociotechnical systems.’ In: Proceedings of the Conference on Fairness, Accountability, and Transparency, in FAT* ’19. New York, NY, USA: Association

for Computing Machinery. 2019. pp. 59–68. https://doi.org/10.1145/3287560.3287598.

84. Madden M, Gilman ME, Levy K, Marwick AE.: 'Privacy, poverty and big data: a matrix of vulnerabilities for poor Americans'. Rochester, NY, 2017. Accessed: May 06, 2023. [Online]. Available: https://papers.ssrn.com/abstract=2930247
85. Van den Broek E, Sergeeva A, Huysman M.: 'Hiring algorithms: an ethnography of fairness in practice.' In: ICIS 2019 Proceedings, Munich. 2019. pp. 1–9
86. Marda V, Acharya B.: 'Identifying Aspects of Privacy in Islamic Law', The Centre for Internet and Society, 2014. https://cis-india.org/internet-governance/blog/identifying-aspects-of-privacy-in-islamic-law (Accessed May 06, 2023).
87. Song B.: 'The West may be wrong about China's social credit system', Washington Post, 2021. Accessed: May 06, 2023. [Online]. Available: https://www.washingtonpost.com/news/theworldpost/wp/2018/11/29/social-credit/
88. Hahm SC.: 'Striving to survive: human security of the Hijra of Pakistan', International Institute of Social Studies, 2010, Accessed: May 06, 2023. [Online]. Available: https://thesis.eur.nl/pub/8652
89. Richards JF.: The Mughal Empire, vol. 51. In: ACLS Humanities E-Book, no. pt. 1, v. 5, vol. 51. Cambridge University Press, 1993. [Online]. Available: https://books.google.pt/books?id=HHyVh29gy4QC
90. Gul M.:'History of marginalized community', *D+C*, 2018. https://www.dandc.eu/en/ article/british-introduced-discrimination-transgender-persons-south-Asia. Accessed 23 May 2023
91. Khanam, A.: Human rights of Hijras in Bangladesh: an analysis. Soc Sci Rev **38**(1), 249–276 (2022). https://doi.org/10.3329/ssr.v38i1.56533
92. U. Institute for Human Rights, 'India's Relationship with the Third Gender—UAB Institute for Human Rights Blog', 2018. https://sites.uab.edu/humanrights/2018/10/29/indias-relationship-with-the-third-gender/ (Accessed May 06, 2023).
93. Hossain, A.: The paradox of recognition: Hijra, third gender and sexual rights in Bangladesh. Cult. Health Sex. **19**(12), 1418–1431 (2017). https://doi.org/10.1080/13691058.2017.1317831
94. Mahapatra D.: 'Supreme Court recognizes transgenders as "third gender"', The Times of India, 2014. Accessed: May 06, 2023. [Online]. Available: https://timesofindia.indiatimes.com/india/supreme-court-recognizes-transgenders-as-third-gender/articleshow/33767900.cms
95. Bansal, V., et al.: Help-seeking behaviours of those experiencing technology-facilitated GBV in Asia: implications for policy and programming. J Gend-Based Violence **1**, 1–12 (2022). https://doi.org/10.1332/239868021X16697232129517
96. Backe, E.L., Lilleston, P., McCleary-Sills, J.: Networked individuals, gendered violence: a literature review of cyberviolence. Violence and Gender **5**(3), 135–146 (2018)
97. Panday J.: 'India's Supreme Court Upholds Right to Privacy as a Fundamental Right—and It's About Time', Electronic Frontier Foundation, 2017. https://www.eff.org/deeplinks/2017/08/indias-supreme-court-upholds-right-privacy-fundamental-right-and-its-about-time (Accessed May 07, 2023).
98. Stinson, C.: Algorithms are not neutral. AI Ethics **2**(4), 763–770 (2022). https://doi.org/10.1007/s43681-022-00136-w
99. Aizenberg, E., van den Hoven, J.: Designing for human rights in AI. Big Data Soc. **7**(2), 2053951720949566 (2020). https://doi.org/10.1177/2053951720949566

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