UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

The Digital Faces of Oppression and Domination

A Relational and Egalitarian Perspective on the Data-driven Society and its Regulation

Naudts, L.

DOI
10.1145/3630106.3658934

Publication date
2024

Document Version
Final published version

Published in
ACM FAccT '24

License
CC BY

Link to publication

Citation for published version (APA):
Naudts, L. (2024). The Digital Faces of Oppression and Domination: A Relational and Egalitarian Perspective on the Data-driven Society and its Regulation. In *ACM FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency : June 3rd-6th 2024, Rio de Janeiro, Brazil* (pp. 701-712). The Association for Computing Machinery. https://doi.org/10.1145/3630106.3658934

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:06 Oct 2026

The Digital Faces of Oppression and Domination

A Relational and Egalitarian Perspective on the Data-driven Society and its Regulation.

Laurens Naudts
AI, Media & Democracy Lab – Institute for Information Law, University of Amsterdam, Centre for IT & IP Law, KU Leuven, The Netherlands, Belgium
l.p.a.naudts@uva.nl

ABSTRACT

Drawing from Iris Marion Young's politics of difference and democratic theory, this contribution formulates a relational and egalitarian account of digital justice to understand and help counter, the social and technical conditions under which data-driven decision-making systems are liable to reinforce and introduce social injustice. To do so, this contribution is structured alongside three axes. First, I present data-driven decision-making systems as socio-technical systems that both take meaning from and co-shape people's relationships and the social structures they are part of. Due to this relational push and pull, I argue, data-driven systems have the potential to restructure society and, consequently, the conditions that govern people's exposure to, and experience of, injustice therein. Second, I transpose Young's ideation of oppression and domination onto the digital ecosystem. Both notions are used to locate within complex, dynamic and automated environments, a series of social and technological conditions that unjustifiably limit people's actions and behaviours. Third, I build on Young's model for an inclusive democracy to propose a series of institutional and procedural practices to ensure that, within the digital ecosystem, each person has the effective opportunity to pursue the life projects they value and to communicate their needs, concerns and experiences in ways that are heard and recognized by others.

CCS CONCEPTS

• **Computing Methodologies – Artificial Intelligence – Philosophical/Theoretical Foundations of Artificial Intelligence**; • **Applied Computing – Law, Social and Behavioural Science**;

KEYWORDS

Digital Justice, Structural Injustice, Algorithmic Fairness, Political Philosophy, Relational Equality, Democratic Theory, Inequality Oppression, Domination

ACM Reference Format:
Laurens Naudts. 2024. The Digital Faces of Oppression and Domination: A Relational and Egalitarian Perspective on the Data-driven Society and its Regulation.. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), June 03–06, 2024, Rio de Janeiro, Brazil.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3630106.3658934

This work is licensed under a Creative Commons Attribution International 4.0 License.

FAccT '24, June 03–06, 2024, Rio de Janeiro, Brazil
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0450-5/24/06
https://doi.org/10.1145/3630106.3658934

1 INTRODUCTION AND STRUCTURE

Social justice questions the proper ordering of social life; it seeks to formulate rules regarding how society, its social, economic and political arrangements, should be structured. Ideally, those rules accommodate, in some broad sense, the interests and needs of people in a manner that does not arbitrarily favor the interests and needs of one group over those of others [75]. Out of the basic premise that people's interests matter and matter equally - the so-called egalitarian plateau – it is possible to draw rival conclusions about how society should be structured [30, 31, 56].

In evaluating the impact of increased digitization on society, much regulatory emphasis has been placed on data-driven systems's ability to produce just material or computational outputs [51]. This outcome-oriented mindset is typically complemented with a belief that prior injustice can be corrected through technological interventions, including the implementation of data quality standards, debiasing strategies and mathematical fairness metrics [4, 9, 51]. Though not without merit, the resulting conceptions threaten to simplify the demands of justice: questions of justice are reduced to questions of local resource distribution with a technological fix [9, 51]. In viewing social injustice as a local problem, we risk to ignore the intricate and interconnected nature of the digital society and fail to target the occurrence of harm at its root [51]. Even more damningly, certain social harms, like misogyny or racism, might be viewed as something that can be patched out of existence [4, 9]. For instance, fairness metrics might realize an equitable or non-discriminatory distribution of a specific resource (credit) or opportunity (a job position), but only at a particular moment in time, and only among a privileged subset of the population. Indeed, people's ability to apply for a job or loan is already predicated on them having the necessary psychological, social and economic prerequisites to do so [101]. In this context, the relational perspective has become a powerful and convincing counternarrative and correction to this dominant AI paradigm [9, 14, 27, 43, 67, 70, 72, 93]. When evaluating justice claims, attention must also be paid to the relationships people maintain with their peers, social groups and collectives, public institutions and private corporations [2, 9, 34, 51, 82–84, 93]. This totality of relationships, including the (latent) norms and values endorsed and the way people are treated within them, condition and constrain people's actions and behaviors, their ability to grow, experience, participate and flourish as equals in social life, where they are treated with mutual respect, their concerns, experiences and needs recognized, valued, heard and listened to. Relationships determine people's opportunities, power, privilege as well as their ability to convert and enjoy these affordances. The efficacy of technical and outcome-oriented solutions therefore too depends in large part on their capacity to be sensitive toward social context.

This contribution focuses on one influential and relational account of justice: Iris Marion Young's politics of difference and model for an inclusive democracy [95, 97]. For Young, social justice concerns the promotion of self-development and self-determination. Each person should have an effective opportunity to develop, train and convert their skills into valuable beings and doings and to express their experiences, feelings and perspectives in socially recognized settings, that is, in a manner that is seen and heard, rather than obscured (self-development). Moreover, to realize their considered interests, each person should have an effective opportunity to exercise control over, that is, be able to participate in the procedures and institutions that govern their living environment (self-determination). Within the AI governance debate, Young's political philosophy appears particularly pertinent. For one, her work on justice and the politics of difference is a seminal response to the failures of the outcome-oriented or distributive justice paradigm. Moreover, in giving concrete expression to the notions of oppression and domination, she articulated the complex relational dynamics that give rise to social injustice. Finally, from this critical perspective, her democratic model explored the positive conditions under which policies can promote social justice. To realize social justice in a digital society however, we should also critically examine technology's influence over people's social surroundings. Modern societies are characterized by an increased reliance on data and algorithms to make sense of and structure social life. In so doing, these knowledge- and data-driven technologies, their ideation, design and deployment, interface with people's effective opportunities and life prospects, their mental and physical well-being, and cultural, social and economic position [37]. This renewed, tech-adapted, relational and situated perspective comprises the vantage point from which normative recommendations should be drawn concerning the regulation of the digital society. This contribution then aims to align Young's theory with modern data practices. To do so, this contribition proceeds in three steps.

First, I explore the relational dynamics between data-driven decision-making systems and social life. In this phase, I position data-driven decision-making systems as socio-technical systems that draw meaning from and actively push against their social surroundings (Section 2). Due to this relational push and pull, I argue that data-driven systems can restructure society and, consequently, the conditions that govern people's exposure to and experience of injustice. In the second phase, I transpose Young's ideation of oppression and domination onto the digital ecosystem (Section 3). My purpose for investigating both notions is two-fold. For one, oppression and domination denote social processes that are particularly inimical to people's ability to participate in social life as equals. Their function, however, is also identificatory and forward-looking. More specifically, both notions can be used to locate within complex, dynamic and automated environments, a series of social and technological conditions that limit people's capacity for self-development and self-determination. Finally, drawing from Young's democratic ideals of inclusion, political equality, reasonableness and publicity, I propose a series of institutional and procedural practices to ensure that, within the digital ecosystem, each person has the effective opportunity to pursue the life projects they value and to communicate their needs, concerns and experiences in ways that are heard and recognized by others (Section 4).

Given Young's foundational status, it is unsurprising that her work has inspired others in the field [51, 93, 101]. Likewise, AI scholarship has explored people's subjugation to structures of oppression and domination in the algorithmic context [43, 67, 71]. Rather than supplant preceding accounts, this contribution hopes to build on, strengthen, complement, engage with, and ultimately enrich ongoing digital justice efforts, relational or otherwise, through its deep and extensive engagement with Young's philosophy for the articulation (Section 2) and governance (Section 3) of digital social injustice.

2 RELATIONSHIPS IN A DIGITAL SOCIETY

In our attempt to evaluate modern society, we should not focus on algorithms as singular, isolated and dyadic decision-making units [57, 78]. Following Gabriel, modern societies comprise both social and technical elements that "interact dynamically to constitute new forms of stable institutional practice and behaviour [37]." Assuming social injustice is best understood as a relational phenomenon, how then, through these social and technical dynamics, do data-driven systems affect people's relationships, and in so doing, instruct social change [90]?

First, during the ideation and design phase, data-driven decision-making systems are informed by, and hence entangled with, the broader social structures in which they are embedded. According to Young, social structures are formed out of the many actions and interactions between individuals, (social) groups, public institutions and private corporations, which "produce and reinforce opportunities and constraints" people have access to, "the physical [and cultural] conditions of [their] future actions" and their "habits and expectations [98: 6]." For instance, algorithms trained on large datasets comprised of content scraped from popular Internet fora will incorporate the discourse found within these social spaces, including the (latent) beliefs, prejudices and stereotypes held by their members [11, 24]. Data-driven decisions thus normalize the structures of injustice that have informed their functioning. Social structures permeate the entire "anatomy of AI", including the labour and production processes that govern hardware and software development [20]. In their field research, Miceli and Posada illustrate how capitalist economic structures encourage an outsourcing of data labour that is built upon the surveillance, alienation and exploitation of workers with no sustainable prospects for economic autonomy, participation and progress [63].

Second, once deployed, socio-technical systems actively engage with their social surroundings. Learning techniques for example, draw cues from the information people (in)directly provide through their actions, preferences, or other monitorable behaviours. This interaction is perhaps still best exemplified by Microsoft's AI Chatbot Tay. Through direct communication with Twitter users, Tay soon radicalized from a friendly neighbour to a fascist nazi [50]. Recommender systems too adapt to their users' preferences [69]. Within these set-ups, people are seldom simply passive subjects but rather dynamic counterparts to the technologies they interact with. Of course, that does not take away from the fact that others typically define the underlying rules of these interactions unilaterally. The values and viewpoints embedded within these systems, over which

end-users generally have little control, represent a set of instructions that act as the boundary of people's actions and behaviours [22: 266]. For this reason, when we assess the dyadic relationship between a decision subject and a decision-making (eco)system, we must examine the conditions under which the former became part of the latter.

Third, data-driven systems might co-shape people's interpersonal relationships. Of course, existing social structures and relationships, including the latent norms and values that are reproduced within them, inform the actions, behaviours, interpretations and choices made during the ideation, development and deployment of data-driven applications. At the same time, within these creative stages, technological innovations, such as inferential analytics, are also relied upon to inform and aid in the actions and decisions made by the many hands involved. And, these technologies have altered the logic and performance of decision-making. Consequently, we should examine how technical affordances exert distinctive pressure on social structures. To help better understand this effect, perhaps a good place to start is the act of classification, which runs through the entire decisional value chain [19, 93]. During data collection, classifications are relied on to label and give meaning to training data. Those data are typically used to obtain population or group-level insights [93]. To be useful, these broad categories need not present an accurate reflection of society nor be causally connected to the tasks for which they are used. Correlations are usually deemed a sufficiently reliable and plausible ground for action [46, 65]. Still, through this process, objects, phenomena, and people become bracketed into various, sometimes ethereal, consequential categories [19]. While at times a valuable heuristic, classes nonetheless construct boundaries that constrain the actions, behaviours, and possibilities of people. And true, whether automated or analogue, most decision-making procedures mandate some form of classification. Employers need to define the desired traits of applicants when hiring, news media determine which political groups are sufficiently newsworthy, etc. What modern technologies did change however, is how class-based decisions can be applied and defined [68]. On the former level, automation enables the application of class-based information with greater uniformity and consistency. Moreover, once datafied, that information can be easily stored, shared, and, if deemed relevant, repurposed [53, 87]. In other words, automation amplifies classifications in reach and impact. Hence, classification systems that reflect histories of disadvantage reinforce, rather than merely replicate, said injustice. On the latter level, analytical techniques have greatly enhanced the level of information decision-makers can draw from in defining class structures. Having access to vast amounts of data, data analytics can be used to make sense of information that would be either too difficult or impossible for humans to parse [94]. In other words, in their capacity to map and redraw information concerning objects, people, phenomena, etc., socio-technical decision-making systems carry within them the potential to redefine, or at least co-construct, society's future boundaries and the future narratives of those who inhabit them, alongside both existing and novel social dimensions. This exact risk has prompted non-discrimination scholars to critique the "social saliency" paradigm of equality laws [40, 94]. Indeed, when people risk being excluded from various life prospects based on how they have been defined, we should question the practice of that definition, that is, of classification itself.

Finally, data-driven systems impact the relationship individuals hold vis-à-vis themselves. Data-driven systems can be optimized to reinforce and exploit people's vulnerabilities and sensibilities [91]. In addition, scholars have warned against technology's alienating effects [6, 12, 28, 88]. For instance, increased exposure to artificial agents might undermine our capacity to socialize with others [28, 88]. In this context, classification has been said to interfere with people's capacity for identity formation, including their sense of self and self-respect [23, 48, 60]. The prospect of datafication might be experienced as dehumanizing and reductive, an assassination of self-authorship [8, 60, 89]. Even if they are inaccurate, labels might still endure and persist and become part of a person's digital and perceived identity. In this sense, classifications are often not only descriptive but also co-constitutive. As labels condition people's opportunities; they can trigger resistance or conformity. The visible and invisible confrontations people have with these labels, Manders-Huits suggests, "implicitly shape their (moral) identities according to the possibilities and constraints [those labels offer] [60]." As those classes are derived from and applied to large groups of people, this representational impact often affects not only single individuals but large groups simultaneously.

The basic structure of society, Gabriel claims, is "best understood as a composite of socio-technical systems [37]." In a digital society, people are the continuous subject of various, isolated or interconnected, data-driven systems. Therefore, a key challenge for combatting the emergence of injustice concerns our ability to identify the influence these socio-technical systems exert in isolation, as part of, and co-constituting larger socio-technical structures. To do so, we must show how each decision-making system, either in isolation or in concert with others, influences their technological and social surroundings. On a technical level, we must define which data-driven processes will be used as input for others. On a social level, we need to gauge how a particular decision-making system or structure might affect a person's overall life prospects rather than focus on the discrete outcome that the system is said to produce. For instance, viewing a hiring algorithm as governing a person's job prospects alone would be wrong. Instead, we should consider that decision to be connected to the entire range of opportunities that having a job facilitates [21]. At the same time, we should not evaluate decision-making systems solely in light of the material outcomes they produce. Instead, these technologies offer renewed ways to represent, structure and understand the world due to their socio-technical character. Consequently, even when systems produce beneficial material outcomes, they might do so in a manner that misrepresents people. In our evaluation of decision-making systems and structures, we should consider their allocative and representational impact, on individuals as individuals and individuals as members of larger groups.

3 ARTICULATING DATA-DRIVEN INJUSTICE

In mapping the relational dynamics of the digital society, we have identified how data-driven technologies mediate people's (future) actions and interactions. Following Young, the next question is to articulate under what conditions these socio-technical mediating

activities, both representationally and allocatively, unjustifiably constrain people's capacity to use and learn expansive skills in socially recognized settings (self-development) and their ability to participate in defining the rules under which those interests can be realized (self-determination). In the Justice and Politics of Difference, Young identified oppression and domination as particularly inimical to both values. Oppression opposes self-development, and domination opposes self-determination. This Section transposes these notions onto the data-driven value chain to articulate and identify a series of harms people risk experiencing within complex socio-technical ecosystems and the socio-technical conditions that engender their emergence.

3.1 Oppression

According to Young, "oppression consists in systematic institutional processes which prevent some people from learning and using satisfying and expansive skills in socially recognized settings, or institutionalized social processes which inhibit people's ability to play and communicate with others or to express their feelings and perspective on social life in contexts where others can listen [95:38]." The relationships out of which these constraints emerge, Young argued, can take five different guises: exploitation, powerlessness, marginalization, cultural imperialism and violence.

3.1.1 Exploitation. Exploitation is the process under which one group's efforts, capacities and energy, including labor, are exercised under the "control, according to the purposes, and for the benefit of other people [95:49]." Exploitation enables "a few to accumulate while they constrain many more [95:49]." Exploitation is at the heart of technology development and the division of labor within it. The aforementioned outsourcing of data work to low-income countries under conditions of surveillance and alienation is an excellent case in point [63]. Even though their efforts are integral to the data-production process, workers remain undervalued, left without agency, and their tasks designated as menial and repetitive. This imagining cultivates cultures of subordination, whereby autonomy and decision-power are reserved only for those at the top [99: 93]. Furthermore, because labor conditions are streamlined and optimized toward economic efficiency, those inhabiting the lower echelons of the "occupational pyramid," often suffer mental and physical taxation. Data workers for instance might be tasked to filter and label content that is sexually graphic, violent or hateful, with little or no psychological support [74]. Whereas low-income workers carry the brunt, the social and economic benefits, such as better working language models or reduced exposure to toxic content, are transferred to others [74]. Moreover, data labor often involves workers subjugation to automated governance mechanisms. As these systems reward labour based upon the number of tasks performed, workers exposure to harm is further maximized. Similar dynamics take place within the gig-economy. Delivery drivers are managed through automated systems, the functioning of which they hold no insight or control over. Managerial proficiency through automation indeed benefits platform owners and consumers but undermines workers in their flexibility and autonomy. Instead, their working environment is defined by dynamic and opaque pricing mechanisms, which drive profit and product delivery but push workers to their limits out of fear of robo-firings [38, 67]. Without solid labour laws, gig ideals of financial independence, autonomy and job security are merely ethereal.

Exploitation, however, also occurs outside of the labour environment. For example, WorldCoin was positioned as a company that aims to realize "fairly distributed, cryptocurrency-based universal basic income." To avoid "double dipping", the novel currency is attached to biometric data points. In 2022, Guo and Renaldi reported that, to realize its start-up, the company targeted lower-income and impoverished communities as its test-users, but not necessarily as its beneficiaries. In practice, "deceptive marketing practices, [the collection of] more personal data than it acknowledged, and [a failure] to obtain meaningful informed consent", "test users were not, for the most part, [. . .] intended *end* users, [but] rather, their eyes, bodies, and very patterns of life were simply grist for Worldcoin's neural networks [39]."

3.1.2 Powerlessness. Powerlessness refers to the condition of being left without authority, respect, status, autonomy, sense of self and knowledge to decide upon or participate in, the definition of the rules, conditions and purposes that govern one's actions or behaviours. Power, or conversely, the lack thereof, is a relational phenomenon created and mediated through the actions and behaviours of many. For instance, those with power and privilege help maintain, among others, the latent norms that construct the social division of labour, including the structure of occupational distinctions, the allocation and definition of tasks within them, and the relationships held between those who carry out these tasks [99: 93]. For instance, Young remarks how society's use of the notion of "menial" helps maintain structures of servitude. Rather than being valued on their terms and contributions, when people designate "unskilled work" or "the provision of care" as menial, they reify the belief that these tasks are merely auxiliary. They are instrumental but not as praiseworthy as the specific goals they help to realize [95: 52-53]. Hence, we can more easily argue why the former can be performed under the instruction of those responsible for the latter. We compensate and listen to the engineers, the ethicists, and the lawyers, but we fail to hold data workers in the same regard. Though Young used powerlessness to denote the precarious position of nonprofessional workers within the traditional division of labour and is, therefore, closely associated with the idea of (economic) exploitation [1, 35], the enjoyment of power, including the lack thereof, is not only formed through and within economic relationships. Power is also culturally and politically mediated. Furthermore, stripping away people's autonomy, creativity, and judgment in one area of life typically carries over to others. In this context, powerlessness not only shares a mutually reinforcing connection with exploitative data practices but also with marginalization, cultural imperialism and violence discussed below.

The notion of powerlessness holds intuitive appeal as inequalities in power, including over data, characterize digital societies. The latter, Fisher and Streinz refer to as data inequality. Not only is there uneven access to data as a resource, but also uneven access to the infrastructures and technological expertise and practice needed for datafication [33]. This power to "datafy" refers to the ability of some to decide upon the what, when and how of data collection and technology creation; how data-driven technologies are imposed

onto others; and the influence these systems are allowed to exert on people's future narratives, including their social and economic visibility. In their capacity to define, generate and apply this knowledge, those with the power to datafy have been granted access to the modern means of interpretation and communication [42]. In determining how and through which tools people are judged, ranked, classified, and categorized, (a select few) actors can establish the conditions that define the losers and victors of the digital society [93]. Perhaps then, when viewed from this perspective, the digital society might have generated a novel category of powerless, i.e., those without any authority or autonomy in the society's increased datafication. Indeed, under current circumstances, those affected by data-driven decisions typically have little recourse to question or contest the choices and decisions made by those who datafy, even in areas where regulation does exist. For example, why is data so easily and cheaply extracted while simultaneously hailed as the new oil? Or, as Couldry and Mejias more strongly word, how come this "continuous appropriation [seems] natural, necessary and somehow an enhancement of, not a violence to, human development? [17]" At the same time, however, as I will explain in Section 3.3, we should remain cautious in expanding our reading of the five faces of oppression. In this context, any lack of power is not necessarily synonymous with powerlessness or oppression [95: 56]. For example, non-digital natives risk data exclusion, but they might hold political power, giving them a voice to represent their interests. For people living in poverty, society paints a different picture. They, too, risk digital exclusion but have little political clout to counter the processes that render them invisible. Those powerless lack authority even in a mediated sense [95: 56].

3.1.3 Marginalization. Marginalization captures the social processes through which a particular person or group becomes excluded from participating in cultural, social and economic life. Due to this exclusion, those marginalized are rendered invisible, voiceless, unrecognized and isolated. Consequently, their risk of being deprived of further socio-economic benefits, including fundamental rights and liberties, intensifies. In a data-driven society, people might be barred from participating in social structures and relationships of communication, cooperation and organization through strategies of exclusion.

Underrepresentation is a common source of marginalization. Even though the interests, opinions, data or information of specific individuals and groups could meaningfully inform the decision-making process and the values they represent, when the procedures and data do not account for their existence, they become digitally barred from having their interests counted. To develop or train a system, whether through data or otherwise, without knowledge of the preferences, desires, or particularities a specific group might have is to build a system ill-adapted to their needs and situation. Even if accommodated along the way, underrepresentation signifies that some people lacked recognition at least once. In her analysis of the Streetbump app, which would enable residents, through accelerometers in their smartphones, to inform governmental repair actions about existing potholes in Boston's city landscape, Crawford rightfully identified a signal problem: "People in lower income groups [were] less likely to have smartphones, and this [was] particularly true of older residents, where smartphone penetration [could] be as low as 16%. For cities like Boston, this [meant] that smartphone data sets are missing inputs from significant parts of the population — often those who have the fewest resources [18]." Though a lack of diversity in data sets is symptomatic of a grander societal problem, its resolution lies not in unbridled data collection. For one, digital underrepresentation can be desired. Not necessarily because people want to remain invisible but because they know that their actions and behaviours might become subject to punishment as soon as their data has been collected. Sex workers, drug addicts, political dissidents, etc., might have a reason to stay under the radar, but that does not mean that their interests no longer matter. The notion of data exclusion should thus extend to instances where society does too little to accommodate the interests of those who wish (or have to) remain silent and invisible.

Data does not equal recognition. People's recognition depends upon their being actively heard and recognized, not dehumanized and rendered abstract through datafication. Social exclusion is even propagated through over-exposure, too. In many situations, marginalized groups, though liable to receive a particular benefit or aid, must endure the arbitrary and invasive authority of those who control access to that aid [95: 54]. In this context, Ignazio and Klein call forth the so-called "paradox of exposure": "the double bind that places those who stand to significantly gain from being counted in the most danger from that same counting (or classifying) act [25]." As Eubanks vividly demonstrates, this paradox is at the centre of the automated welfare state: for the old, people with low incomes and people with disabilities, to gain access to social security safeguards, they must upend their rights to privacy and data protection [32, 48]. Moreover, in their propensity to produce self-fulfilling feedback loops, data-driven systems produce dynamics of over-exposure. Surveillance techniques are disproportionately placed, tested and trained in demographic areas constituted by racial and ethnic disparity. And under the wrongful assumption that higher crime identification implies a higher rate of criminality, predictive policing technologies engender excessive surveillance as much as they are mismanagement of resources that could be used elsewhere.

3.1.4 Cultural Imperialism. Cultural imperialism refers to the social processes under which a dominant group can universalize and establish their experiences and culture as the norm. Conversely, the particular perspectives and lived experiences of less privileged groups become obscured, stereotyped or marked out as the "Other [95: 59, 29]." The dominant majority's ability to do so is tied to their access, control and possession over the most critical means of interpretation and communication. As also explained elswhere, in a digital society, data-driven decision-making processes constitute a modern means of interpretation and communication, and hence, instruments to normalize a particular world view [70]. Through their ability to discover and apply (new) knowledge, analytical techniques offer decision-makers renewed ways to interpret and structure society. These newfound interpretations can be communicated to the outside world as part of algorithms' mediating function. Hence, those in charge of these systems can (unknowingly) shape the world according to their perspectives, experiences, and meanings. As mentioned above, classification encodes and naturalizes, through persistent automation, a specific way of ordering the world [19, 20]. Whereas certain classification acts might

appear trivial and devoid of risk (e.g. does a picture contain a dog or apple), others are anything but. For example, systems that look to constrain, define and classify socially constructed categories, such as gender, race and sexuality, Crawford notes, present these categories as detectable, static and fixed [19]. Those not discretely captured by a given class are rendered invisible and designated as an Other. Those value-laden choices gain additional traction upon deployment, and their impact can be lasting and profound [44]. For instance, reports have emerged on how scanning algorithms at airports misidentify trans- and gender-nonconforming persons and wrongfully categorize them as a security threat. These errors occur because these systems have been built with a specific, non-inclusive, universal definition of the human body or gender in mind [62]. Beyond the latter's questionable, pseudo-scientific basis in physiognomy, any classification systems risk obscuring and denying the lived physical and psychological experiences of those not caught by rigid categories.

Likewise, when decision-makers define (highly) subjective target variables, such as "the ideal applicant" or "the trustworthy creditor", their definitions risk gaining an aura of objectivity once integrated into administration. Though decisional criteria must be created, the problem lies in claiming that these attributes represent scientific neutrality or merit-based impartiality. As Young highlights, the criteria we use to assess and evaluate others are always normative and value-laden. More specifically, they serve to verify whether "the person evaluated supports and internalizes specific values, follows implicit or explicit social rules of behaviour, supports social purposes, or exhibits specific traits of character, behaviour, or temperament that the evaluators find desirable [95: 204]." Moreover, some of these classification acts could be seen as inherently demeaning, such as systems designed to classify individuals as beautiful or ugly [58], or stigmatizing, as is the case for fraud detection programmes that replicate the prejudicial assumption that impoverished people are more likely to commit financial crimes [7, 49].

Viewpoints are not only consequential when they pertain to statements about people. For instance, OpenAI's foundation models have been trained on data that refer to artificial intelligence primarily as AGI rather than socio-technical [52, 81]. Such terminology can be used to further technology's ideation as presenting long-term existential threats, which in turn can be abused for regulatory capture, that is, to avoid regulation that confronts the existing and pressing harms technologies cause today.

3.1.5 Violence. Once data-driven systems have normalized a particular set of norms, values and beliefs in a specific setting, deviancy can be identified and made subject to interference. While (singular) acts of physical and psychological violence are reprehensible, violence becomes oppression "due to the social context surrounding [these acts], which makes [them] possible and even acceptable [95: 61-62]." Consider the above-mentioned security example. Sure, any traveller might face an unwarranted and invasive security check following a glitch in the system. Trans travellers however, were Othered by default, and hence, more vulnerable to the threat of having to expose their bodies. Certain social groups come to know and anticipate the threat of violence, pushing them toward invisibility, isolation and alienation, depriving them of the freedom and pleasures associated with various forms of social, cultural and economic engagement. Moreover, because the existence of violence is tolerated and left unpunished, perpetrators are encouraged in their actions, which further legitimizes its occurence [95:62–63]. Technology embeds violence as a systemic social practice [47]. For transpersons, data-driven security scanners are yet another manifestation of the threat of violence they already encounter in their analogue, daily lives. Likewise, people of colour who live through police bias and brutality might know they will have to face biased predictive policing systems, which, despite known error rates, remain in use. Data-driven technologies can also be purposefully used to impose violence onto others, as is the case with surveillance tech, that has been used to aid in the enforcement of anti-trans laws [55], as well as the maintenance of Apartheid regimes [59].

3.2 Domination

The injustice of domination arises when people's social position renders them vulnerable to having their actions and choices arbitrarily interfered with by others [76, 77, 97:259-260]. Drawing from Pettit, Young views domination as undermining people's capacity for self-determination, that is, their ability to exercise agency over the conditions that shape their actions. In an era where algorithms are a modern means of production, communication and representation, those who exercise agency over their ideation, design and deployment exercise agency over the spaces in which people exercise choice, and over the conditions under which doing so is accepted, permissible or impossible. In this context, digital technologies can give additional resources to those with power and privilege to expand their control over the choices of others. Despite promises of enhanced user agency and personalization, data-driven ecosystems constitute choice architectures, the logics and values of which, are, generally speaking, chosen without the active participation of those they affect, nor with the latter's considered interests in mind. Those affected, and oppressed groups in particular, are generally left without any actionable tools to meaningfully contest the choices that make up their increasingly data-driven living environment. Instead, those choices are determined behind a veil of opacity by those with the power to datafy.

Data-driven systems can be used to interfere with people's actions in an active, dynamic, deceitful or manipulative manner. Mass-surveillance programs are perhaps a most evident form of domination. They enable the state to track and monitor citizen's daily activities, and if need be, interfere in their actions in highly dynamic, adaptive and/or autonomous ways [22, 41, 67]. In turn, these techniques have a chilling effect on the way people enjoy and exercise their fundamental rights to privacy, data protection, freedom of speech, etc. Yet, data inequalities threaten to create spaces of unwarranted domination in more general ways [41, 80]. Gräf for example, has argued that algorithmic structures that limit, replace or generate the options of decision-subjects (based upon their data, behaviour or actions), such as video portals or search engines, facilitate domination if end-users have no deliberative control over these (eco)systems' action radius. Importantly, though, domination is not simply a matter of active interference. Relationships of domination exist when the ability to interfere arbitrarily exists, regardless of whether that power is actually exercised. For example, without

strong and enforceable consumer and data protection laws, live-service providers can unilaterally change an algorithm's underlying logic. Sure, they might choose not to exercise this power, but if they were to change their mind, they could.

Whereas the previous examples characterize technology as a tool for domination, data-driven systems also indirectly sustain or introduce relationships of domination by gatekeeping people's future life prospects or range of effective opportunities [21, 77]. News recommenders, for example, affect people's ability to represent and mobilize themselves politically [69]. In this context, data-driven systems increasingly govern people's access to resources instrumental for their navigation of social and economic life. And as these resources serve as proxies for power and privilege, they help decide who can dominate and who will be dominated.

3.3 Mind The Language

The notions of domination and oppression denote specific forms of injustice that are systemic, structural, and often also historic in nature. Therefore, they should be used with care and not function as broad descriptors for any disadvantage a person or group might suffer. Young argued that oppression can only be experienced by social groups. Social groups are fluid rather than static, "collective[s] of persons differentiated from others by cultural forms, practices, special needs or capacities, structures of power and privilege" [97: 90, 95: 43]. Those shared affinities and experiences are connected to people's social positioning and relations, which condition their opportunities and life prospects and co-constitute their history and sense of identity [98: 6]. In this context, oppressed groups can be identified in relation to other groups, which, depending on the context and culture under analysis, hold more power and privilege: whites versus people of colour, men versus women, the middle class versus the working class, hetero-cis persons versus members of the LGBTQIA2S+ community, etc. In our evaluation of social and economic inequality, or the identification of unjust, discriminatory actions and behaviours, we often revert back to these salient and ascriptive group identities, such as gender or ethnicity, as proxies for the "structural social relationships that tend to privilege some more than others [96:2]." Like social praxes, social groups, too, however, evolve and fade. Hence, one should be careful not to define groups as characterized by substantive or essential attributes. In fact, in doing the latter, we might engender oppression. Indeed, we tend to think of stereotypes as perfidious exactly when they are used as unalterable fiction. People must thus remain individuals with their independent desires, needs and preferences. And, while salient attributes can be essential to one's core identity, people should be nonetheless be given the space to "transcend or reject [their associated] group identity [95]."

While cognate group membership does facilitate people's political collectivization and mobilization, Young rightfully points out that "most group-conscious political claims [. . .] are not claims to the recognition of identity as such, but rather claims for fairness, equal opportunity, and political inclusion [97: 107]." At the heart of social justice claims lies a concern with the relative disadvantage people experience in their cultural, social, economic and political recognition and participation following differing positions of power and privilege [54]. Placing disadvantage at the core of social justice has an additional benefit for this particular discussion. In Section 2, I explained how automated decision-making systems categorize and place persons within larger collectives based on their preferences, interactions or other monitorable behaviour. As data-driven systems dynamically adapt to their environments, these categories are typically ever-changing. Though consequential, these groups are seldom socially salient, nor do they always correspond or overlap with socially salient groups. Instead, they are simply aggregates, a classification of persons according to some attributes identified as relevant by the decision-maker or the inferential analytics techniques they rely on. Though Young placed social groups at the centre of social justice claims, if we value the egalitarian ideal of realizing self-development and self-determination for all, should we not also show regard for the socio-technical conditions under which algorithmic aggregates, and their members, are unjustifiably limited in their enjoyment of these values? I believe so. Yet, in the latter case, these limitations should not be labeled as oppression.

Oppression is faced by structurally disadvantaged social groups, for whom the digital environment presents an overarching system in which their likelihood of being labelled as deviant, fraudulent, anomalous, or simply the Other is heightened. Social and political underrepresentation leads to statistical underrepresentation. How can the systems build upon this data recognize their humanity if there is no trace of their existence? It cannot: the digital realm is another environment they must navigate with prejudice, stigma or other negative labels attached to them. The digital domain builds upon and reinforces unjust systems and is both systemic and structural in reproducing injustice. For instance, while specific data-driven structures, such as mass surveillance schemes, are indiscriminately imposed onto people, even in these scenarios of shared harm across large groups, positions of privilege can be identified. While everyone surveilled sacrifices part of their privacy, practice shows how, in their application, such regimes tend to disproportionately harm those less privileged, like cultural minorities and low-income neighbourhoods [61, 92]. Certain groups, Frye observes, are caught "between or among forces and barriers which are so related to each other that jointly they restrain, restrict or prevent [their] motion or mobility [36]." This is not the case for (members of) aggregates (insofar as the latter do not overlap with social groups). Indeed, (the members of) aggregates might lack the persistent social embedding to render the forms of injustice they suffer systemic and structural.

That said, in exploring digital structures of oppression and domination, areas of power and privilege could be located alongside the AI value chain. In so doing, we can more easily identify who (or what components within that chain) exerts more influence on people's resources for self-development and self-determination, and hence, where exactly people's risks of having these values interfered with lie or originate from. For example, if affected persons cannot question and contest the ideation, development and deployment of data-driven systems, others, and those with the power to datafy in particular, are left to dictate by which means, through whose efforts, for whose benefit, and in accordance to which values and world views, people's social and economic narratives are decided upon. However, rather than label all these relationships as oppressive, perhaps we could refer to the socio-technical conditions that place people at risk of suffering undue interference with their capacity

for self-development and self-determination as engendering a form of digital vulnerability, a notion that Helberger and others have used to describe "a universal state of defencelessness and susceptibility to (the exploitation of) power imbalances [45]." Whereas oppressed groups are digitally vulnerable and face domination, the opposite need not be true. Indeed, many people face domination and are digitally vulnerable in the above-mentioned sense, and likely to suffer injustice as a result. Still, they might not endure the pervasive social, structural and systemic constraints of oppression. Consequently, it is far harder to break down the barriers faced by those who are oppressed, the dismantling of which (rightfully) mandates more social, cultural and economic attention and resources. Finally, while the concepts of oppression, domination, and even digital vulnerability, do denote specific forms of digital injustice, they should be used without paternalism, undue victimization and fatalism, and always with a view toward empowerment [13, 16, 36].

4 RESOURCING AGAINST DIGITAL INJUSTICE

The relational egalitarian plight of digital social justice is twofold. First, to dismantle within the socio-technical living sphere, people's exposure to relationships that engender oppression, domination and other digital vulnerabilities. Second, to ensure that, in navigating digitally mediated spaces, people stand in relations of equality to others [2]. Social equality will only be attained when each person is given the effective opportunity to pursue the life projects they value and to communicate their needs, concerns and experiences in ways that are heard and recognized by others. And those ambitions can only be realized when people have been given an effective opportunity to actively participate in the infrastructural, architectural, and decisional socio-technical dynamics that shape society's current and future narrative. In this final Section, and drawing from Young's model for an inclusive democracy in particular [97], the goal is to formulate a series of non-exhaustive recommendations to foster people's ability to participate in the politics of the digital society.

In our attempts to regulate the digital society, we should first acknowledge that people take on different roles and positions within their relationships. Consequently, people's vulnerabilities and associated needs differ and shift. To realize social equality, regulatory efforts should therefore acknowledge how injustice takes on different guises, impacts people and groups differently, and necessitates a differentiated response. And, in societies characterized by structural inequality, historically disadvantaged, oppressed and otherwise marginalized communities should be given more resources to amplify their voice and recognition to identify and correct prior and future injustice.

If we want to guard people against data-driven decisions' potential to introduce or reinforce digital injustice, we should first investigate their restructuring potential, that is, to gauge the societal changes these systems might instil. To do so effectively, however, we must locate and examine relationships of power and privilege alongside the decisional value chain and how the choices made therein can influence technology's push and pull. Following in the footsteps of D'Ignazio and Klein, "who" questions are a valuable heuristic: Who decides and who is left powerless? Who benefits, and who risks exploitation? Who is prioritized, and who remains overlooked and risks marginalization? Whose viewpoints are taken as the norm, and who will be turned into something Other [26]? Their purpose is not limited to the inquisitive. These "who questions" double as a mode of reflection. They are a tool for people to interrogate their responsibilities and influence [15, 26]. Such a process of identification and reflection certainly does not guarantee accuracy in predicting the impact of data-driven decisions on our digital futures. Still, we might succeed in identifying the underlying norms and assumptions that co-shape their narrative and structure [64]. While these singular acts of reflection help, cultures of reflexivity should be promoted. Within cultures of reflexivity, the impact of data-driven decisions is assessed across actors and institutions rather than viewed as the single responsibility of a select few individuals or departments. While not a panacea, when social justice concerns become shared, it might enable collective acts of resistance in the face of injustice. The latter could prove vital, especially in areas where lawmakers, industry leaders or other public exemplars - whose influence, and hence responsibility is more significant - fail to take up their social responsibility. In asking these questions, we not only reflect on relationships of oppression and domination in the strict sense. Instead, they enable us to imagine who, within a particular decision-making context, is subject to experience relative disadvantage and, hence, who is more likely to become more vulnerable on a social and economic level. And while a society might not do without decisions that advantage some over others, we can at least question the conditions under which those decisions have been made. The next step, then, is to ensure each person has equal ability to exercise agency as to the digital society's direction and the impact this has on their life's prospects and considered goals, which entails a democratization of some sort to pushback against data inequalities.

Because data-driven decisions co-shape the basic structure of society, their design and deployment should be informed by a plurality of voices, perspectives and experiences [97]. In general, this norm of inclusion entails political equality and freedom from domination. For Young, each person should thus have an "equal right and effective opportunity to express their interests and concerns" regarding the rules that govern their living environment rather than have those rules unilaterally imposed onto them [97:24]. Political equality opposes legislation in which the general rules applicable to data-driven systems become privatized, left to the sole discretion of industry actors or outsourced to opaque standardization bodies. Even if a general set of principles is agreed institutionally, efforts should be made to maintain these egalitarian aspirations downstream, especially where singular decision-making systems have the potential to structure and constrain people's future actions and opportunities in significant ways. Because social justice claims must address the plurality of Others, democratic data governance strategies should foster political participation based on reasonableness. This reasonableness must not be understood as implicating some thin and disembodied ideal of rationality but instead constitutes a willingness for openness: open to listen, show respect for, understand, and be persuaded by Others, even when their position is communicated disorderly or through emotion, anger, hurt or passion [97]. Under conditions of social inequality, however, claims of inclusivity and participation are prone to abuse. For example, a call for diversity should not be co-opted and understood as one

for more data collection [5]. The same applies to participatory design strategies, which could effectively allow affected persons to exercise their voice and choice. Sloane and others have warned against participation-washing [85]. Even if the input and lived experiences of those invited meaningfully inform the ideation and design of a product or system, these efforts might remain unrewarded and unrecognized. Or, as Arnstein observes, participation might be a pretext for those in power to disregard but declare that all perspectives have been considered [3, 10]. These practices render participatory strategies extractive and exploitative. Bridging the virtues (and risks) of participation to those of reflexivity, Birhane and others offer a set of guiding questions for cooperative tech development to ensure participation is not transactional but reciprocal, vibrant and constant, allows room for disagreement, is built to increase the knowledge and empowerment of affected persons and communities [10]. Importantly, when negotiating tech, we should be open to the possibility that certain systems should not be built [10, 101]. As to how these discussions ought to be constructed, Young warns against placing transcendental goals as the subject of the debate. For instance, AI, it is often claimed, should be directed toward the common or public good. Talk of shared ideals, however, creates the impression that people's lived experiences can be suspended in favour of some generalized interest. Agreement must thus be found in, rather than outside, conflict and diversity [66].

Though cooperative design strategies offer a contextual and situated space where diverse needs can be discussed and addressed, their outcomes typically remain local, aimed to inform a particular decision-making context (e.g. automated hiring) or a particular phase of the decision-making chain (e.g., the design stage). The same applies even more to technological design strategies, such as formal fairness metrics, whose use remains limited to course-correct insular resource allocation problems [51]. While we cannot do without an evaluation of the impact specific decision-making systems exert, we should also examine the influence socio-technical structures exert as structures. At the same time, given the complexity of the digital environment, we should not expect people to individually assess the impact data-driven systems might have on their future actions and interactions and decide what course of regulatory action would be desirable. Of course, increased transparency and digital literacy will go a long way in resourcing people to regain control over their surroundings. Yet, even if people do possess the requisite faculties, they might not have the time to participate [73]. Additionally, as the digital landscape is characterized by significant asymmetries in political power, it would be foolish to expect reflexivity and democratization to appear wilfully. To overcome the deficits just described, we must rely on the regulatory support of public institutions. First, legal initiatives should further enhance people's collective ability and facilitate their coordination and organization through civil society initiatives, where others can and will listen to their lived experiences. Such spaces of knowledge building can be harnessed to understand better technology's societal push and pull and, in turn, mobilized to counter power imbalances and hold those who control the modern means of interpretation and communication accountable. To foster knowledge sharing and public accountability, lawmakers should enforce documentation and publicity regarding the "who's" of decision-making and enable their contestability. In addition, those in power should be willing to listen to the right people at the right time [79]. For example, in 2023, the LAION-5B machine learning dataset was taken down after reports emerged that it contained harmful content [86], but similar warnings had already been issued and reported on by Birhane and others in 2021 [11]! Likewise, publicity enables external actors to assess whether claims of inclusion, diversity, and human rights are truthful rather than lip service. Though collective mobilization must be encouraged, individual contestation mechanisms remain a valuable asset in people's regulatory toolbox, among others, because individuals might not always identify themselves with the viewpoints or perspectives of groups they are a member of. Though this contribution cannot address tensions that might arise between the individual and the collective, a digital society needs tools so both individuals, whether as individuals or as members of social groups or aggregates, and groups can have their interests represented.

5 CONCLUDING REMARKS

Data-driven systems have the potential to reproduce and restructure social relationships in ways that can either promote or limit people's capacity for self-development and self-determination. Fortunately, the relationships people maintain are not static; they are fluid and subject to change, and therefore, it is possible to transform them as long as many individuals, in concert, take up the responsibility to make this change possible. Some agents and entitites hold more institutional and infrastructural power than others. As they have a greater ability to co-shape society's future narratives, they are the rightful subject of increased regulation. At the same time, citizens should acknowledge that using their services might contribute to systems that benefit the interests, values, democratic norms, rights, and freedoms of some while diminishing those of others. While some hold greater moral and legal responsibility, anyone who sustains socio-technical structures and systems that produce digital oppression and vulnerabilities is obligated to transform the status quo [51, 100]. Efforts to transform society, Young contends, shall always remain a struggle [97]. The AI governance debate will cause friction. Even if legislators can create meaningful spaces of public contestability, people and interest groups will still struggle to get their views across. Under conditions of social inequality, oppression, domination and digital vulnerability, however, those views would never make it to the surface in the first place.

ACKNOWLEDGMENTS

This contribution is based in part on the lead author's doctoral research "Fair or Unfair Differentiation? Reconsidering the Concept of Equality for the Regulation of Algorithmically Guided Decision-Making." I want to thank Plixavra Vogiatzoglou, Ljubiša Metikoš, Linda Weigl, and the anonymous reviewers for their valuable feedback, which has contributed to the quality of this manuscript. This manuscript has been informed by discussions held with Marijn Sax, Michael Veale and Natali Helberger, as well as the participants of the NoBIAS Doctoral School. This publication has been supported via the AI, Media & Democracy Lab - Dutch Research Council project number: NWA.1332.20.009.

RESEARCH POSITIONALITY AND ADVERSE IMPACT STATEMENT

The main author is a European citizen who, at the moment of writing, works and resides in the Netherlands. He is an interdisciplinary scholar with a background in fundamental rights and information law, whose work is strongly rooted in political philosophy and legal theory. This contribution draws on prior research into social inequality and discrimination in the digital society, as well as the authors personal views and concerns regarding (structural) social injustice. It is the author's intention to build on, strengthen, complement, engage with, and ultimately enrich ongoing digital justice efforts. These arguments can be misconstrued or misinterpreted, or fail to reach their desired impact, that is, to increase the voice and choice of those often unheard and unrecognized. However, insofar as this contribution represents the interpretations and argumentation of the lead author, others are invited to critically assess claims made in this contribution, including their repercussions for the field of fair machine learning, and policy making in general. Drawing from real-world examples, the research conducted for this work was primarily theoretical in nature. To better capture the societal consequences of AI, future research would benefit from more integrated engagement with affected groups, and marginalised and underrepresented communities especially.

REFERENCES

[1] Amy Allen. 2008. Power and the Politics of Difference: Oppression, Empowerment, and Transnational Justice. *Hypatia* 23, 3 (2008), 156–172. https://doi.org/10.1111/j.1527-2001.2008.tb01210.x

[2] Elizabeth S. Anderson. 1999. What Is the Point of Equality? *Ethics* 109, 2 (1999), 287–337. https://doi.org/10.1086/233897

[3] Sherry R. Arnstein. 1969. A Ladder Of Citizen Participation. *Journal of the American Institute of Planners* 35, 4 (July 1969), 216–224. https://doi.org/10.1080/01944366908977225

[4] Agathe Balayn and Seda Gurses. 2021. *Beyond Debiasing: Regulating AI and its inequalities*. EDRi.

[5] Solon Barocas. 2019. Machine Learning Is a Co-opting Machine. *Public Books*. Retrieved February 15, 2023 from https://www.publicbooks.org/machine-learning-is-a-co-opting-machine/

[6] Siri Beerends and Ciano Aydin. 2021. Negotiating Authenticity in Technological Environments. *Philos. Technol.* 34, 4 (December 2021), 1665–1685. https://doi.org/10.1007/s13347-021-00480-5

[7] Marvin van Bekkum and Frederik Zuiderveen Borgesius. 2021. Digital welfare fraud detection and the Dutch SyRI judgment. *European Journal of Social Security* (August 2021), 13882627211031257. https://doi.org/10.1177/13882627211031257

[8] Reuben Binns, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 2018. "It's Reducing a Human Being to a Percentage"; Perceptions of Justice in Algorithmic Decisions. In *Conference on Computer-Human Interaction*, 2018. ACM Press, New York, New York, USA, 1–14. https://doi.org/10.1145/3173574.3173951

[9] Abeba Birhane. 2021. Algorithmic injustice: a relational ethics approach. *Patterns (N Y)* 2, 2 (February 2021), 100205. https://doi.org/10.1016/j.patter.2021.100205

[10] Abeba Birhane, William Isaac, Vinodkumar Prabhakaran, Mark Díaz, Madeleine Clare Elish, Iason Gabriel, and Shakir Mohamed. 2022. Power to the People? Opportunities and Challenges for Participatory AI. https://doi.org/10.1145/3551624.3555290

[11] Abeba Birhane, Vinay Uday Prabhu, and Emmanuel Kahembwe. 2021. Multimodal datasets: misogyny, pornography, and malignant stereotypes. https://doi.org/10.48550/arXiv.2110.01963

[12] Albert Borgmann. 1984. *Technology and the Character of Contemporary Life: A Philosophical Inquiry*. University of Chicago Press, Chicago.

[13] Anna Carastathis. 2013. Basements and Intersections. *Hypatia* 28, 4 (2013), 698–715.

[14] Mark Coeckelbergh. 2011. Human development or human enhancement? A methodological reflection on capabilities and the evaluation of information technologies. *Ethics Inf Technol* 13, 2 (June 2011), 81–92. https://doi.org/10.1007/s10676-010-9231-9

[15] Patricia Hill Collins. 1990. *Black feminist thought: knowledge, consciousness, and the politics of empowerment*. Unwin Hyman, Boston.

[16] Patricia Hill Collins. 1998. *Fighting Words: Black women and the search for justice*. University of Minnesota Press, Minneapolis. Retrieved February 7, 2022 from https://www.upress.umn.edu/book-division/books/fighting-words

[17] Nick Couldry and Ulises A. Mejias. 2019. Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. *Television & New Media* 20, 4 (May 2019), 336–349. https://doi.org/10.1177/1527476418796632

[18] Kate Crawford. 2013. The Hidden Biases in Big Data. *Harvard Business Review*. Retrieved October 14, 2019 from https://hbr.org/2013/04/the-hidden-biases-in-big-data

[19] Kate Crawford. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press. https://doi.org/doi:10.12987/9780300252392

[20] Kate Crawford and Trevor Paglen. 2021. Excavating AI: the politics of images in machine learning training sets. *AI & Soc* 36, 4 (December 2021), 1399–1399. https://doi.org/10.1007/s00146-021-01301-1

[21] Kathleen Creel and Deborah Hellman. 2021. The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision Making Systems. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (*FAccT '21*), March 03, 2021. Association for Computing Machinery, New York, NY, USA, 816. https://doi.org/10.1145/3442188.3445942

[22] John Danaher. 2022. Freedom in an Age of Algocracy. In *The Oxford Handbook of Philosophy of Technology*, Shannon Vallor (ed.). Oxford University Press, 249–272. https://doi.org/10.1093/oxfordhb/9780190851187.013.16

[23] Sylvie Delacroix and Michael Veale. 2020. Smart technologies and our sense of self: Going beyond epistemic counter-profiling. *Life and the Law in the Era of Data-Driven Agency* (January 2020). Retrieved April 5, 2020 from https://www.elgaronline.com/view/edcoll/9781788971997/9781788971997.00011.xml

[24] Emily Denton, Alex Hanna, Razvan Amironesei, Andrew Smart, and Hilary Nicole. 2021. On the genealogy of machine learning datasets: A critical history of ImageNet. *Big Data & Society* 8, 2 (July 2021), 20539517211035955. https://doi.org/10.1177/20539517211035955

[25] Catherine D'Ignazio and Lauren F. Klein. 2020. 4. "What Gets Counted Counts." *Data Feminism* (March 2020). Retrieved December 8, 2023 from https://data-feminism.mitpress.mit.edu/pub/h1w0nbqp/release/3

[26] Catherine D'Ignazio and Lauren F. Klein. 2020. *Data feminism*. The MIT Press, Cambridge, Massachusetts. Retrieved from https://mitpress.mit.edu/books/data-feminism

[27] Virginia Dignum. 2022. Relational Artificial Intelligence. *arXiv:2202.07446 [cs]* (February 2022). Retrieved February 17, 2022 from http://arxiv.org/abs/2202.07446

[28] Taylor Dotson. 2014. Authentic Virtual Others? The Promise of Post-Modern Technologies. *AI and Society* 29, 1 (2014), 11–21. https://doi.org/10.1007/s00146-012-0435-x

[29] W. E. B. Du Bois. 2007. *The souls of Black folk*. Oxford University Press, Oxford [England]; New York.

[30] Ronald Dworkin. 1977. *Taking Rights Seriously*. Harvard University Press, Cambridge.

[31] Ronald Dworkin. 1983. Comment on Narveson: In defense of equality: Ronald Dworkin. *Social Philosophy and Policy* 1, 1 (1983), 24–40.

[32] Virginia Eubanks. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press, New York. Retrieved from https://dl.acm.org/doi/book/10.5555/3208509

[33] Angelina Fisher and Thomas Streinz. 2021. *Confronting Data Inequality*. Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.3825724

[34] Carina Fourie, Fabian Schuppert, and Ivo Wallimann-Helmer (Eds.). 2015. *Social equality: on what it means to be equals*. Oxford University Press, Oxford; New York.

[35] Nancy Fraser. 1997. *Justice Interruptus: Critical Reflections on the "Postsocialist" Condition*. Routledge, New York.

[36] Marilyn Frye. 2000. Oppression. In *Encyclopedia of Feminist Theories*, Lorraine Code (ed.). London & New York: Routledge, 370.

[37] Iason Gabriel. 2022. Towards a Theory of Justice for Artificial Intelligence. *Daedalus* 151, 2 (forthcoming spring 2022), 12.

[38] Raphaël Gellert, Marvin van Bekkum, and Zuiderveen Borgesius. 2021. EU Law Analysis: The Ola & Uber judgments: for the first time a court recognises a GDPR right to an explanation for algorithmic decision-making. *EU Law Analysis*. Retrieved December 22, 2023 from http://eulawanalysis.blogspot.com/2021/04/the-ola-uber-judgments-for-first-time.html

[39] Eileen Geo and Adi Renaldi. Deception, exploited workers, and cash handouts: How Worldcoin recruited its first half a million test users. *MIT Technology Review*. Retrieved December 21, 2023 from https://www.technologyreview.com/2022/04/06/1048981/worldcoin-cryptocurrency-biometrics-web3/

[40] Janneke Gerards and Frederik Zuiderveen Borgesius. 2022. Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence. *Colorado Technology Law Journal* 20, (June 2022), 56.

[41] Eike Gräf. 2017. When Automated Profiling Threatens Our Freedom: *European Data Protection Law Review* 3, 4 (2017), 441–451. https://doi.org/10.21552/edpl/2017/4/6

[42] Blayne Haggart and Natasha Tusikov. 2023. *The New Knowledge: Information, Data and the Remaking of Global Power*. Rowman & Littlefield, London. Retrieved

February 15, 2024 from https://rowman.com/ISBN/9781538160879/The-New-Knowledge-Information-Data-and-the-Remaking-of-Global-Power
[43] Lelia Marie Hampton. 2021. Black Feminist Musings on Algorithmic Oppression. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, March 03, 2021. ACM, Virtual Event Canada, 1–1. https://doi.org/10.1145/3442188.3445929
[44] Alex Hanna, Emily Denton, Andrew Smart, and Jamila Smith-Loud. 2020. Towards a Critical Race Methodology in Algorithmic Fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (*FAT* '20*), 2020. Association for Computing Machinery, Barcelona, Spain, 501–512. https://doi.org/10.1145/3351095.3372826
[45] Natali Helberger, Marijn Sax, Joanna Strycharz, and Hans-Wolfgang Micklitz. 2022. Choice Architectures in the Digital Economy: Towards a New Understanding of Digital Vulnerability. *J Consum Policy (Dordr)* 45, 2 (2022), 175–200. https://doi.org/10.1007/s10603-021-09500-5
[46] Mireille Hildebrandt. 2008. Defining Profiling: A New Type of Knowledge? In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, Mireille Hildebrandt and Serge Gutwirth (eds.). Springer Netherlands, Dordrecht, 17–45. https://doi.org/10.1007/978-1-4020-6914-7_2
[47] Kashmir Hill. 2020. Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match. *The New York Times*. Retrieved May 12, 2022 from https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html
[48] Anna Lauren Hoffmann. 2022. Rawls, Information Technology, and the Sociotechnical Bases of Self-Respect. *The Oxford Handbook of Philosophy of Technology*. https://doi.org/10.1093/oxfordhb/9780190851187.013.15
[49] Charlotte Huisman. 2019. SyRI, het fraudesysteem van de overheid, faalt: nog niet één fraudegeval opgespoord. *de Volkskrant*. Retrieved January 22, 2020 from https://www.volkskrant.nl/gs-b789bc3a
[50] Elle Hunt. 2016. Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter. *The Guardian*. Retrieved July 5, 2023 from https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter
[51] Atoosa Kasirzadeh. 2022. Algorithmic Fairness and Structural Injustice: Insights from Feminist Political Philosophy. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, July 26, 2022. 349–356. https://doi.org/10.1145/3514094.3534188
[52] Nicolas Kayser-Bril. Using ChatGPT for (re)search is a philosophical break – Automated Society. *AlgorithmWatch*. Retrieved January 22, 2024 from https://r.algorithmwatch.org/nl3/q_7P5fwseQoWIpcLRde4Kw
[53] Os Keyes. 2019. Counting the Countless. Retrieved January 5, 2024 from https://ironholds.org/counting-writeup/
[54] Tarunabh Khaitan. 2015. *A Theory of Discrimination Law*. Oxford University Press. Retrieved June 1, 2020 from http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199656967.001.0001/acprof-9780199656967
[55] René Kladzyk. 2023. Policing Gender: How Surveillance Tech Aids Enforcement of Anti-Trans. . .. *POGO*. Retrieved December 21, 2023 from https://www.pogo.org/investigations/policing-gender-how-surveillance-tech-aids-enforcement-of-anti-trans-laws
[56] Will Kymlicka (Ed.). 1992. *Justice in Political Philosophy: Volume 1*. Edward Elgar Publishing Limited, Cambridge.
[57] Paddy Leerssen. 2023. Seeing what others are seeing: Studies in the regulation of transparency for social media recommender systems.
[58] Sam Levin. 2016. A beauty contest was judged by AI and the robots didn't like dark skin. *The Guardian*. Retrieved August 3, 2022 from https://www.theguardian.com/technology/2016/sep/08/artificial-intelligence-beauty-contest-doesnt-like-black-people
[59] Antony Loewenstein. 2023. The Palestine Laboratory: How Israel Exports the Technology of Occupation Around the World. Verso Books.
[60] Noëmi Manders-Huits. 2010. Practical versus moral identities in identity management. *Ethics Inf Technol* 12, 1 (March 2010), 43–55. https://doi.org/10.1007/s10676-010-9216-8
[61] Surya Mattu, Jeff Larson, Lauren Kirchner, and Julia Angwin. 2016. How We Analyzed the COMPAS Recidivism Algorithm. *ProPublica*. Retrieved August 8, 2020 from https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token$=$e3k6mE9frXgAOt3zvKfhyLDrUGgrhmOb
[62] Lucas Waldron Medina Brenda. TSA's Body Scanners Are Gender Binary. Humans Are Not. *ProPublica*. Retrieved February 7, 2022 from https://www.propublica.org/article/tsa-transgender-travelers-scanners-invasive-searches-often-wait-on-the-other-side?token$=$7bjY-MRzWk5Ed4DCZRvFVYwt2HBrAFXd
[63] Milagros Miceli and Julian Posada. 2022. The Data-Production Dispositif. Retrieved April 4, 2023 from http://arxiv.org/abs/2205.11963
[64] Milagros Miceli, Tianling Yang, Laurens Naudts, Martin Schuessler, Diana Serbanescu, and Alex Hanna. 2021. Documenting Computer Vision Datasets: An Invitation to Reflexive Data Practices. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, March 03, 2021. ACM, Virtual Event Canada, 161–172. https://doi.org/10.1145/3442188.3445880
[65] Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society* 3, 2 (December 2016), 205395171667967. https://doi.org/10.1177/2053951716679679
[66] Chantal Mouffe. 1999. Deliberative Democracy or Agonistic Pluralism? *Social Research* 66, 3 (1999), 745–758.
[67] James Muldoon and Paul Raekstad. 2022. Algorithmic Domination in the Gig Economy. *European Journal of Political Theory* (February 2022), 14748851221082078. https://doi.org/10.1177/14748851221082078
[68] Laurens Naudts and Ana Maria Correa. forthcoming. Data-Driven Inequality and Discrimination: Challenges and Opportunities for Regulating AI-systems in the COE and EU. In *The Legal Consistency of Technology Regulation in Europe*, Inge Graef and Bart van der Sloot (eds.). Hart Publishing, 9–42.
[69] Laurens Naudts, Natali Helberger, Marijn Sax, and Michael Veale. 2024. Toward Constructive Optimisation: a new perspective on the regulation of recommender systems and the rights of users and society. In *Digital Fairness for Consumers*, Natali Helberger, Betül Kas, Hans-W Micklitz, Monika Namysłowska, Laurens Naudts, Marijn Sax and Michael Veale (eds.). BEUC - The European Consumer Organisation, Brussels, 280.
[70] Laurens Naudts and Anton Vedder. Forthcoming. Fairness and Artificial Intelligence. In *Cambridge Handbook on the Law, Ethics and Policy of AI*, Nathalie A. Smuha (ed.). Cambridge University Press.
[71] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press, New York. Retrieved August 6, 2019 from https://www.jstor.org/stable/j.ctt1pwt9w5
[72] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press. https://doi.org/10.18574/9781479833641
[73] Jonathan A Obar. 2015. Big Data and 'The Phantom Public': Walter Lippmann and the fallacy of data privacy self-management. *Big Data & Society* 2, 2 (December 2015), 205395171560887. https://doi.org/10.1177/2053951715608876
[74] Billy Perrigo. 2023. The $2 Per Hour Workers Who Made ChatGPT Safer. *Time*. Retrieved July 5, 2023 from https://time.com/6247678/openai-chatgpt-kenya-workers/
[75] Philip Pettit. 1980. *Judging Justice. An Introduction to Contemporary Political Philosophy*. Routlegde & Kegan Paul, London, Boston and Henley.
[76] Philip Pettit. 1999. *Republicanism: A Theory of Freedom and Government*. Oxford University Press. Retrieved February 29, 2020 from https://www.oxfordscholarship.com/view/10.1093/0198296428.001.0001/acprof-9780198296423
[77] Philip Pettit. 2012. *On The People's Terms: A Republican Theory and Model of Democracy*. Cambridge University Press, Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paolo, Delhi, Mexico City.
[78] Bernhard Rieder and Jeanette Hofmann. 2020. Towards platform observability. *Internet Policy Review* 9, 4 (December 2020). Retrieved July 19, 2023 from https://policyreview.info/articles/analysis/towards-platform-observability
[79] Eryk Salvaggio. 2024. LAION-5B, Stable Diffusion 1.5, and the Original Sin of Generative AI | TechPolicy.Press. *Tech Policy Press*. Retrieved January 3, 2024 from https://techpolicy.press/laion5b-stable-diffusion-and-the-original-sin-of-generative-ai
[80] Marijn Sax. 2021. *Between empowerment and manipulation: The ethics and regulation of for-profit health apps*. Retrieved September 18, 2023 from https://dare.uva.nl/search?identifier$=$52225d37-e7e1-4883-9dab-a3f5d3a063d8
[81] Kevin Schaul, Szu Yu Chen, and Nitasha Tiku. Inside the secret list of websites that make AI like ChatGPT sound smart. *Washington Post*. Retrieved January 22, 2024 from https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning/
[82] Samuel Scheffler. 2003. What Is Egalitarianism? *Philosophy & Public Affairs* 31, 1 (2003), 5–39.
[83] Christian Schemmel. 2011. Distributive and relational equality: *Politics, Philosophy & Economics* (September 2011). https://doi.org/10.1177/1470594X11416774
[84] Fabian Schuppert. 2015. Non-domination, non-alienation and social equality: towards a republican understanding of equality. *Critical Review of International Social and Political Philosophy* 18, 4 (July 2015), 440–455. https://doi.org/10.1080/13698230.2015.1033863
[85] Mona Sloane, Emanuel Moss, Olaitan Awomolo, and Laura Forlano. 2022. Participation Is not a Design Fix for Machine Learning. In *Equity and Access in Algorithms, Mechanisms, and Optimization*, October 06, 2022. ACM, Arlington VA USA, 1–6. https://doi.org/10.1145/3551624.3555285
[86] David Thiel. 2023. Identifying and Eliminating CSAM in Generative ML Training Data and Models. (2023). https://doi.org/10.25740/kh752sm9123
[87] Tommy Tobin. 2018. Health Insurance Coverage For Healthy Groceries? More Food-Based Interventions May Be Coming. *Forbes*. Retrieved January 3, 2024 from https://www.forbes.com/sites/tommytobin/2018/05/24/health-insurance/
[88] Sherry Turkle. 2007. Authenticity in the age of digital companions. (2007).
[89] Anton Vedder. 1999. KDD: The challenge to individualism. *Ethics and Information Technology* 1, 4 (1999), 275–281.
[90] Peter-Paul Verbeek. 2006. Materializing Morality: Design Ethics and Technological Mediation. *Science, Technology, & Human Values* 31, 3 (May 2006), 361–380. https://doi.org/10.1177/0162243905285847

[91] Tim Verheyden. 05:22+02:00. Is het echt zo dat enkel een chatbot een man heeft aangezet tot zelfdoding? *vrtnws.be*. Retrieved December 20, 2023 from https://www.vrt.be/vrtnws/nl/2023/04/04/chatbot/
[92] Koen Vervloesem. 2020. How Dutch activists got an invasive fraud detection algorithm banned. *AlgorithmWatch*. Retrieved October 16, 2020 from https://algorithmwatch.org/en/story/syri-netherlands-algorithm/
[93] Salomé Viljoen. 2021. A Relational Theory of Data Governance. *the yale law journal* (2021), 82.
[94] Sandra Wachter. 2023. The Theory of Artifical Immutability: Protecting Algorithmic Groups Under Anti-Discrimination Law. *Tulane Law Review* 97, 2 (March 2023), 149.
[95] Iris Marion Young. 1990. *Justice and the Politics of Difference*. Princeton University Press, Princeton, New Yersey.
[96] Iris Marion Young. 2001. Equality of Whom? Social Groups and Judgments of Injustice. *Journal of Political Philosophy* 9, 1 (March 2001), 1–18. https://doi.org/10.1111/1467-9760.00115
[97] Iris Marion Young. 2002. *Inclusion and Democracy*. Oxford University Press. https://doi.org/10.1093/0198297556.001.0001
[98] Iris Marion Young. 2003. Political Responsibility and Structural Injustice. May 05, 2003. The University of Kansas, 25.
[99] Iris Marion Young. 2006. Taking the Basic Structure Seriously. *PPS* 4, 01 (March 2006). https://doi.org/10.1017/S1537592706060099
[100] [100] Iris Marion Young. 2011. *Responsibility for Justice*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195392388.001.0001
[101] [101] Annette Zimmermann and Chad Lee-Stronach. 2021. Proceed with Caution. *Canadian Journal of Philosophy* (2021), 1–20. https://doi.org/10.1017/can.2021.17