JOURNAL OF E-LEARNING AND KNOWLEDGE SOCIETY
Vol. 18, No. 3 (2022), pp. 83-95

# Digital marginalization, data marginalization, and algorithmic exclusions: a critical southern decolonial approach to datafication, algorithms, and digital citizenship from the Souths

Chaka Chaka[a,1]

[a]*University of South Africa, Dept. of English Studies* – Pretoria (South Africa)

*(submitted: 3/5/2022; accepted: 14/11/2022; published: 31/12/2022)*

**Abstract**

This paper explores digital marginalization, data marginalization, and algorithmic exclusions in the Souths. To this effect, it argues that underrepresented users and communities continue to be marginalized and excluded by digital technologies, by big data, and by algorithms employed by organizations, corporations, institutions, and governments in various data jurisdictions. Situating data colonialism within the Souths, the paper contends that data ableism, data disablism, and data colonialism are at play when data collected, collated, captured, configured, and processed from underrepresented users and communities is utilized by mega entities for their own multiple purposes. It also maintains that data coloniality, as opposed to data colonialism, is impervious to legal and legislative interventions within data jurisdictions. Additionally, it discusses digital citizenship (DC) and its related emerging regimes. Moreover, the paper argues that digital exclusion transcends the simplistic haves versus the have nots dualism as it manifests itself in multiple layers and in multiple dimensions. Furthermore, it characterizes how algorithmic exclusions tend to perpetuate historical human biases despite the pervasive view that algorithms are autonomous, neutral, rational, objective, fair, unbiased, and non-human. Finally, the paper advances a critical southern decolonial (CSD) approach to datafication, algorithms, and digital citizenship by means of which data coloniality, algorithmic coloniality, and the coloniality embodied in DC have to be critiqued, challenged, and dismantled.

**KEYWORDS:** Digital Citizenship, Digital Marginalization, Data Marginalization, Algorithmic Exclusions, Data Colonialism, Critical Southern Decoloniality.

**DOI**
https://doi.org/10.20368/1971-8829/1135678

**CITE AS**
Chaka, C. (2022). Digital marginalization, data marginalization, and algorithmic exclusions: a critical southern decolonial approach to datafication, algorithms, and digital citizenship from the Souths. *Journal of e-Learning and Knowledge Society*, 18(3), 83-95. https://doi.org/10.20368/1971-8829/1135678

## 1. Introduction

This paper explores how communities (also regarded as users or citizens) from the Souths, especially indigenous, subaltern, and underprivileged communities, tend to be marginalized by big data in all its multiple digital configurations [The other co-references of these indigenous, subaltern, and underprivileged communities in the paper are Black, Indigenous and People of Color (BIPOC) communities, Southern societies, and societies in the Souths. A further co-reference of these communities is *the others* even though this co-reference has not been used in this paper]. It also examines how such communities in the Souths often get excluded by algorithms through their multifarious uses. Those who collect, collate, capture, configure, process, and preserve data do so for various purposes: advertising, tracking, monitoring, surveillance, credit control, population census, and decision making. To this end, there are different types of data. All of these data processes get passed off as big data and datafication (Andrejevic, 2014; Charitsis & Lehtiniemi, 2022; Milan & Treré, 2019, 2021; Van Dijck, 2014). Central to collecting, collating, capturing, configuring, processing, and preserving data, to the purposes that data serve, and to big data and datafication, are algorithms. That is, data has to be *big* for it to undergo these data processes and for it to be subjected to algorithms. If it is not big, it cannot be data, and it cannot have data infrastructure (cf. Milan & Treré, 2021). If it is not big, it is worthless and unusable.

Data, particularly big data, benefit societies that are data- and digitally-savvy, and that have unlimited access

[1] corresponding author - email: chakachaka8@gmail.com

to such data and to associated technologies. Such societies are, however, confined to the Norths – in their diverse and multiple configurations (see Milan & Treré, 2019) – and not to the Souths – also in their diverse and multiple configurations (again, see Milan & Treré, 2019). In this data sphere, Northern societies tend to be more privileged in terms of big data and datafication than Southern societies as characterized above. This means that for societies in the Norths, big data and datafication imply empowerment and affirmation, what Charitsis and Lehtiniemi (2022) refer to as data *ableism*. Conversely, for societies in the Souths, especially Black, Indigenous and People of Color (BIPOC) (Terp, 2020) communities, big data and datafication entail marginalization and exclusion from digital citizenship if data is deemed to be a passport to being a citizen in the digitally datafied world. Charitsis and Lehtiniemi (2022) regard this state of affairs as data *disablism*, while Lerman (2013) refers to it as a *perspective of exclusion*.

In a similar vein, the deployment of algorithms in harvesting and mediating big data has spawned a parallel process in which Northern societies are privileged and affirmed by algorithmic inclusions as they serve as a model society for machine learning, while Southern societies tend to be disadvantaged and disaffirmed by algorithmic exclusions as they are a non-model society for machine learning. A corollary of this is that, right from the onset, in an increasingly automated world and in a world where what Janssen and Kuk (2016) call Big and Open Linked Data (BOLD) is readily available, Southern societies are denied digital citizenship by algorithmic exclusions even if they were all to be data- and digitally-savvy. This is a consequential issue as who gets excluded by both big data and algorithms has their life chances negatively impacted by exclusions perpetuated by automated algorithms. This is also a problematic issue as those whose data is harvested and utilized by algorithms have no control and decision-making capacity over how their data is used. The point here is, as Janssen and Kuk (2016) pertinently argue, even though algorithms are thought to belong to the domain of computer programming, they nonetheless percolate into social and economic spheres. In fact, there is no gainsaying that big data and algorithms have almost colonized the life worlds of modern-day, automated societies, wherever their locations are.

Against the background sketched above, this paper has the following sections: the Souths and data colonialism; digital citizenship and emerging regimes of digital citizenship; digital marginalization, data marginalization, and algorithmic exclusions; and critical southern decolonial approach to datafication, algorithms, and digital citizenship.

## 2. The Souths and data colonialism

The phrase, *the Souths*, builds on and departs from the Global South, whose counterpoise is the Global North. It builds on the Global South in line with how the latter (the Global South) has been conceptualized and utilized by different scholars from various disciplines (see, for example, Benabdallah et al., 2017; Chaka, 2020; Clarke, 2018; Dados & Connell, 2012; Kloß, 2017; Lazar, 2020; Mahler, 2018; Milan & Treré, 2019; Wolvers et al., 2015). Sometimes, tying down a concept to a definition yields the opposite: prescriptiveness, essentialism, and definitional opaqueness. This is particularly the case with concepts such the Global South and the Souths. In trying to define them, one may end up being prescriptive, essentializing them, or making them appear more opaque. So, there is no one straightforward or no one-size-fits-all definition that can be attached to these two concepts. Rather, definitional perspectives from which these terms are conceptualized are more helpful in this context. Concerning the Global South, and without delving deeper into a historical evolution of the term, three definitional perspectives are relevant [for the historical evolution of the term, the Global South, see Clarke (2018), Dados and Connell (2012), Kowalski (2020), and Mahler (2017)]. First, it is a metaphor for countries characterized by persistent inequalities and asymmetrical power relations owing to imperialism and neocolonialism, irrespective of their spatial locations. Second, it refers to global subaltern communities (subjugated peoples and Indigenous Peoples), irrespective of their geographical locations, whose knowledges are often marginalized by the Global North. This is an equivalent of what Kloß (2017) calls *the global peripheries*. Third, it refers to countries whose economies are less developed when compared with those of the countries situated in the Global North (see Dados & Connell, 2012; Kloß, 2017; Lazar, 2020; Milan & Treré, 2019).

As noted from the three foregoing definitional perspectives, there is a sense of reductionism and essentialism about them: reducing the countries and the peoples deemed to belong to the Global South to a homogeneous whole and to subalternity, and essentializing them territorially, racially, demographically, and economically along that reductionist axis. Additionally, there is a sense of romanticizing about the countries and their peoples in the Global South, and about the latter itself: that these countries and their peoples are ideal polities situated in the ideal spatial locality (the Global South). Owing to this, the paper prefers to use the phrase, *the Souths*, to refer to the Global South. It does so in keeping with Mahler's (2017) and Milan and Treré's (2019) view of the Souths (also see Armillas-Tiseyra & Mahler, 2021). In its plural form, the Souths, conceptually and metaphorically, signals the multiplicities of the Souths across the geographic globe. It also signifies the heterogeneity, the diversity, the severalness, and the situatedness of the countries and the peoples of the Souths: they are Indigenous Peoples; they are the erstwhile colonized; they are the subaltern and the *othered* peoples; they are the peoples whose knowledge systems are marginalized; they are the peoples with varying underdeveloped and developing economies and with limited and varying access to the fourth industrial revolution (4IR) technologies; and they are the peoples defined by, as pointed out by Francis (2021), their positionalities relative to global capitalism. Moreover, they are countries that have differing pockets of the Norths in them. Mahler (2017) aptly captures this geography-defying, and sometimes nation-state-incompatible Souths and Norths by asserting that: “there are economic Souths in the geographic North and Norths in the geographic South” (p. 32). Likewise, Francis (2021, p. 689) avers that “there are Global Souths in the geographic North and Global Norths in the geographic South”. This is a deterritorialized conception of the term, *the Souths* together with its alter ego, *the Norths*.

Alongside the Souths are the notions of the big data from the Souths and data colonialism. Data colonialism is the practice of unilaterally extracting computational data from mainly the Souths that parallels predatory extractive and exploitative tendencies of geopolitical colonialism (see Couldry and Mejia, (2019a) to which the majority of the Souths were subjected. It is, on the one hand, one of the salient features of global colonialism that is aided and mediated by digital technologies. On the other hand, it is a central part of datafication in which big data algorithms are indispensable technologies. Datafication itself has to do with the pervasive digital harvesting and use of big data and its impact (Heeks & Shekhar, 2019) on the social life worlds of individuals in a borderless digital world. Elsewhere, Ricaurte (2019) and Zembylas (2021) talk about digital colonialism and digital neocolonialism, respectively. Data colonialism is part of this overarching neocolonialism.

In this context, big data from the Souths refers to different forms of big datasets digitally harvested or extracted from the Souths and how these datasets are appropriated and exploited by the Norths in a manner similar to historical and geopolitical colonialism (cf. Couldry & Mejia, 2019a; Milan & Treré, 2019; Mumford, 2021). Driven mainly by big corporations and big money from the Norths – a proxy for global capitalism – data colonialism has as one of its key purposes extracting and aggregating big data for profit-making and for other purposes that serve the various digital dividends for such big corporations and big money. One of these digital dividends is marketing, commodifying, and monetizing personal data and behavioral data. Another digital dividend is surveilling individuals by governments or by private entities, a data practice that Greenwood (2020), Hintz et al. (2019), and Zuboff (2019) refer to as *surveillance capitalism*, and which Van Dijck (2014) calls *dataveillance*.

## 3. Digital citizenship and emerging regimes of digital citizenship

Conventionally, digital citizenship (DC) refers to a situation in which users of digital technologies are presumed to possess a wide range of skills that enable them to competently, progressively, and critically engage with and use such technologies. This includes the ability to meaningfully participate, learn, socialize, work, play, and communicate in diverse digital environments (Richardson & Milovidov, 2019). DC comprises the following nine features: digital access; digital literacy; digital communication; digital commerce; digital health; digital law; digital ethics; digital rights and duties; and digital security (Mukhametzyanov, 2022). It also consists of two features: state organization and citizens’ self-organization; and human activity. State organization refers to the manner in which the state organizes and regulates digital technologies in its jurisdiction, and the type of access to digital environments it allows its citizens to have. Self-organization has to do with how citizens leverage digital technologies as individuals and as groups of people. For its part, human activity is related to citizens’ digital presence and the digital traces citizens leave online. Additionally, this has to do with whether the digital presence and the attendant online traces of citizens are anonymized or not. This latter point is crucial as citizens’ digital presence is a virtual copy of citizens’ persona. Their presence and traces online serve as a gateway to and as a source of their personal data that is tracked, used, and surveilled by the state and corporations for various purposes. In this way, citizens’ online presence and traces are also part of their digital footprint. Most importantly, DC is tied to citizens’ digital literacy, digital rights, and digital freedom (Mukhametzyanov, 2022; also see Richardson & Milovidov, 2019).

Viewed from another perspective articulated by Richardson and Milovidov (2019), DC is, as represented in a temple-like model, underscored by four sets of competences, five framing pillars, and ten DC domains. The four sets of competences are as follows:

- *Values* – democracy, fairness, equality, human rights, and cultural diversity
- *Attitudes* – openness, self-efficacy, civil mindedness, respect, tolerance, and responsibility
- *Skills* – communication, plurilingual skills, listening, observing, cooperation, empathy, flexibility, adaptability, autonomous learning, analytical and critical thinking skills, and conflict resolution
- *Knowledge and critical understanding* – knowledge and critical understanding of: self, language and communication, cultures, politics, human rights, religions, law, media, environment, and sustainability.

These core competences, which are the foundational or bottom layer, are followed by five constructs: policies, stakeholders, strategies, infrastructures and resources, and evaluation. These constructs are regarded as framing pillars, and constitute the middle layer. At the top layer are ten DC domains. These are as follows:

- *Students* – empowering, educating, and protecting themselves
- *Parents* –participating in citizenship and Internet debate, and helping children find a balance between interpersonal and social lives when using digital technologies
- *Teachers* – upskilling in terms of digital competences and reviewing teachers' role in the digital era
- *School management* – ensuring that all the relevant stakeholders are part of a decision-making process concerning safe, ethical, and legal use of both digital technologies and digital information
- *Academia* – developing local resources to ensure maximum engagement by all stakeholders, highlighting the positive and negative implications of digital technologies and digital information; and conducting research related to DC
- *Private sector* – creating conditions conducive to effective DC; initiating a multi-stakeholder and cross-media approach to dealing with digital technologies and digital information with a view to empowering users and protecting minors; and putting in place appropriate terms and conditions that are user-centric
- *Civil sector* – providing new directions and future orientation for DC education
- *Local educating communities* – developing a framework for formal, informal, and non-formal education that speaks to DC, and initiating civic tech to respond to and address different aspects of DC
- *Regulatory authorities* – encouraging education authorities to embrace DC education, and ensuring that users' and children's rights are respected
- *National/international authorities* – promoting democratic values and human rights for multi-stakeholder consultative and governance structures (Richardson & Milovidov, 2019).

While the points attributed to DC above are crucial, the notion *DC* is very complex, especially given the nuances and challenges associated with it. This is more so given the rapidity with which digital technologies evolve and the new ones come into play; and also given the fact that digital environments are ever-changing minefields in terms of user data, which feeds into datafication. In this case, acquiring digital competences, which include digital literacy, and knowing about one's digital rights and privacy protection are not sufficient safeguards against an unauthorized use of personal data, or against a nefarious use of such data. For this reason, it is equally crucial for users to acquire critical digital literacies, critical technology education (Pötzsch, 2019), and critical data literacies, and to develop an understanding of critical data infrastructure literacies (cf. Chaka, 2019; Gray et al., 2018; Pötzsch, 2019) and datafication. Nonetheless, all of this becomes tricky and challenging when users are children.

Reflecting on the complexity of DC in the ever-evolving digital age and taking into consideration the era of the COVID-19 pandemic and its ramifications both on digital environments and on DC, Calzada (2022) proposes and discusses five emerging DC regimes. These DC regimes can also be taken to be the modes of DC or the personas of DC users assume in various digital environments. The main drivers of these emerging DC regimes are different digital technologies and the practice of datafication. These five emerging DC regimes are: pandemic citizenship, algorithmic citizenship, liquid citizenship, metropolitan citizenship, and stateless citizenship. Pandemic citizenship is a global, generalizable, emerging DC regime that reflects how datafication practices during and post-COVID-19 have engendered interwoven, techno-politically and city-regionally driven, unique DC regimes in certain urban parts of European nation-states.

In this context, algorithmic citizenship is mainly powered by big data algorithms. Similarly, liquid citizenship is driven by the big data ideology or dataism. For its part, metropolitan citizenship is powered by data cooperatives in response to Brexit. Finally, stateless citizenship is driven by data sovereignty. Even though the last two regimes of DC are specific to conditions related to European nation-states, and by extension to the Norths, the first three regimes have applicability to other nation-states, including those in the Souths. This is especially so when taking into account the datafication and algorithmization that are key in mediating these DC regimes. Another factor to note is that these regimes of DC are not necessarily mutually exclusive: they can overlap and co-exist within one user. Their interlocking reflects how intricate living digitally can be in the face of what Calzada (2022) calls *algocracy*, *algorithmic surveillance*, *dataveillance*, and *digital panopticon* (also see Floridi, 2020; Geeker and Hind, 2019). Moreover, it highlights how DC is inextricably linked to data citizenship, with the latter underscoring the need for users to display active and critical agency when online, particularly when datafication and algorithmization have become so naturalized and normalized (see Pawluczuk et al., 2020).

## 4. Digital marginalization, data marginalization, and algorithmic exclusions

Digital users can and do get marginalized when online, when accessing digital environments, or due to lack of access to digital technologies and to the Internet connectivity. This constitutes, the paper contends, digital marginalization. Digital marginalization entails digital exclusion and discrimination. For example, Gangadharan (2021) maintains that digital exclusion is linked to issues about Internet infrastructure access, Internet technologies adoption, marginalization caused by socio-economic conditions and forms of historical oppression (also see Martin et al., 2016; Tomczyńska, 2017). All of these factors have a significant bearing on whether or not citizens have a meaningful and active participation in digital technologies, or whether or not they have a meaningful and active digital participation. According to Tomczyńska (2017), digital exclusion, whose origin he traces to the United States, has much to do with information-poor societies versus information-rich societies, or with information *have-nots* versus information *haves*. It is an equivalent of an erstwhile digital divide. Nonetheless, as Tomczyńska (2017) points out, this dualism tends to oversimplify a very complex phenomenon. This oversimplified dualism is rooted in technological determinism that views digitality and digitalization [digitality refers to a condition in which everything a user does (communicating, writing, purchasing, creating content, etc.) happens exclusively online through digital technologies (see Fund, 2022). Segura and Waisbord (2019) calls it digitalism. For its part, digitalization is a process in which a user's social life domains, and the information related to such domains, are structured around and mediated by digital communication and media infrastructure. In it, social interactions such as work and leisure occur solely on digital platforms as opposed to analog platforms. It also relates to an environment in which business operations happen on digital platforms, thereby blurring the physical and digital worlds (see Bloomberg, 2018)] in terms of haves and have nots, while ignoring the factors engendering this binarism such as political, social, racial, cultural, educational, economic, institutional, infrastructural, geographical, and ideological factors. These factors are not binary, but multilayered, multidimensional factors; they are also inextricably intertwined, and have embedded or underlying subsets.

At a more intricate and nuanced level, digital exclusion transcends the binarism and both the multilayerism and multidimensionality portrayed above. For example, it can occur at the level of what Sin et al. (2021) call digital design marginalization. The latter refers to a situation in which certain digital interface designs are configured in such a way as to exclude particular users, especially underrepresented users such as those in the Souths, thereby contributing to their being marginalized in certain aspects of their digital lives. As a result of such non-inclusive designs, these users encounter digital barriers when trying to access essential services such as shopping, healthcare, and personal finance. A similar process is the one in which underrepresented users, owing to their socioeconomic, cultural, and historical marginalization, manage to possess only low-level digital devices that are not fitted with user interface designs that can allow them to access essential services online. Another instance is the one in which certain underrepresented users may possess relevant digital devices, but may still not be able to access online essential services due to some of the marginalizing and exclusionary factors mentioned in the preceding paragraph.

Digital users can further be marginalized when their data or the data they generate online is used by various data exploiters, whenever they (users) access digital technologies and digital environments through any form of Internet connectivity. This practice engenders data marginalization. The practice is so called because users become marginalized from the very data they generate, particularly in automated and datafied societies. It is a practice that typifies data colonialism, whose central logics are algorithmizing, commodifying, and monetizing data (big and small data) within the broader datafication process as defined and discussed earlier. It is a scenario, in which, to repurpose Charitsis and Lehtiniemi's (2022) thoughts, market-driven norms and standards trump the privacy and the sanctity of personal data for marketization purposes.

It also a situation in which certain individuals and communities, especially the underrepresented communities from the Souths, get excluded and marginalized, while others, particularly the data- and digitally-rich users from the Norths, are privileged and rewarded. Herein lies the notions of *data ableism*, *data disablism* (see Charitsis & Lehtiniemi, 2022), data capitalism (Charitsis and Lehtiniemi, 2022 also refer to it as *data-based capitalism*; see Couldry and Mejias, 2019a; Segura and Waisbord, 2019), and data coloniality (Mumford, 2021). The first two concepts are, as argued by Charitsis and Lehtiniemi (2022), more than just ability (efficiency) versus disability (deficiency) and more than just tropes especially when they are viewed from both critical disability scholarship and critical technology scholarship. In line with this dual view, data ableism refers to practices, processes, and politics of data whose primary purpose is to privilege and affirm particular data-related abilities and digitalization practices that are expected in certain data subjects. These data abilities and digitalization practices include, the paper argues, the skillsets of competences and their attendant sub-skillsets possessed by mainstream digital citizens in the Norths as outlined earlier. They also include these digital citizens' digital habituses. Moreover, they relate to the five emerging DC regimes mentioned earlier into which digital citizens in the Norths are categorized.

In this case, data disablism has to do with practices, processes, and politics of data that tend to exclude and marginalize individuals and communities who are perceived to lack the requisite digital skillsets of competences and their attendant sub-skillsets as mentioned earlier, and who are deemed not to display the digital habituses often exhibited by the data-savvy digital users. Such individuals and communities are, additionally, construed as having data-based deficiencies (see Charitsis & Lehtiniemi, 2022). The corollary of the two processes, data ableism and data disablism, is the parallel processes that Charitsis and Lehtiniemi (2022) call data (in)visibility and data (un)desirability. The former refers to the ability users have to generate data that makes them visible or invisible within a data ecosystem, while the latter is related to the ability users have to generate data that is construed to be (in)valuable or (un)desirable within a data ecosystem. These two processes underscore the value and the normalizing/de-normalizing logic of data in automated, datafied societies and of the attendant data economies of such societies in the Norths. In such societies, visibility and desirability through data becomes the norm, while data invisibility and data undesirability become an aberrance. Therefore, data invisibility and data undesirability lead to data marginalization and exclusion, or to what Lerman (2013) calls a perspective of exclusion. In a data visibility/desirability - data invisibility/undesirability continuum, the first end of the continuum gets more privileged and validated than the last end of the continuum in terms of the data produced by users (cf. Charitsis & Lehtiniemi, 2022). In addition, the first end of the continuum is often associated with the automated, datafied users and communities in the Norths, whereas the last end is seen to be linked to the less automated and the less datafied users and communities in the Souths.

What is intriguing in the data colonialism, data ableism, and data disablism equation in which underrepresented individuals and communities are marginalized and excluded, is data coloniality or the coloniality of data. While data colonialism, like its alter ego, historical colonialism, can be dealt with within legal and legislative frameworks (e.g., personal data privacy rights, data protection laws, data sovereignty, and digital citizenship rights) in given data jurisdictions, data coloniality is impervious to any legal and legislative interventions. That is, as pointed out by many scholars such as Escobar (2007), Grosfoguel (2007), Hsu (2017), Maldonado-Torres (2007; 2018), Mignolo (2007), Núñez-Pardo (2020), Quijano and Ennis (2000), coloniality, unlike colonialism, persists in postcolonial jurisdictions well after colonialism has ended. It does so in multiple variants like the coloniality of power, of knowledge, of being, and of thought. To this, can be added the coloniality of data, of algorithms, and of digitality. Underscoring these three forms of coloniality is Eurocentrism, which projects Euro-American worldviews as centers of universal, objective, zero-point epistemes (Mumford, 2021) against which all subaltern knowledges can be judged and benchmarked. With reference to data and algorithmic coloniality, "the heteronormative … White … modern subject" (Mumford, 2021, p. 4) together with its gendered, racialized, and classed (Mumford, 2021) Euro-American representation is the basis of both data representation and algorithmic configuration. Human features and characteristics of non-European subjects become excluded and *peripheralized* in this data and algorithmic setup.

Over and above digital and data marginalization, there are algorithmic exclusions. At a basic level, algorithms[5] are abstract, formalized, automated, rules-based descriptions of computer procedures for processing data [algorithms are more complex than they have been presented in this paper. They do not operate only in digital gadgets, but also in large machines and in super computers, where in tandem with AI, they perform complex functions and tasks that human brains cannot ordinarily perform. For some of the examples of algorithms and their related methods and tools, especially within the educational data mining field, see Chaka (2021); Jago and Laurin (2022); also see Cofone (2019)]. They are sets of procedural steps intended to solve certain problems based on inputs and outputs that regulate and ensure the functioning of automated tasks. Simply put, they are computer programs (Borgesius, 2018; also see Orwat, n.d.). Mostly, algorithms operate through very complex and coded procedures that are invisible to users. They are often required to remotely execute coded and automated decision-making based on the types of datasets that they are assigned to collect or work on. Many of big data-driven algorithms tend to operate predictively in real-time by learning from previous and existing observations with a view to perfecting their predictions (Tenney & Sieber, 2016). Overall, algorithms have wide-ranging applications in different contexts such as generating and distributing online data, surveilling and policing citizens, carrying out employee assessments, marketing and advertising, financial and purchasing transactions, stock trading, and fraud detection (see Ulbricht & Yeung, 2022). To these algorithmic applications can be added online personal profiling, hiring, and university student admissions as well (cf. Orwat, n.d.; Ulbricht & Yeung, 2022; Williams et al., 2018). One key issue worth mentioning is that algorithms are intended to computationally optimize things: solve problems; carry out or complete tasks; and save lives. At a very innocuous level, and with the aid of artificial intelligence (AI), algorithms help different digital devices sort photos, recognize human faces; respond to voice commands; drive cars (Rainie & Anderson, 2017); personalize learning and adverts; harvest and match publications against authors; or diagnose illnesses. This includes their nefarious use in activities such as cyberattacking, hacking, and code-breaking (Rainie & Anderson, 2017).

To this end, big data-driven algorithms are constantly employed by organizations, corporations, institutions, and governments in different jurisdictions, or in different data jurisdictions (Ulbricht and Yeung, 2022), both in the Norths and in the Souths, to access citizens' datasets, either innocuously or nefariously, for various decision-making purposes such as the ones mentioned above. Inherently, algorithms operate on a discriminatory and differentiating computational logic. That is, they have to recognize and discriminate patterns on the basis of what they have by ignoring that which they do not have. Jago and Laurin (2022) point out that even though applying algorithms has led to a new hope in various human domains, nonetheless, algorithms are capable of both systematizing discrimination and obscuring its presence. This, the paper contends, is tantamount to normalizing discrimination, while simultaneously *invisibilizing* it. This is also equivalent to naturalizing exclusion, while pretending that it does not exist because algorithms are autonomous, neutral, rational, objective, fair, unbiased, and non-human. No, they are not necessarily so, especially social algorithms and algorithms meant to regulate and monitor human behavior! They are, as Mattiuzzo (2019) maintains, designed and created by humans. This is a point that she aptly frames as follows: "results provided by algorithms have a façade of objectivity, which runs from their use of mathematics … [c]urrent algorithmic systems are mostly concerned with finding correlation in data, not causation" (p. 3; also see Borgesius, 2018, pp. 7 and 9; Madden et al., 2017; cf. Cahan et al., 2019; Cofone, 2019, pp. 1409-1410). To add to this, they are concerned with identifying and recognizing familiar and relatable patterns from a sea of datasets.

The picture painted above underlines algorithmic exclusions, especially the exclusions of marginalized and underrepresented users such as BIPOC communities both in the Souths and in the Norths by algorithms. This is particularly the case when algorithms are programed in such a way as to replicate given historical human biases embodied in their input datasets. For instance, algorithms trained using unrepresentative, incomplete, insufficient, or biased datasets in which men are inferred and evaluated more positively than women in performance variables are likely to perpetuate negative evaluations of women as they employ gender as their predictive input variable (Jago & Laurin, 2022; also see Borgesius, 2018; Gilman & Green, 2018; Lee et al., 2019; Noble, 2018; Williams et al., 2018). In a different but related scenario, Noble (2018) points out that when querying the phrase, *black girls*, on a Google search, the information returned was *Big Booty* and other terms that sexually depicted black girls. Conversely, she contends that when the string, *white girls*, was queried, completely different results were returned. In the current paper, when a search string, *race and crime*, was queried by the author into the Google search engine (to 26 April 2022), the piece of information that was returned was: "According to the FBI, African-Americans accounted for 55.9% of all homicide offenders in 2019, with whites 41.1%, and "Other" 3.0% in cases where the race was known. Among homicide victims in 2019 where the race was known, 54.7% were black or African-American, 42.3% were white, and 3.1% were of other races". This was out of 3,710,000,000 returned results. The primary source of this returned information was Wikipedia (2022), which had last been updated on 18 March 2022 (also see Lee et al., 2019).

By contrast, when the same search string was queried into the Microsoft Bing search engine (to 26 April 2022), the first result was: "Race is one of the correlates of crime receiving attention in academic studies, government surveys, media coverage, and public concern". The primary source of the returned result, which was out of 134,000,000 results, was the same Wikipedia (2022) referenced by the Google search engine above. While the comparison of the two sets of results on the same search string from the two Internet search engines is not intended to imply that one search engine is more racist or discriminatory than the other, or the algorithms of one search engine are more racist or discriminatory than those of the other, the results emphasize what can happen when search engine algorithms are fed source datasets that potentially replicate human-induced biases. This significantly compromises the objectivity and fairness of such algorithms. This algorithmic replication of racist or discriminatory human biases have dire ramifications for BIPOC people as digital citizens in both the Souths and the Norths.

The point is, when algorithms have, as their predictive input, training data sources that exclude variables or that discriminate against variables related to underrepresented users and communities, they are likely to exclude those users and communities in their predictive pattern recognition and correlation. Terp (2020) contends that technologies, together with their associated AI and algorithms, can reinforce prevailing racist human biases by entrenching them in machine learning systems through biased input data encoded in algorithms. Or, by perpetuating racial and gender biases embedded in interactions that are mediated in the way designed technology interfaces and datasets are presented. Moreover, she talks about technology that is accidentally racist because it is designed to recognize only monoculture, and about technology and data science that reinforce racist biases, but which are regarded as neutral technology. To this, needs to be added technologies and datasets that are deliberately racist and exclusionary, but which are passed off as neutral, objective, and fair. Cave and Dihal (2020) berates the racial slant of AI as the racialization and the Whiteness of AI. Lee et al. (2019) aptly contextualize how bias can emerge from algorithms:

Bias in algorithms can emanate from unrepresentative or incomplete training data or the reliance on flawed information that reflects historical inequalities. If left unchecked, biased algorithms can lead to decisions which can have a collective, disparate impact on certain groups of people even without the programmer's intention to discriminate (n.p.).

They go on to assert that algorithmic bias finds its way into online recruitment tools, online advertisements, word associations, facial recognition technologies, and criminal justice algorithms. All of this relates mainly to BIPOC people as online algorithmic subjects.

In this regard, we have a situation in which data and algorithms are weaponized to discriminate against and to exclude BIPOC people. In this context, Raghuveera and Koch (2020) discuss how data has been weaponized against underrepresented communities, or what they call communities of color in South Africa and in the United

 

States. They argue that "[d]ata is weaponized whenever it is used to inflict harm, well-intentioned or not" (n.p.). Indeed, Treré and Milan (2021) point out that even in Latin America, there is a tendency to replicate social asymmetries, which are a colonial legacy, in automated, data-driven systems that are often amenable to data manipulation and corruption. For this paper, the point is that both data and algorithms can be weaponized to discriminate against and to exclude BIPOC people as digital citizens in both the Souths and the Norths. This is what Calzada (2022) refers to as algocracy, which for this paper, is the rule and government by algorithms. Its alter ego is *datacracy*: the rule and government by big data, whose other variants are data colonialism and data capitalism as discussed earlier. In fact, with respect to the algorithmic colonization of Africa as part of the Souths, Birhane (2020) argues that algorithmic domination and colonialism as driven by corporate monopolies has come to replace traditional or historical colonialism, and is passed off and marketed as "state-of-the-art algorithms", cutting-edge "AI solutions", and "technological innovation[s]" (p. 391).

## 5. Critical southern decolonial approach to datafication, algorithms, and digital citizenship

This section of the paper proposes a critical southern decolonial (CSD) approach to datafication, algorithms, and digital citizenship. This particular approach seeks to build on the work of researchers such as Adams (2021), Ali (2017), Couldry and Mejias (2021), Mohamed et al. (2020), Ricaurte (2019), and Zembylas (2021). However, these researchers' work focuses on the decolonization of or the decolonial approach to one of these three aspects. For example, Couldry and Mejias (2021) deal with a decolonial turn to data and technology; Ali (2017) advocates decolonizing information narratives in algorithmic racism; Adams (2021) argues for decolonizing AI; and Mohamed et al. (2020) and Zembylas (2021) propose a decolonial AI. For her part, Ricaurte (2019) explores data epistemologies, the coloniality of power, and resistance. This section of the paper, therefore, argues for a CSD approach to datafication, to algorithms, and to digital citizenship, simultaneously. CSD integrates critical scholarship and southern decoloniality. Briefly, the former entails a critical approach to existing forms of scholarship, while the latter is a decolonial approach as framed and theorized from the Souths (Chaka, 2022; Ndlangamandla & Chaka, 2022). In this context, CSD advocates, on the one hand, a critical view of: data, datafication, data literacy and infrastructure, algorithms, digital citizenship, digitality, and technology. On the other hand, it calls for challenging, interrogating, problematizing, critiquing, and decolonizing of all of these aspects and the Euro-American colonialist orientations on which they are founded.

The picture painted above, applies to both the Souths and the Norths, even though more so to the former than to the latter. In relation to data coloniality, BIPOC communities, as marginalized and underrepresented users both in the Souths and in the Norths, should cease serving as passive purveyors of data to organizations, to corporations, to institutions, and to governments that utilize and exploit their data for their own purposes. Also, the manner in which these users' datasets are harvested, extracted, appropriated, and represented by big and small tech companies and by governments must be challenged and criticized. Persistent calls need to be made to involve BIPOC users (see Karumbaiah and Brooks, 2021) in deciding the fate and the endgame of their extracted data. This fate should not be left to the whims and dictates of data privacy and security regimes or of data legal and regulatory frameworks. In fact, CSD questions and challenges the very existence of these regimes and frameworks as, in most cases, they are formulated without involving end users, from whom datasets are extracted. Needless to say that this tendency has to be flagged as a classic example of a data colonialist practice. Couldry and Mejias' (2019b) view that colonization by data at the point of getting connected to digital technologies is an entry point to a costly appropriation of human life, becomes more instructive in this case.

A CSD approach in this context, then, advocates an epistemic disobedience to data and datafication, and to data surveillance. It challenges the epistemic enterprise that underpins and informs data configuration and datafication, and argues that this epistemic enterprise is disproportionately ethnocentric as it is biased toward Euro-American, White, middle-class, racial demographics (see Arora, 2018; also see Ali, 2017; Mumford, 2021; Raghuveera and Koch, 2020; Ricaurte, 2019; Terp, 2020). To this effect, Ricaurte (2019) opines how dominant data epistemologies, based on Western rationality, tend to perpetuate serial marginalization of and to reproduce multiple injustices to underrepresented users in multicultural countries that have huge social inequality levels. In doing so, these data epistemologies promote a misrepresentation and a mischaracterization of underrepresented users in the Souths, while subjecting their beings, their languages, and their cultures to data violence, oppression, and alterity. It is this data-based epistemic violence and oppression, which relegates BIPOC users in the Souths to data subalterns and to data purveyors and, which is anchored on Euro-American data infrastructures that CSD rejects and challenges.

Moreover, CSD resists and questions algorithmic coloniality. Sustained by its symbiotic relationship with data coloniality, algorithmic coloniality adds another layer to the colonialty of modern-day, automated, data-driven, machine-learning decision-making process. As mentioned earlier, algorithmic coloniality comes into play when algorithms have as their source and utilize as their sole predictive basis, datasets and AI configurations modeled on heteronormative, White, modern subjects (see Mumford, 2021), and this to the exclusion of the human attributes of underrepresented, non-European subjects in the Souths. The endgame of all of this, is the algorithmic bias against and the algorithmic marginalization and exclusion of the underrepresented users and communities in the Souths. Additionally, this algorithmic coloniality leads to harm and violence being algorithmically inflicted on such users and communities. Gangadharan and Niklas' (2019) assertion that "systems powered by bad data, bad algorithmic models, or both lead to 'high-tech' discrimination – misclassifications, over target, disqualifications, and flawed predictions that affect some groups, such as historically marginalized ones, more than others" (p. 883) becomes relevant in this regard. A stand for CSD, in this case, is that the Euro-Americanism, the *Westernness*, and the Whiteness built into algorithms employed by organizations, corporations, institutions, and governments in different data jurisdictions in both the Souths and the Norths, must be dismantled and replaced by algorithms that are capable of recognizing the human attributes and the peculiarities of diverse BIPOC communities in the Souths and in the Norths. Importantly, CSD calls for the dismantling of algorithmic coloniality, or what Karumbaiah and Brooks (2021) refer to as the rootedness of algorithms in coloniality, and of what Birhane (2020) calls "the West's algorithmic invasion" (p. 389) of Africa, and by extension, of the Souths.

Furthermore, CSD impugns and rejects the coloniality embedded in digital citizenship (DC) or in data citizenship, and as embedded in its five emerging regimes (pandemic citizenship, algorithmic citizenship, liquid citizenship, metropolitan citizenship, and stateless citizenship) discussed earlier. It argues that in its current form, DC is conceptualized from a colonialist framework that uses as its prototypes, White, modern, middle-class, Euro-American subjects with the Western, digital, data, and algorithmic infrastructures to which they have access. This colonialist framing also applies to the five emerging regimes of DC. But, however, it excludes and is oblivious to BIPOC users in its canvas. If anything, such users feature as subalterns in it. What is missing and excluded from this framing in terms of the five emerging DC regimes are rural citizenship, Indigenous citizenship, nomadic citizenship, immigrant citizenship, refugee citizenship, diasporic citizenship, crossborder citizenship, and transnational citizenship that are characteristic of most BIPOC users in both the Souths and the Norths. These missing regimes of DC, are what CSD insists should be considered and included when DC is conceptualized and theorized across digital and data spheres both in the Souths and in the Norths.

## 6. Conclusion

This paper has focused on digital marginalization, data marginalization, and algorithmic exclusions in the Souths. To this end, it has explored how underrepresented users and communities, especially BIPOC communities, tend to be marginalized and excluded by digital technologies, by big data, and by algorithms employed by organizations, corporations, institutions, and governments in different data jurisdictions. Framing data colonialism within the Souths, the paper has pointed out that data ableism, data disablism, and data colonialism are at work when data collected, collated, captured, configured, and processed from these users and communities is utilized by these mega entities for their own multiple purposes. Some of these purposes are advertising, profit-making, monetization, tracking, surveillance, and decision-making. The paper has also highlighted how data coloniality is immune to legal and legislative interventions within data jurisdictions. In addition, it has discussed digital citizenship (DC), specifically foregrounding pandemic citizenship, algorithmic citizenship, liquid citizenship, metropolitan citizenship, and stateless citizenship as emerging regimes of DC.

Moreover, the paper has argued that even though there is a nexus between digital marginalization and exclusion and digital infrastructural underdevelopment, digital exclusion transcends the *haves* versus the *have nots* binarism as it manifests itself in multiple layers. Furthermore, it has characterized how algorithmic exclusions tend to replicate historical human biases despite the contention that algorithms are autonomous, neutral, rational, objective, fair, unbiased, and non-human. Finally, the paper has proposed a critical southern decolonial (CSD) approach to datafication, algorithms, and digital citizenship in terms of which data coloniality, algorithmic coloniality, and the coloniality embedded in DC have to be critiqued, challenged, and dismantled.

## References


Adams, R. (2021). Can artificial intelligence be decolonized? *Interdisciplinary Science Reviews*, 46(1–2), 176–197. https://doi.org/10.1080/03080188.2020.1840225

Ali, s. M. (2017). Decolonizing information narratives: Entangled apocalyptics, algorithmic racism and the myths of history. *Proceedings*, 1(50), 1–3. https://doi.org/10.3390/IS4SI-2017-03910

Andrejevic, M. (2014). Big data, big questions: The big data divide. *Journal of International Communication*, 8, 1673–89.

Armillas-Tiseyra, M., & Mahler, A. G. (2021). Introduction: New critical directions in Global South Studies. *Comparative Literature Studies*, 58(3), 465–484. doi: 10.5325/complitstudies.58.3.0465

Arora, P. (2018). Decolonizing privacy studies. *Television & New Media*, 1–13. https://doi.org/10.1177/1527476418806092

Benabdallah, L., Murillo-Zamora, C., & Adetula, V. (2017). Global south perspectives on international relations theory. https://www.e-ir.info/2017/11/19/global-south-perspectives-on-international-relations-theory/

Birhane, A. (2020). Algorithmic colonization of Africa. *SCRIPTed*, 17(2), 389–409.

Bloomberg, J. (2018). Digitization, digitalization, and digital transformation: Confuse them at your peril. https://www.forbes.com/sites/jasonbloomberg/2018/04/29/digitization-digitalization-and-digital-transformation-confuse-them-at-your-peril/?sh=53855bce2f2c

Borgesius, F. Z. (2018). Discrimination, artificial intelligence, and algorithmic. https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73

Cahan, E. M., Hernandez-Boussard, T., Thadaney-Isran, S., Rubin, D. L. (2019). Putting the data before the algorithm in big data addressing personalized healthcare. *NPJ Digital Medicine*, 2(78), 1–6. https://doi.org/10.1038/s41746-019-0157-2

Calzada, I. (2022). Emerging digital citizenship regimes: Pandemic, algorithmic, liquid, metropolitan, and stateless citizenships. *Citizenship Studies*. https://doi.org/10.1080/13621025.2021.2012312

Carmi, E., Yates, S. J., Lockley, E., & Pawluczuk, A. (2020). Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation. *Internet Policy Review*, 9(2). https://doi.org/10.14763/2020.2.1481

Cave, S., & Dihal, K. (20202). The Whiteness of AI. *Philosophy & Technology*, 33, 685–703. https://doi.org/10.1007/s13347-020-00415-6

Chaka, C. (2019). Re-imagining literacies and literacies pedagogy in the context of semio-technologies. *Nordic Journal of Digital Literacy*, 14(1-2), 53–68. https://doi.org/10.18261/issn.1891-943x-2019-0102-05

Chaka, C. (2020). Translanguaging, decoloniality, and the Global South: An integrative review study. *Scrutiny 2*, 25(1), 6–42. https://doi.10.1080/18125441.2020.1802617

Chaka, C. (2021). Educational data mining, student academic performance prediction, prediction methods, algorithms and tools: An overview of reviews. https://www.preprints.org/manuscript/202108.0345/v1/download

Chaka, C. (2022). The geopolitics of knowledge production in applied English language studies: Transknowledging and a two-eyed critical southern decoloniality (An inaugural lecture delivered on 22 August 2022 at the University of South Africa). https://youtu.be/xb9seQEUR3M OR https://www.youtube.com/watch?v=qu7zUH9lAbE&list=PL4cHAn5OGG5NXSagMU_wiCDczpQOUaCym&index=1

Charitsis, V., & Lehtiniemi, T. (2022). Data ableism: Ability expectations and marginalization in automated societies. *Television & New Media*, 1–16. https://doi.org/10.1177/15274764221077660

Clarke, M. (2018). Global South: What does it mean and why use the term? https://onlineacademiccommunity.uvic.ca/globalsouthpolitics/2018/08/08/global-south-what-does-it-mean-and-why-use-the-term/

Cofone, I. N. (2019). Algorithmic discrimination is an information problem. *Hastings Law Journal*, 70(6), 1389–1444. https://repository.uchastings.edu/hastings_law_journal/vol70/iss6/1

Couldry, N., & Mejias, U. A. (2019a). Data colonialism: Rethinking big data’s relation to the contemporary subject. *Television & New Media*, 20(4), 336–349. https://doi.org/10.1177/1527476418796632

Couldry, N., & Mejias, U. A. (2019b). *The costs of connection: How data is colonizing human life and appropriating it for capitalism.* Stanford: Stanford University Press.

Couldry, N., & Mejias, U. A. (2021). The decolonial turn in data and technology research: What is at stake and where is it heading? *Information, Communication & Society*, 1–18. https://doi.org/10.1080/1369118X.2021.1986102

Dados, N., & Connell, R. (2012). The Global South. *Contexts*, 11(1), 12–13. https://doi.org/10.1177/1536504212436479

Escobar, A. (2007). Worlds and knowledges otherwise: The Latin American modernity/coloniality research program. *Cultural Studies*, 21(2-3), 179–210.

Floridi, L. (2020). The fight for digital sovereignty: What it is, and why it matters, especially for the EU. *Philosophy & Technology*, 33(3), 369–378. https://doi.org/10.1007/s13347-020-00423-6

Francis, A. R. (2021). Global Southerners in the North. *Temple Law Review*, 93, 689–711.

Fund, S. (2022). Digitality vs digitisation: New thinking, new opportunities. https://www.canon-europe.com/business/insights/articles/digitality-digitisation-mind-set/

Gangadharan, S. (2021). Digital exclusion: A politics of refusal. In L. Bernholz, H. Landemore & R. Reich. (Ed.), *Digital technology and democratic* (pp. 113–140). Chicago: University of Chicago Press. https://doi.org/10.7208/9780226748603-005

Gangadharan, S. P., & and Niklas, J. (2019). Decentering technology in discourse on discrimination. *Information, Communication & Society*, 22(7), 882–899. https://doi.org/10.1080/1369118X.2019.1593484

Geeker, A., & Hind, S. (2019). Infrastructural surveillance. *New Media & Society*, 2(8), 1414–1436. https://doi.org/10.1177/1461444819879426

Gilman, M., & Green, R. (2018). The surveillance gap: The harms of extreme privacy and data marginalization. *New York University Review of Law & Social Change*, 42(253), 253–307.

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, 5(2). https://doi.org/10.1177/2053951718786316

Greenwood, F. (2020). Data colonialism, surveillance capitalism and drones. In D. Specht (Ed.), *Mapping crisis: Participation, datafication and humanitarianism in the age of digital mapping* (pp. 89–117). London: University of London Press.

Grosfoguel, R. (2007). The epistemic decolonial turn. *Cultural Studies*. 21(2-3), 211–223.

Heeks, R., & Shekhar, S. (2019). Datafication, development and marginalised urban communities: An applied data justice framework. *Information, Communication & Society*, 22(7), 992–1011. https://doi.org/10.1080/1369118X.2019.1599039

Hintz, A., Dencik, L., & Wahl-Jorgensen, K. (2019). *Digital citizenship in a datafied society*. Cambridge, UK.: Polity Press.

Hsu, F. (2017). Resisting the coloniality of English: A research review of strategies. *The CATESOL Journal*, 29(1), 111–132.

Jago, A. S., & Laurin, K. (2022). Assumptions about algorithms' capacity for discrimination. *Personality and Social Psychology Bulletin*, 48(4), 582–595. https://doi.org/10.1177/01461672211016187

Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*, 3, 371–377. http://dx.doi.org/10.1016/j.giq.2016.08.011

Karumbaiah, S., & Brooks, J. (2021). How colonial continuities underlie algorithmic injustices in education. IEEE Research in Equity and Sustained Participation in Engineering, Computing, and Technology (IEEE RESPECT). https://doi.org/10.1109/RESPECT51740.2021.9620605

Kloß, S. T. (2017). The Global South as subversive practice: Challenges and potentials of a heuristic concept. The Global South, 11(2), 1–17. https://doi.org/10.2979/globalsouth.11.2.01

Kowalski, A. M. (2020). Global South-Global North differences. In F. W. Leal, A. M. Azul, L. Brandli L., P. Özuyar & T. Wall (Eds.), *No poverty, encyclopedia of the UN sustainable development goals* (pp. 1–13). https://doi.org/10.1007/978-3-319-69625-6_68-1

Lazar, M. M. (2020). "Politics of the South": Discourses and praxis. *Discourse & Society*, 31(1), 5–18. https://doi.org/10.1177/0957926519886126

Lee, N. T., Resnick, P., & Barton, G. (2019). Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

Lerman, J. (2013). Big data and its exclusions. *Stanford Law Review Online*, 66(55), 55–63.

Madden, M., Gilman, M., Levy, K., & Marwick, a. (2017). Privacy, poverty, and big data: A matrix of vulnerabilities for poor Americans. *Washington University Law Review*, 95(053), 53–125.

Mahler, A. G. (2017). Global South. In E. O'Brien (Ed.), *Oxford bibliographies in literary and critical theory*. New York: Oxford University Press. doi: https://doi.org/1093/OBO/9780190221911-0055

Maldonado-Torres, N. (2007). On the coloniality of being. *Cultural Studies*, 21(2), 240–270. http://dx.doi.org/10.1080/09502380601162548

Maldonado-Torres, N. (2018). Outline of ten theses on coloniality and decoloniality. http://fondation-frantzfanon.com/wp-content/uploads/2018/10/maldonado-torres_outline_of_ten_theses-10.23.16.pdf

Martin, C., Hope, S., Zubairi, S. (2016). The role of digital exclusion in social exclusion. https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2016/09/LOW-2697-CUKT-Digital-Participation-Report-REVISE.pdf

Mattiuzzo, M. (2019). Algorithms and big data: Considerations on algorithmic governance and its consequences for antitrust analysis. *Revista de Economia Contemporânea*, 23(2), 1–19. http://dx.doi.org/10.1590/198055272328

Mignolo, W. (2007). Delinking: The rhetoric of modernity, the logic of coloniality and the grammar of de-coloniality. *Cultural Studies*, 21(2), 449–514. http://dx.doi.org/10.1080/09502380601162647

Milan, S., & Treré, E. (2019). Big data from the South(s): Beyond data universalism. *Television & New Media*, 20(4), 319–335. https://doi.org/10.1177/1527476419837739

 

Milan, S., & Treré, E. (2021). Big data from the South(s): An analytical matrix to investigate data at the margins. In D. Rohlinger & S. Sobieraj (Eds.), *The Oxford handbook of sociology and digital media* (pp. 1–21). Oxford University Press. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780197510636.001.0001/oxfordhb-9780197510636-e-7

Mohamed, S., Png, M-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33, 659–684. https://doi.org/10.1007/s13347-020-00405-8

Mukhametzyanov, I. (2022). Digital citizenship and the student's digital footprint: Questions of application, promotion and data protection. In Proceedings of the International Scientific and Practical Conference on Computer and Information Security (INFSEC 2021) (pp. 12–16). https://doi.org/10.5220/0010616600003170

Mumford, D. (2021). Data colonialism: compelling and useful, but whither epistemes? *Information, Communication & Society*. https://doi.org/10.1080/1369118X.2021.1986103

Ndlangamandla, S., & Chaka, C. (2022). Relocating English studies and SoTL in the Global South: Towards decolonizing English and dismantling the coloniality of language. *Journal of Contemporary Issues in Education*, 17(2), 39-56. https://doi.org/10.20355/jcie29495

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York Press.

Núñez-Pardo, A. (2020). Inquiring into the coloniality of knowledge, power, and being in EFL textbooks (Indagar la colonialidad del saber, del ser y del poder en libros de texto de inglés). *HOW Journal*, 27(2), 113–133. https://doi.org/10.19183/how.27.2.566

Orwat, C. (n.d.). Risks of discrimination through the use of algorithms. https://www.antidiskriminierungsstelle.de/EN/homepage/_documents/download_diskr_risiken_verwendung_von_algorithmen.pdf?__blob=publicationFile&v=1

Pawluczuk, A., Yates, S. J., Carmi, E., Lockley, E., & Wessels, B. (2020). Data citizenship framework: exploring citizens' data literacy through data thinking, data doing and data participation. https://drive.google.com/file/d/1tDq670KtfYqnxkBovcaWP2GpUBcx0ASh/view

Pötzsch, H. (2019). Critical digital literacy: Technology in education beyond issues of user competence and labour-market qualifications. *tripleC*, 17(2), 221–240.

Quijano, A., & Ennis, M. (2000). Coloniality of power, Eurocentrism, and Latin America. *Nepantla: Views from South*, 1(3), 533–580.

Raghuveera, N. & Koch, T. (2020). #GoodTechChoices: Addressing unjust uses of data against marginalized communities. https://www.atlanticcouncil.org/blogs/geotech-cues/goodtechchoices-address-unjust-uses-of-data/

Rainie, L., & Anderson, J. (2017). Code-dependent: Pros and cons of the algorithm age. http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age

Ricaurte, P. (2019). Data epistemologies, the coloniality of power, and resistance. *Television & New Media*, 20(4), 350–365. https://doi.org/10.1177/1527476419831640

Richardson, J., & Milovidov, E. (2019). *Digital citizenship education handbook*. Strasbourg: Council of Europe.

Segura, M. S., & Waisbord, S. (2019). Between data capitalism and data citizenship. *Television & New Media*, 20(4) 412–419. https://doi.org/10.1177/1527476419834519

Sin, J., Franz, R. M., Munteanu, C., & Neves, B. B. (2021). Digital design marginalization: New perspectives on designing inclusive interfaces. . In Proceedings of the CHI '21. https://doi.org/10.1145/3411764.3445180

Tenney, M., & Sieber, R. (2016). Data-driven participation: Algorithms, cities, citizens, and corporate control. *Urban Planning*, 1(2), 101–113. https://doi.org/10.17645/up.v1i2.645

Terp, S-J. (2020). #GoodTechChoices: Addressing past and current racism in tech and data. https://www.atlanticcouncil.org/blogs/geotech-cues/goodtechchoices-address-racism-tech-and-data/

Tomczyńska, W. (2017). Digital exclusion: Definitions, causes, countermeasures. *Adeptus*, 10, 1–26. https://doi.org/10.11649/a.1503

Treré, E. & Milan, S. (2021). Latin American perspectives on datafication and artificial intelligence: Traditions, interventions, and possibilities. *Palabra Clave*, 24(3), e2431. https://doi.org/10.5294/pacla.2021.24.3.1

Ulbricht, L., & Yeung, K. (2022). Algorithmic regulation: A maturing concept for investigating regulation of and through algorithms. *Regulation & Governance*, 16, 3–22. https://doi.org/10.1111/rego.12437

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & society*, 12(2), 197–208.

Wikipedia. (2022). Race and crime in the United States. https://en.wikipedia.org/wiki/Race_and_crime_in_the_United_States#:~:text=According%20to%20the%20FBI%2C%20African,3.1%25%20were%20of%20other%20races

Williams, B. A., Brooks, C. F., & Shmargad, y. (2018). *Journal of Information Policy*, 8, 78–115.

Wolvers, A., Tappe, O., Salverda, T., & Schwarz, T. (2015). Introduction*. Concepts of the global south: voices from around the world*. https://kups.ub.uni-koeln.de/6399/1/voices012015_concepts_of_the_global_south.pdf

Zembylas, M. (2021). A decolonial approach to AI in higher education teaching and learning: Strategies for undoing the ethics of digital neocolonialism. *Learning, Media and Technology*. https://doi.org/10.1080/17439884.2021.2010094

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