| SEEING | STONES | AND |
|--------|--------|-----|
| SPACES | BEYOND | THE |
| VALLEY |        |     |
|        |        |     |
| 03.06. |        | BNN |
| 17.07. |        | WWA |

### The decolonial turn in data and technology research: what is at stake and where is it heading?

#### NICK COULDRY, ULISES ALI MEJIAS

### Introduction

While the business and government celebrations of Big Data continue everywhere (Manyika et al., 2011; World Economic Forum, 2013; Siebel, 2019; Siggelkow & Terwiesch, 2019), the era when interventions in academic fields largely echoed that celebration is thankfully over. As evidence of this, there has emerged in the last decade or so something that could be termed a 'decolonial turn' within the diverse field of critical studies of data. The voices contributing to this movement hail from disparate camps, and interdisciplinary dialogue has been somewhat limited, since each intervention uses different terminology aimed at overlapping audiences. Nevertheless, cross-disciplinary conversations are starting to converge. Our own work is part of this. While writing our book, The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism (2019),<sup>1</sup> we were aware of many of these voices, and we used their arguments to expand our own. But since the book's publication, we have found further authors that had applied - or are currently applying - a decolonial framework to the study of data. Hence, part of the motivation for writing this article is to situate our argument in the context of this wave of converging ideas. This survey will inevitably be shaped by the perspective of our own work, but we hope to convey both the wider debate and those features of our model that can help in expanding the conversation further. This is an opportune moment to take stock of this critical mass of research whose relevance has become all the more apparent during a global pandemic when ever more aspects of daily life have been conducted on data-extracting digital platforms.

This article is in four parts. First, we present a genealogy (by necessity, short) of research that predates or parallels the concept of data colonialism, or to varying degrees converges with it. Second, we identify the differences and common ground that underlie all these approaches, while also responding briefly to some critics

<sup>1</sup> Our own work in the area began in January 2016, and was first presented at IAMCR, Colombia in July 2017.

of our own particular position. Third, we summarize the most important advantages of a decolonial approach to data. In conclusion, we unpack some theoretical and practical implications that follow if the growing force of data colonialism (or some cognate concept) is acknowledged.

Before that, for context, we will briefly summarize the main features of our own approach to data colonialism (Couldry & Mejias, 2019a, 2019b). In common with other critical approaches to data - in particular platform capitalism (Srnicek, 2017), surveillance capitalism (Zuboff, 2019), and work in critical data studies (Benjamin, 2019; Broussard, 2019; Eubanks, 2018; Noble, 2018) - the data colonialism thesis foregrounds the continuous extraction of economic value from human life through data. But it is distinctive in repositioning those critiques explicitly within colonialism's centuries-old relations to capitalism. It is beyond question that contemporary data practices are important for the late modern capitalism that some call 'digital' or 'informational' (Schiller, 2000; Castells, 1996; Cohen, 2019). But to frame contemporary data practices only in terms of capitalism's dynamics ignores how capitalism itself emerged on the basis of colonialism's detailed histories, in particular the European colonial powers' vast aggregation of global resources that fueled industrial capitalism (Beckert, 2014; Beckert and Rockman, 2016; Pomeranz, 2001). The concept of data colonialism therefore expands, not narrows, our understanding of data practices within an arc of historical comparison that includes both China and North America/Europe (Couldry & Mejias, 2019b, chapters 2 and 3).

More specifically, the data colonialism thesis echoes critiques of traditional Marxism (Robinson, 2000; Williams, 1994) for assuming that core features of 18th and nineteenth century colonialism (slavery, the plantation) were irrelevant to industrial capitalism, being at most a form of 'primitive accumulation' (Marx, 1976, p. 714) transcended by the 'higher' historical stage of capitalism. On the contrary, as the historian of cotton Sven Beckert (2014, p. 44) puts it, 'slavery, colonialism and forced labor, among other forms of violence, were not aberrations in the history of capitalism, but were at its very core'. Far from being 'primitive', innovations in colonial plantation economies were core to the explosive growth of industrial capitalism in late nineteenth century North America and Europe (Beckert, 2014, chapters 4 and 5; Cooke, 2003; Baptist, 2016). It is therefore the integrated history of colonialism and capitalism, not capitalism alone, that helps us understand the intensified economic and social management underlying capitalism's huge nineteenth century growth.<sup>2</sup> Similarly, modern science and technology emerged within a colonial and imperial context, as techniques of governmentality on a global scale (Jasanoff, 2006). Expanding such arguments, we ask: what if we interpret data and technology today in terms of not just historic, but contemporary and evolving, relations between colonialism and capitalism?

Admittedly, colonial and post-colonial studies have long emphasized the importance of interpreting the present in light of the past, influencing many disciplines. For example, David Harvey (2004) repurposed 'primitive accumulation' into the more useful notion of 'accumulation by dispossession', translating colonialism's core act (dispossession) into a modality of capitalist expansion. This inspired Thatcher

2 The all-encompassing control of workers – a core characteristic of capitalism – experienced its first great success on the cotton plantations of the American South' (Beckert, 2014, p. 115; compare Baptist, 2016).

et al.'s (2016) initial proposal of 'data colonialism': while they foreground data extractivism (cf. Mezzadra & Neilson, 2018), they propose the term 'data colonialism' as just a metaphor (Thatcher et al., 2016, p. 992), confirming Harvey's implication (2004) that the main explanatory model for understanding 'accumulation by dispossession' remains capitalism, not colonialism.

But what if, consistent with the historical arguments just reviewed, we insist on an explanatory model for Big Data practices in which colonial extractivism remains a real, not metaphorical, feature of capitalist accumulation? There are two ways of formulating this. First, that a neo-colonial legacy still shapes capitalism's development: consider Facebook's Free Basics programme in Africa and elsewhere (Nothias, 2020). Second, more directly, that the extraction of value through data represents a new form of resource appropriation on a par with the landgrab (the seizure of land, resources and labor) that kicked off historical colonialism. If so, data practices today would represent not just a continuation of colonialism/ capitalism, but a distinctive new stage of colonialism that lays the foundations for new developments in capitalism, just as colonialism's original landgrab enabled capitalism's emergence and subsequent centuries of colonial oppression. This is not to deny that 'landgrabs' are important also to capitalism (Dörre et al., 2015, chapter 1), but to insist their most obvious historical reference-point lies in colonialism, not capitalism. Either way, understanding the contemporary data landgrab requires us to understand how colonialism and capitalism intersect in new ways. Explain- ing how has been the goal of decolonial data studies for at least a decade, under differently named proposals.

## 1. Converging decolonial perspectives on data and technology

We first find a linkage between computing and colonialism in the 2010 proceedings of ACM's (Association for Computing Machinery) Conference on Human Factors in Computing Systems (CHI). There, Irani, Dourish et al. proposed the idea of 'postcolonial computing' as a way to think of 'the many ways histories, power relations, and epistemology tacitly underpin engagements in [computer system] design' (Irani et al. 2010).<sup>3</sup> Two years later, the proceedings of ACM's Conference on Ubiquitous Computing contain a visionary paper by Dourish and Mainwaring (2012) that critiqued the 'colonial impulse' of ubicomp, drawing on perspectives from history, philosophy of knowledge, development theory, and the quantification of the social. It is impossible to think even of these early interventions outside the long history of related critiques of global power from within geography, science and technology studies, and critical political economy. Even before this decolonial turn, 'critical data studies' was applying lessons from these fields to examine the uneven development of data production (Dalton et al., 2016). Another early application of the framework of colonialism to digital technologies was Syed Mustafa Ali's (2016) work on 'decolonial computing', based on a conference presentation from 2014. This concept frames computing itself

3 This important paper has stimulated subsequent work on the role that race plays in shaping the conditions of computing and IT work (for example Amrute, 2016). We return to the issue of race in the next section.

as a modern – and therefore colonial – phenomenon, and (as had feminists of technological practices generally) asks 'who is doing computing, where they are doing it, and, thereby, what computing means...' (2016, para. 17).

In the same year as Thatcher et al. introduced the term 'data colonialism', Aouragh and Chakravartty (2016) published an important piece on the geopolitics of media and information studies, using a colonial lens to look at infrastructures in the Global South (cf. Alhassan & Chakravartty, 2011). Antonio Casilli's, 2017 article on digital labor studies advocated 'a digital decolonial turn' while rejecting other less coherent interpretations of contemporary data practices as 'colonial'. Casilli's focus was work on digital platforms (from microwork platforms to smart platforms, and social media generally). Noting Casati's (2013) early use of the term 'digital colonialism' to capture platform language about technology, Casilli argued this general usage of the term 'colonialism' falls prey to 'the neocolonialism pitfall', which assumes 'that any form of inter- national power relation can be conflated with neocolonial dynamics' (Casilli, 2017, p. 3945). Crude uses of the term 'colonialism' in the realm of data, Casilli argued, exhibit an 'orientalism' that 'situat[es] countries with a history of colonization outside change and agency' (Casilli, 2017, p. 3946), that is, as mere victims of external power. Casilli proposed his own decolonial perspective on digital labor that emphasized the coloniality of knowledge, voice and power (2017, pp. 3946-3048).

Meanwhile, Renata Avila Pinto (2018) described Big Tech's relations to countries in the Global South in terms of the dependency their tools create, the platforms' influence on social and political processes, the population surveillance they enable, and the extraction of value from social data, using a broad language of feudalism and colonialism.

In 2019, Stefania Milan and Emiliano Treré edited a special issue on Big Data from the South in the journal *Television & New Media*. While not all chapters employed a colonial thesis, some did, including ours (Couldry & Mejias, 2019a, published first online in 2018), and Paula Ricaurte's (2019). We have described our core argument already. Ricaurte engages in analysis of the continued epistemic violence against women, indigenous peoples, and ethnic groups from the start of historic colonialism to today's data practices within an explicitly 'decolonial, intersectional and feminist analysis of data colonialism' (especially p. 353). Common to our approaches is an emphasis on the long-term global asymmetries in economic, cultural and knowledge production conceptualized by Peruvian sociologist Aníbal Quijano as 'coloniality' (Quijano, 2007 [orig in Spanish 1989]; Maldonado-Torres, 2007; Wynter, 2003).

Mirca Madianou's article on 'technocolonialism' advanced the debate (Madianou, 2019), as with Casilli's 'decolonial turn', through a specific focus: humanitarian practice. Here, Madianou argues, Global North power meets the economic and social conditions of the Global South, shaped by centuries of colonial government and neocolonial policies, and triggered by the world's ongoing refugee and migration crisis and the management technologies that emerged under neoliberalism. She shows how data practices play a considerable role in 'humanitarian' settings by transforming 'development' relations, extracting value and discriminating between populations within a wider 'coloniality'. Madianou affirms the continued agency of peoples of the Global South in resisting technocolonialism, thus avoiding Casilli's

'orientalism' trap. However, a potential worry with the technocolonialism thesis, as we see it, is to leave unclear how it applies to parallel processes of data extraction aimed at disadvantaged populations in the Global North (Eubanks, 2018). Do these fall outside technocolonialism? If they are part of it, does this call into question technocolonialism's special ties to the Global South?

A further decolonial approach to data and technology is Michael Kwet's work on 'digital colonialism' (Kwet, 2019). This term, as already noted, had been used by Casati in 2013 and also proposed by the Global Voices project (Can Facebook Connect the Next Billion?, 2017; Solon, 2017), but Kwet offers a more developed version, proposing that data and technology practices in the Global South, particularly Southern Africa, are directly continuous with earlier economic imperialism. Kwet's focus is exclusively on US Big Tech, saying little about China's Big Tech in Africa and offering a model of 'imperial control' exercised through the influence of the digital ecosystem over 'political, economic and cultural domains of life' (2019, p. 3, added emphasis). Kwet (2021) argues eloquently for the greater global force of US Big Tech compared with China's, but an overall worry with this account is that it can be redescribed as just 'global surveillance capitalism' (Kwet, 2019, p. 3) in a neo-colonial setting. It therefore says little about the role of coloniality in knowledge, and specifically data, production. That said, Kwet offers important examples of how datafication operates in Global South settings (the practices of Uber, Google and Facebook and educational technology platforms).

Relevant here too is recent work that affirms a decolonial approach to artificial intelligence. Google Deep Mind scientist Shakir Mohamed and colleagues have argued for the role of decolonial theory and decolonial science in generating a more ethical artificial intelligence (2020). Mohamed, Png & Isaac's target is not the wider transformation of data colonialism or indeed surveillance capitalism, but rather 'algorithmic coloniality', the reproduction of colonial power structures in algorithmic practice. They operationalize the idea of decoloniality by reforming Al's fairness, accountability and transparency. Meanwhile Sabelo Mhlambi (2020) has argued that artificial intelligence governance needs to be enriched by 'Ubuntu', the African conception of relational community, which not only challenges the specific market-driven priorities of colonial data practices and the narrow philosophical sources recognized in computer science. Nhemachena et al. (2020), in turn, offer a wide-ranging critique of the Internet of Things in African society, drawing on data colonialism, digital colonialism and coloniality, and calling for African data sovereignty in response to a threatened colonization of being (Bulhan, 2015) driven by data power.

This summary necessarily leaves out still more work that, while not primarily concerned with decoloniality and data, has made important contributions to the debate. Consider Julie E. Cohen's argument (based on the colonial concept of *terra nullius*) that data is 'just there' for the taking by corporations and governments, as colonies' land once was (Cohen, 2019, p. 59). We must also acknowledge Zuboff's (2019) surveillance capitalism thesis, debated at length elsewhere. Although it acknowledges colonial parallels to the actions of Big Tech corporations like Google (Zuboff, 2019, pp. 12, 178–179), it never offers a developed notion of what might be colonial about contemporary data practices. At most, it uses colonial history to suggest a striking echo. Moreover, the global and historical breadth of colonialism does not fit well with Zuboff's core argument that surveillance capitalism is a recent and deviant version of contemporary US capitalism that, while alarming, can be reined in to restore capitalism to an unproblematic path (Zuboff,

2019, p. 22, 194). By contrast, interpreting data and technology through the combined lens of colonialism and capitalism's interrelations over five centuries affords no such easy escape route from our current problems, including those that point to the racism embedded in our technologies. The latter requires a separate discussion.

#### 2. Connecting data colonialism to racism and discrimination

Recent critical studies of technology, whether or not they discuss colonialism, have productively focused on racism and discrimination as it relates to artificial intelligence and datafication, and with special intensity in the USA. Here Yarden Katz's book *Artificial Whiteness* (2020) acts as a bridge from the critical debates on AI just discussed. Katz's argues that AI serves imperial and capitalist projects by privileging the ideology of whiteness in military-industrial-academic projects that seek to dispossess people of their resources, normalize the mass incarceration and surveillance of black and brown bodies, and universalize the shape of the (white) self by defining how computer systems should interact with it. What relation to such debates does data colonialism have, now that we have broadened our understanding of it?

Two writers, Safiya Noble (2018) and Ruha Benjamin (2019), have had enormous impact by focusing on the workings of algorithmic power and code. Whereas race occurs only incidentally in accounts of data power such as Zuboff's,Benjamin and Noble argue that, because algorithms and code are produced in the USA under conditions of deep and continuing racism, they reproduce and amplify racial differentiation under a dangerous veneer of objectivity and scientific truth (Noble, 2018, p. 24; Benjamin, 2019, pp. 5–6). The result poorly serves 'multiracial democracy' (Noble, 2018, p. 186) and facilitates racial modes of social control (Benjamin, 2019, p. 6). The *implicit* links here to the colonial history which produced the USA's deeply racialized society and economy are clear, as are those to how knowledge, power and race have been intertwined throughout colonial history (what Quijano calls 'coloniality'). Strikingly, however, in both books explicit discussion of colonialism is relatively scarce.<sup>4</sup>

A more direct connection between contemporary technology (specifically for surveillance) and the colonial history of surveilling black bodies, with or without technology, is found in Simone Browne's important book Dark Matters<sup>5</sup> (2015). Browne explains how surveillance in history has worked to sustain racism and produce raced subjects under colonialism (colonial New York City) and capitalism (contemporary biometric imaging). Drawing on Fanon, she argues that raced surveillance enables a 'sociogeny', 'the organizational framework... that names what is, and is not, bounded within the category of the human' (Browne, 2015, p. 7). Earlier we noted the importance of the slave plantation to the history of capitalist management. There is a strong convergence between Browne's work on the long-term surveillance of black bodies and the insistence of decolonial approaches on placing capitalism within the longer history

4 But see Noble (2018, p. 86); Benjamin (2019, p. 44, 135-126, 176).

5 Cited Benjamin (2019, p. 44).

of colonialism. Seeta Peña Gangadharan's work on race and data discrimination within everyday life in the contemporary USA sees fundamental continuities between contemporary data practices and long histories of race-based injustice and marginalization (Gangadharan & Niklas, 2019; cf. Our Data Bodies project, https://www.odbproject.org/). Important also is Charlton McIlwain's (2020) history of the intersection of computing's and datification's development in the USA and widespread racism and white supremacism.

Does reference to capitalism's raced nature make discussion of colonialism's relation to data and technology redundant? That seems to be the position taken in a recent article by Tressie Cottom. Criticizing the sociology of race for failing to deal with platform capitalism and accounts of capitalism for failing to deal adequately with race, Cottom usefully identifies processes of 'predatory inclusion' (2020, p. 443) at work within platform capitalism. These processes bind racialized populations into a social world that continues to discriminate against them. Cottom theorizes these developments exclusively in terms of Gargi Bhattacharyya's concept of 'racial capitalism' (2018), not colonialism. But Bhattacharyya had offered an unusually broad account of how capitalism was organized through race via the global division of labor (2018, chapter 1), locating the origins of 'racial capitalism' in the colonial 'pursuit of economic interests' (what earlier we called the 'appropriation of resources') that developed existing forms of racism to achieve its goals (2018, pp. 77-78). Bhattacharyya quotes Patrick Wolfe: 'colonisers did not set out to create a racial doctrine. They set out to create wealth' (Wolfe, 2016, p. 52; quoted Bhattacharyya, 2018, p. 77). Racism, Bhattacharyya argues, was used to rationalize the dispossession of territory, a move which she links to the wider 'logic of coloniality', drawing on Walter Mignolo (Bhattacharyya, 2018, p. 74, 99; citing Mignolo, 2011) and the historical framework that is core to data colonialism.

The concept of racial capitalism itself therefore makes most sense if explicitly reconnected to the history of colonialism and coloniality (cf. Melamed, 2015). Meanwhile, there are other fundamental dimensions of social discrimination which data and technology currently reproduce and intensify – class (Eubanks, 2018), gender (D'Ignazio & Klein, 2020), and the intersectionality of class, gender and race (Madden et al., 2017). A more comprehensive formulation, we suggest, would contextualize the unquestionable racism of contemporary data relations within colonialism's longer-term relations to what subsequently became capitalism, in all the latter's dimensions of social structuring. This both grounds the raced dimensions of today's capitalist economy and foregrounds the deep historical, indeed colonial, connections between the production of data and the broader stratification of knowledge production that Benjamin (2019) and Noble (2018) so vividly describe.

All this suggests an exciting new context for decolonial research about data: the beginnings, at least, of a movement – or a network of movements – which challenges the rationalities of commercial- and state-driven data extraction by reframing them in terms of colonialism's long entanglement with both capitalism and knowledge production throughout modernity. It is vital that this movement is not conceived as a 'discovery' of the Global North, or as unfolding exclusively in the Global South. For that would reproduce precisely the colonial knowledge hierarchies that needs to be challenged (Todd, 2016). On this, the theses of data colonialism and racial capitalism converge. Before we synthesize the perspectives emerging through the decolonial turn, it is worth noting some specific critiques of our own data colonialism thesis.

#### 3. Critiques of the data colonialism thesis

In the introduction to their special issue, Milan and Treré (2019) argue that critical approaches to datafication must reflect the diversity of the geographical, political and social contexts captured by the umbrella term 'Global South'. While noting the dangers to knowledge production from 'the relentless push towards datafication' across borders (2019, p. 321), they ask: 'does *the fabric* of datafication shift alongside with inequality, geography, class, race, and culture...?' Ignoring this possibility would be to commit 'the original sin of Western interpretations' of knowledge, 'data universalism' (2019, p. 325).<sup>6</sup>

Milan & Treré introduce 'data universalism' to capture the epistemic hubris of those who see Big Data 'as something operating outside of history and of specific socio-political, cultural, and economic contexts' (p. 325). While they endorse the notion of data colonialism (p. 326), others have interpreted Milan & Treré as implicitly critiquing the data colonialism thesis for not taking sufficient account of the actual contexts of datafication. Halkort (2019) insists that the analysis of 'data' must be 'better situate[d] within ontologies of the social... in historically and geographically specific contexts' (cf. Soledad & Waisbord, 2019, p. 417). We agree, and argue in *The Costs of Connection* that any social order wrought through data colonialism must take into account the diversity of subjects and histories, which shape specific possibilities of resistance (Couldry & Mejias, 2019a, chapter 6; Couldry & Mejias, 2019b, pp. 345–346).

But the trap of data universalism is subtle, especially when it extends, as Milan & Treré propose, to interpretations of 'datafication-related dynamics'. Consider two different uses of theory, one necessary, the other problematic. The first would recognize - as both we and Milan & Treré do - that Big Data discourse really does seek to impose its validity on countless local contexts across the world. Given this, the data colonialism thesis seeks to identify the force of this universal discourse and oppose it wherever it seeks influence. Ignoring that force for fear of contributing to a 'general discourse' risks missing the forest for the trees: business, governmental and social discourses and practices today do support datafication almost everywhere. A second, more problematic use of theory would assume that datafication plays out exactly the same way everywhere, ignoring detailed struggles over data extractivism in countless contexts across the world. But to confuse the second theoretical move with the first unwittingly blocks off the role which general conceptual frameworks can play as theoretical horizons for specific resistance against datafication (Connell, 2007, p. 207).<sup>7</sup> It is curious that this move to make general conceptual observations has been interpreted (Calzati, 2020) as itself a colonizing move, somehow 'reducing' the complexity and diversity of contemporary struggles over data.

6 Milan and Treré build on Chan's (2014) concept of 'digital universalism'.

 As Raewyn Connell writes, 'only knowledge produced on a planetary scale is adequate to support the self-understanding of societies now being forcibly reshaped on a planetary scale' (2007, p. vii).
Of course such 'knowledge' must also make itself accountable on a local scale. For a parallel defence of theory's role in critiquing the discourse of data 'optimization', see Powell (2021). Segura & Waisbord's (2019) critiques of the data colonialism thesis do not depend on criticisms of universal discourse as an automatic erasure of the specific. They argue, first, that Latin America's less developed informational infrastructure 'during the past half century' (2019, p. 417) prevents datafication rolling out as in Europe and North America. Leaving aside the fact that China achieved intense datafication in less than half a century, this is an important point about the unevenness of how data colonialism may spread globally. But there are countervailing arguments, too. Consider the efforts of US Big Tech in many countries to pave the way, through enhanced digital connectivity and proprietorial internet access portals, for precisely the infrastructures of connection which will make wide-ranging datafication possible (Arora, 2019; Nothias, 2020). Consider also the efforts of US Big Tech to offer its machine learning expertise to help retool social welfare systems in Latin America through, for example, the Horus project (Magalhães & Couldry, 2021). Research is yet to be done on parallel efforts by China in Africa and other parts of the Global South.

Segura and Waisbord's second critique is that the 'methods' of datafication (for example, the expansion of platform power) are not 'colonialist' (2019, p. 417) because, manifestly, they are not characterized by brutal physical violence. 'Other concepts', they insist, 'such as 'data extractivism'... correctly describe the processes of exploitation and subjectification' (2019, p. 417). We do not deny – indeed we emphasize – that the presence/absence of physical violence appears a point of difference between historical and contemporary data colonialism (Couldry & Mejias, 2019b, p. xviii). The question is what that difference signifies. As social relations between the colonized and the colonized developed, physical violence was replaced by other methods, including symbolic forms of violence such as legal and technological constructs of 'scientific' data collection (Jasanoff, 2006). The absence of physical violence in today's data colonialism merely confirms the multiplicity of means by which dispossession can, as before, unfold.

What emerges from this brief review is a recognition of the general danger inherent in the discourses of datafication and Big Data, even if it it is actualized and resisted in specific contexts.<sup>8</sup> A disagreement with those who argue that contemporary data practices' lack of brutal physical violence disqualifies them from being colonial brought out a key feature of our approach, which is to see at colonialism's core not particular methods (its methods were always complex and variable), but a consistent vision of the world's resources, economic and cognitive, that grounds universal claims to those resources by a particular few. While colonial practice always has a material basis - a landgrab of particular types of resource - the underlying rationale remains the same: as Aníbal Quijano put it in his discussion of coloniality, 'the specific cosmic vision of a particular ethnie... [understood] as universal rationality' (2007, p. 177). This is how the provincial discourse of Big Data today claims privileged application everywhere. The data colonialism thesis foregrounds this continuity with the epistemic violence (Ricaurte, 2019) of earlier colonialism. Here, not just in physical violence, lies the core point exposed by data's decolonial turn: the self-rationalizing appropriation of resources on a vast global scale. Far from serving to 'occlude' this or other aspects of historical colonialism (Crawford, 2021, p. 254), decolonial approaches highlight those continuities, so exposing the colonial core of today's data practices.

8 Compare Kotliar (2020) on how cultural difference is reproduced and constructed in data processes.

#### 4. What's at stake in a decolonial approach to data?

We have mapped the wider terrain of work that has contributed to the decolonial turn in data research and, in the last section, clarified the epistemic violence whose recognition is at the core of that turn. Now we can summarize what is at stake in this decolonial turn more generally. First, it offers an overall theory of data extractivism that can articulate practices both in the Global North and the Global South, both in the West and in China and India. The decolonial turn links data extraction everywhere not just to capitalism, but to capitalism's colonial underpinnings; not just to neo-colonial locations, but to Global North sites where the legacy of colonialism lives on. It sees the epochal expansion of data accumulation since World War II as a stage in the evolution of both colonialism and capitalism understood together as tools for exploiting human life for power and for capital.

Second, through the double lens of colonialism and capitalism, this approach foregrounds data extraction's implications not just for profit, but for the governing of human life and freedom. While freedom is also a concern of Zuboff's 'surveillance capitalism' thesis, her critique addresses the citizens of the USA and Europe, and is abstracted from the long histories of hierarchization between and among peoples that emerged through colonialism (Ricaurte, 2019). By emphasizing data extraction's continuity with colonialism's histories of ruling and dividing people on a global scale, the decolonial turn foregrounds datafication's continuities with historic forms of inequality, rather than seeing datafication as an aberration of late modern Western democracies.

Given that the roots of today's Big Data discourse can be located in the longer history of colonialism's distortions of knowledge through power (as uncovered by Quijano's concept of coloniality),<sup>9</sup> it is vital that challenges to data colonialism are grounded not in supposed universal 'Western' values, but in attempts to forge a common basis for conceptualizing freedom and autonomy on a global scale that work against, rather than with, the history of coloniality (Couldry & Mejias, 2019 chapter 5, drawing on Dussel, 1985; cf. Wynter, 2003; Escobar, 2018). The deep links between contemporary data practices and historic inequalities in the distribution of the world's resources over the past five centuries are obscured by analyzing data and technology exclusively from the perspective of capitalism. Meanwhile the deep racialization of contemporary technological practices can be captured through the double lens of colonialism's and capitalism's evolving interrelations, an approach also informed by the concept of coloniality developed by decolonial scholars (Quijano, 2007; Mignolo & Walsh, 2018).

There remain differences in the exact causal weight that various proponents of the decolonial turn give to data's role in these transformations. One version ('digital colonialism': Kwet, 2019) argues that data practices sustain themselves, like all other capitalist practices, in inherited neo-colonial contexts (for example, the advantages Facebook inherits when it does business in Africa). The 'technocolonialism' thesis (Madianou, 2019) argues by contrast that data practices take a specific form in neo-colonial settings, because of the distinctive

<sup>9</sup> Norbert Weiner's reflections on the risks of computer networks at the dawn of the computer age remain relevant (Wiener, 2013, p. 27 [o.p. 1948]).

types of social and economic power which characterize those settings. The most ambitious causal claim is made by the data colonialism thesis (Couldry & Mejias, 2019a, 2019b; Ricaurte, 2019), namely that practices of data appropriation and processing are themselves a distinctive new type of resource extraction with global significance and, as such, represent a historic new phase of colonialism: a new asymmetric mode of dispossession through data based in an enduringly colonial conception of the world's material and immaterial resources, and the entitlements that supposedly flow from them. Its proposal is that a new form of colonialism emerges not simply when things are digitized, but when aspects of our lives are converted into digital data for the purpose of generating profit: hence the term 'data colonialism', not digital colonialism. Because it offers a general theory of extraction, not just a geographically specific one, the data colonialism thesis is not limited to particular historical sites of colonial extraction.<sup>10</sup> This new colonial appropriation can occur wherever the resources it extracts are situated, which means potentially anywhere (since human life is everywhere), even if its consequences are particularly malign where it overlaps with historic colonialism's legacy.

All three theses – data colonialism, technocolonialism and digital colonialism – regard the racialized inequalities found, for example, in artificial intelligence practices as not incidental, but inherent to the data practices they analyze. But only the data colonialism thesis recognizes a further possibility: that additional segmentations of the world's resources will emerge through data extractivism that both build upon and extend the racisms inherited from historic colonialism.

Decolonial approaches to data have other broad advantages. They contextualize today's data practices and uses of technology in terms of the historical *longue durée* of the past 500 years, not just the past 15 years of social media, or even the past 40 years of internet development. This allows richer explanations of the force and effectiveness, for example, of Facebook's and Google's claims to be privileged players in Africa and Asia, as supposed facilitators of global community. The approach also helps us better grasp the demands of Western corporations and donors in so-called 'underdeveloped' settings, and it generates striking historical parallels for the extraordinary claims to justify access to data which have become standard in the Global North over the past decade.

A colonial reading of what's going on with data, in essence, better addresses the scope of today's transformation of society through data extraction. Just as historical colonialism was an extractive model that reorganized life at every level, the decolonial turn focuses not just on particular vectors, such as social media and search engines, but on wider habits of data collection right across economic life. Solutions to the harm done by data colonialism therefore require profound societal change, and not just the reform of particular corporations, as Zuboff (2019) appears to suggest. In addition, just as historical colonialism represented an entire social and economic order and not merely a set of specific economic rules and government mechanisms, a decolonial approach sees in contemporary data practices the potential emergence of a new way of governing and distributing

Anticipating this point, Achille Mbembe reflects on 'the very distinct possibility that human beings will be transformed into animate things made up of coded digital data' in a 'new collusion between the economic and biological', affirming the global nature of the struggle against data colonialism: 'across early capitalism, the term 'Black' referred only to the condition imposed on peoples of African origin... Now, for the first time in human history, the term 'Black' has been generalized' (2017, pp. 5–6). power in societies and economies. Resistance to such data practices must, in turn, be equally transformational. Lastly, a decolonial approach takes seriously the excuses and justifications offered during historic colonialism for massive resource extraction, and their anchoring in larger frameworks of rationality, progress, order, science and modernity (even salvation!). Those justificatory frameworks are still of course evolving, just as they took decades to stabilize following Spain and Portugal's seizure of the 'New World' (Pagden, 1987). It is therefore not yet possible to offer a definitive reading of those rationalities. All we insist upon is to view the contemporary discourses of Big Data, AI for Social Good, and the like through the long-term historical lens of attempts to justify the unequal distribution of the world's resources that began in earnest 500 years ago. The most useful concept for doing this is not capitalism (or even its ideological cognate term, neoliberalism) but coloniality. No other concept captures better the hubris of Big Data rhetoric which insist that only through the maximal collection and concentration of data can the world be developed, understood, governed, and saved. The decolonial turn treats the ideologies of 'dataism' (Van Dijck, 2014; Harari, 2016) with the utmost seriousness, but understands them in the extended historical perspective of colonialism, just as capitalism itself can best be understood through the additional lens of colonial relations.

What the decolonial turn (including our particular version of it, the data colonialism thesis) does not do, however, is ignore or minimize the differences between data colonialism and historic colonialism. It would be absurd to expect large-scale transformations on the scale of historic colonialism to repeat themselves in some neat mirror-image. Alongside the continuities just noted, important differences between, for example, data colonialism and historic colonialism are clear. First, because data colonialism builds on the social relations of more than two centuries of capitalism, physical violence plays a lesser role in establishing it, which is not to deny the violence of data colonialism in other dimensions (symbolic, economic), or the fact that it can combine in forms of rule that amount to physical force. Second, because of the globalization of science and economy, data colonialism is necessarily global, but operates around two distinctive poles, China and the West, with other intermediate players. Third, data colonialism involves new modalities of oppression. Historic colonialism was based on the violent seizure of land and physical resources, managed through the oppression of bodies. Data colonialism may be apparently less disruptive of ways of life, and unfold within very different temporalities, but promises to have fundamental long-term impact.

Notwithstanding these differences from historic colonialism, the data colonialism thesis, and the decolonial turn generally, offer the most comprehensive perspective – historically, geopolitically, societally – on the unfolding contemporary developments of data and technology. Such inclusivity enables us to imagine a common reference-point for global resistance to data and technology practices. That is the possibility to which we turn in our conclusion.

# 5. What's Next for a decolonial approach to data and technology?

By reviewing debates around the coloniality of data, we have sought to clarify the broader common ground that underlies critical work that insists, implicitly or explicitly, on a decolonial approach to data and technology. But what are the practical and policy implications of taking a decolonial perspective on data and technology? To begin, any struggle against the colonialism advanced through data practices must be global in its framing. The corporate ambitions of data industries are always global, just as capitalism and colonialism are global, and the framing of those ambitions appeals to a mandate that pretends to address problems for the whole of humanity. Following Milan and Treré (2019), we have emphasized that resistance against data colonialism is always local and must use, and contend with, local conditions. But having acknowledged that, the frame within which such activism unfolds must always account for the global scale of what it is opposing, if it is not to miss its target. Whatever the local dynamics, data colonialism – like historical colonialism – operates within and takes advantage of a global framing, and it is there that its contestation must begin.

We should also acknowledge that decolonial struggles against data extraction must always be double: struggles over particular practices of technology and struggles over knowledge and rationality (that is, the deep narratives that help frame those specific practices and uses). Regardless of how intense their engagement with the materiality of tools and infrastructures, these decolonial struggles are always also acts of imagination, construction alternative imaginings of data and technology (Ricaurte, 2019). Sometimes they will build on existing social imaginaries (Taylor, 2004); other times, they will reject all existing alternatives and forge a different direction.

Next, we need to keep in mind that given the massive global inequalities and asymmetries of cognitive and economic production at least since the start of historic colonialism, decolonial struggles against data extractivism must include, and be defined by, the widest possible range of peoples. The involvement of indigenous and marginalized peoples, for example, is essential (one representative example is the Abolitionist Futures project articulated by Emmer et al., 2020, that seeks to analyze how 'Queer, Trans, Two-Spirit, Black, Indigenous, and People of Color communities are disproportionately impacted by surveillance and criminalization' technologies).<sup>11</sup> Otherwise the practice of struggle will only serve to reinforce the colonialism it claims to challenge. This will require new opportunities to enable those who are typically excluded from debates on the global economy, legal policy, and computing standards to lead those debates.

To conclude, and following from the last point, we must recognize that a new conceptual space must be built that seeks to define and claim techno-social spaces beyond the profit-motivated model of Silicon Valley and the control-motivated model of the Chinese Communist Party, the two centers of power of the new colonial extractivist order. That is why we are also involved, along with our colleague Juan Ortiz Freuler, in the creation of a Non-Aligned Technologies Movement (NATM), a movement which mirrors in some ways the original collective of non-aligned states which, during the Cold War, sought to define an alternative to capitalism and communism.

Whether this reincarnation of the non-aligned movement will succeed, only time will tell. At a time when even conservative states are seeking to disentangle themselves from their dependency on the USA and China, it is clear that 'technological

<sup>11</sup> Compare Salas and Srivastava (2020). Important also here is the broader 'design justice' movement (Chock, 2020, pp. 190–204).

sovereignty', when practiced as nationalism and protectionism, can be as much of a threat as data colonialism (cases like Russia and Iran are illustrative). In contrast, NATM is predicated on the following principles: the boycotting of extractivist technologies and the use of alternative tools; divestment at local and national government levels from Big Tech (by not buying or accepting their 'free' products); the re-appropriation of data (and the products of data) on behalf of those who generate it; the implementation of taxes and sanctions against Big Tech to repair the damage done by their technologies; the bolstering of public education – in the form of citizen research, literacy campaigns, decolonial thinking – to understand the dangers of data colonialism; the promotion of a broad and diverse culture of non-alignment to reimagine new forms of community without extractivist technologies and their heavy costs, not least to the physical environment; and finally, the building of a solidarity that joins non-aligned individuals and communities globally through collective imagination and action.

The NATM is one initiative among many. But it exemplifies how the questions of definition, coherence, and transhistorical relevance with regard to data extractivism on which this article has concentrated are more than topics for academic debate. They are central questions around which practical alliances can be built everywhere for resisting the intensified colonial processes in the field of data and technology that we and many others have sought to name.

#### Disclosure statement

No potential conflict of interest was reported by the author(s). Note on contributors

#### References

Alhassan, A., & Chakravartty, P. (2011). Postcolonial Media policy under the long shadow of empire. In R. Mansell, & M. Raboy (Eds.), *The handbook of global media and communication policy* (pp. 366–382). Wiley-Blackwell.

Ali, S. M. (2016). A brief introduction to decolonial computing. *XRDS: Crossroads, The ACM Magazine for Students*, 22(4), 16–21. https://doi.org/10.1145/2930886

Amrute, S. (2016). Encoding race, Encoding class. Duke University Press.

Aouragh, M., & Chakravartty, P. (2016). Infrastructures of empire: Towards a critical geopolitics of media and information studies. *Media, Culture & Society*, 38(4), 559–575. https://doi.org/10. 1177/0163443716643007.

Arora, P. (2019). The Next Billion users. Harvard University Press.

Avila Pinto, R. (2018). Digital sovereignty or digital colonialism? *SUR*, 27. https://sur.conectas.org/ en/digital-sovereignty-or-digital-colonialism/

Baptist, E. (2016). Toward a political economy of Slave labor: Hands, whipping machines and modern power. In S. Beckert & S. Rockman (Eds.), *Slavery's capitalism* (pp. 31–61). University of Pennsylvania Press.

Beckert, S. (2014). Empire of cotton. Vintage.

Beckert, S., & Rockman, S. (Eds.). (2016). Slavery's capitalism. University of Pennsylvania Press.

Benjamin, R. (2019). Race after technology. Polity.

Bhattacharyya, G. (2018). Rethinking racial capitalism. Rowman and Littlefield.

Broussard, M. (2019). Artificial unintelligence. MIT Press.

Browne, S. (2015). Dark matters. Duke University Press.

Bulhan, H. (2015). Stages of colonialism in Africa: From occupation of land to occupation of being. *Journal of Social and Political Psychology*, 3(1), 239–256. https://doi.org/10.5964/jspp.v3i1.143

Calzati, S. (2020). Decolonising "data colonialism": propositions for investigating the realpolitik of today's networked ecology. *Television & New Media*. https://doi.org/10.1177/1527476420957267

Can Facebook Connect the Next Billion? (2017, July 27). Global voices Advox. https://advox.globalvoices.org/2017/07/27/can-facebook-connect-the-next-billion/

Casati, R. (2013). Contro il colonialismo digitale. Laterza.

Casilli, A. (2017). Digital labor studies go global: Towards a digital decolonial turn. *International Journal of Communication*, 11, 3934–3954.

Castells, M. (1996). The rise of the network society. Blackwell.

Chan, A. (2014). Networking peripheries. MIT Press.

Chock, S. C. (2020). Design justice. MIT Press.

Cohen, J. (2019). Between truth and power. Oxford University Press.

Connell, R. (2007). Southern theory. Polity.

Cooke, B. (2003). The Denial of Slavery in Management Studies. *Journal of Management Studies*, 40(8), 1895–1918.

Cottom, T. (2020). Where platform capitalism and racial capitalism meet: The Sociology of race and racism in digital society. Sociology of Race and Ethnicity, 6(4), 441–449. https://doi.org/10.1177/2332649220949473

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. Stanford University Press.

Crawford, K. (2021). Atlas of Al. Yale University Press.

Dalton, C. M., Taylor, L., & Thatcher, J. (2016). Critical data studies: A dialog on data and space. *Big Data & Society*, 3(1), 205395171664834. https://doi.org/10.1177/2053951716648346

D'Ignazio, C., & Klein, L. (2020). Data feminism. MIT Press.

Dörre, K., Lessenich, S., & Rosa, H. (2015). Sociology, capitalism, critique. Verso.

Dourish, P., & Mainwaring, S. D. (2012). Ubicomp's colonial impulse. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 133–142. https://doi.org/10.1145/2370216.2370238

Dussel, E. (1985). Philosophy of Liberation. Eugene, OR: Wipf and Stock.

Emmer, P., Rivas, C., Salas Neves, B., & Schweidler, C. (2020). *Technologies for Liberation: Toward Abolitionist futures*. Astraea Lesbian Foundation for Justice. https://www.astraeafoundation.org/FundAbolitionTech/

Escobar, A. (2018). Designs for the pluriverse. Duke University Press.

Eubanks, V. (2018). Automating inequality. St. Martin's Press.

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

Halkort, M. (2019, October 15). On the Coloniality of Data Relations: Revisiting Data Colonialism as Research Paradigm. *Datactive*. https://data-activism.net/2019/10/bigdatasur-on-the-coloniality-of-data-relations-revisiting-data-colonialism-as-research-paradigm-12/

Harari, Y. (2016). Homo deus. Vintage.

Harvey, D. (2004). The 'new' imperialism: Accumulation by dispossession. Socialist Register, 40, 63-87.

Irani, L., Vertesi, J., Dourish, P., Philip, K., & Grinter, R. (2010). Postcolonial computing: A Lens on Design and Development. *Proceedings of the SIGCHI Conference on Human Factors in Computing* (pp. 1311–1320).

Jasanoff, S. (2006). Biotechnology and empire: The global power of seeds and science. Osiris, 21(1),273–292. https://doi.org/10.1086/507145

Katz, Y. (2020). Artificial whiteness. Columbia University Press.

Kotliar, D. (2020). Data orientalism: On the algorithmic construction of the non-western other. *Theory and Society*, 49(5–6), 919–939. https://doi.org/10.1007/s11186-020-09404-2

Kwet, M. (2019). Digital colonialism: US empire and the new imperialism in the global south. *Race & Class*, 60(4), 3–26. https://doi.org/10.1177/0306396818823172

Kwet, M. (2021). Digital colonialism: the evolution of American empire. *Roar*, 3 March. https://roarmag.org/essays/ digital-colonialism-the-evolution-of-american-empire/

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, 53–125.

Madianou, M. (2019). Technocolonialism: Digital innovation and data practices in the humanitarian response to refugee crises. *Social Media & Society*, https://doi.org/10.1177/2056305119863146

Magalhães, J., & Couldry, N. (2021). Giving by taking away: Big tech, data colonialism and the reconfiguration of the social good. *International Journal of Communication*, 15, 343–362.

Maldonado-Torres, N. (2007). On the coloniality of being: Contributions to the development of a concept. *Cultural Studies*, 21(2–3), 240–270. https://doi.org/10.1080/09502380601162548

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big Data: the next frontier for innovation, competition, and productivity*. McKinsey Global Institute. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation#

Marx, K. (1976). Capital volume I. Harmondsworth.

Mbembe, A. (2017). Critique of black reason. Duke University Press.

McIlwain, C. (2020). Black software. Oxford University Press.

Melamed, J. (2015). Racial capitalism. *Critical Ethnic Studies*, 1(1), 76–85. https://doi.org/10.5749/jcritethnstud.1.1.0076

Mezzadra, S., & Neilson, B. (2018). The politics of operation. Duke University Press.

Mhlambi, S. (2020). From rationality to relationality. Carr Center Discussion Paper Series, 2020–009. https://carrcenter.hks.harvard.edu/publications/rationality-relationality-ubuntu-ethical-and-human-rights-framework-artificial

Mignolo, W. (2011). The darker side of modernity. Duke University Press.

Mignolo, W., & Walsh, C. (2018). On decoloniality. Duke University Press.

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

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy and Technology*, 33(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

Nhemachena, A., Hlabangane, N., & Kaundjua, M. (2020). Relationality or hospitality in twenty-first century research? Big data, internet of things, and the resilience of coloniality in Africa. *Modern Africa: Politics, History and Society*, 8(1), 105–139. https://doi.org/10.26806/modafr.v8i1.278

Noble, S. (2018). Algorithms of oppression. NYU Press.

Nothias, T. (2020). Access granted: Facebook's free basics in Africa. *Media Culture & Society*, 42 (3), 329–348. https://doi.org/10.1177/0163443719890530 Pagden, A. (1987). The fall of natural Man (2nd ed). Cambridge University Press.

Pomeranz, K. (2001). The great divergence (Revised ed.). Princeton University Press.

Powell, A. (2021). Undoing optimization. Yale University Press.

Quijano, A. (2007). Coloniality and modernity/rationality. *Cultural Studies*, 21(2–3), 168–178. https://doi.org/10.1080/09502380601164353

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

Robinson, C. (2000). Black marxism (2nd ed.). University of North Carolina Press.

Salas, B., & Srivastava, M. (2020). *Technologies for Liberation: Toward abolitionist futures*. Astraea Lesbian Foundation for Justice. https://www.astraeafoundation.org/FundAbolitionTech/

Schiller, D. (2000). Digital capitalism. MIT Press.

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

Siebel, T. (2019). Digital transformation. Rosetta Books.

Siggelkow, N. & Terwiesch, C. (2019, May–June). The age of continuous connection. *Harvard Business Review*, 64–73.

Soledad, M., & Waisbord, S. (2019). Between Data Capitalism and Data Citizenship. *Television and New Media*, 20(4), 412–419.

Solon, O. (2017, July 27). 'It's digital colonialism': How Facebook's free internet service has failed its users. *The Guardian*. https://www.theguardian.com/technology/2017/jul/27/facebook-free-basics-developing-markets

Srnicek, N. (2017). Platform capitalism. Polity.

Taylor, C. (2004). Modern social imaginaries. Duke University Press.

Thatcher, J., O'Sullivan, D., & Mahmoudi, D. (2016). Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environment and Planning D: Society and Space*, 34(6), 990–1006. https://doi.org/10.1177/0263775816633195

Todd, Z. (2016). An indigenous feminist's take on the ontological turn: 'ontology' is just another word for colonialism. *Journal of Historical Sociology*, 29(1), 4–22. https://doi.org/10.1111/johs.12124.

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. Surveillance & Society, 12(2), 197–208. https://doi.org/10.24908/ss.v12i2.4776

Wiener, N. (2013). Cybernetics (2nd ed.). Martino Publishing.

Williams, E. (1994). Capitalism and slavery. University of North Carolina Press.

Wolfe, P. (2016). Traces of history. Verso.

World Economic Forum. (2013). Unlocking the value of personal data. http://www3.weforum.org/docs/WEF\_IT\_ UnlockingValuePersonalData\_CollectionUsage\_Report\_2013.pdf

Wynter, S. (2003). Unsettling the coloniality of being/power/truth/freedom. CR: *The New Centennial Review*, 3(3), 257-337. https://doi.org/10.1353/ncr.2004.0015

Zuboff, S. (2019). The age of surveillance capitalism. Profile Books.