Critiques of Data Colonialism

*Benedetta Brevini, Irene Fubara-Manuel, Clément Le Ludec,
Jakob Linaa Jensen, Andrea Jimenez,[1] and Jo Bates[2]*

Introduction

In this chapter, we critically interrogate the concept of data colonialism developed by Couldry and Mejias (2019). For them, data production activities occur casually as Internet users engage in social activities on the services of big tech companies. As with other research on digital capitalism, their focus is on the process of appropriation and value creation of data produced by users of social networks (Scholz, 2017). They see it as 'an emerging order' of appropriating and extracting 'social resources for profit through data', without any recognition and valorization of the work behind those data.

Many scholars in the last two decades have investigated the unprecedented consolidation of communication systems, exploring the increasing globalization of capitalist relations and rationales; the concerted pursuit of neoliberal economic policies, and the roll-out of digital technologies with its progressive datafication of everyday life. It is hard to think of an aspect of life that has not been affected by the use of algorithms, automation, and big data, including medicine, education, welfare, voting, dating, communication, law enforcement, warfare, and cyber-security. The focus on data accumulation has occupied centre stage in several recent works, including Fuchs' big data capitalism (2019), Zuboff's surveillance capitalism (2018), and Srnicek's platform capitalism (2017). The common themes of these works are that data has a crucial value, and data collection is highly unequal, tending to replicate existing power asymmetries.

Nick Couldry and Ulises Mejias go further and, drawing on their recent book, *The Costs of Connection* (2019), argue that the specificity of contemporary forms of data extraction is best understood by seeing it as a genuinely new stage of colonialism which is based around the appropriation

not (as in historic colonialism) of land, its resources, and the bodies to work those resources, but the flow of human life itself, made valuable in the form of data. This thesis of 'data colonialism' is offered as a characterization not just of what is happening with data in historic colonies, but right across the world, including the Global North. These processes of data appropriation and exploitation are the source of the valuation of major big tech companies such as Facebook or Google through their stock market value. In view of their systemic dimension in terms of the number of users and as 'lead firms' structuring the web, these companies have been described as 'sovereign', encroaching on competences traditionally reserved to the state. In this chapter, we each present a critique of Couldry and Mejias's arguments about the emergence of a new stage of colonialism.

Fubara–Manuel argues that before data capture became computational, it was colonial. They interrogate early scholarship that introduces the metaphor of computation as colonization linking this history to contemporary concepts such as 'data colonialism'. Fubara–Manuel argues for scholarship and research that examines colonization as a tangible and pre-existing system of oppression with a violent data legacy that is continued through computation. They posit that computational methods of data capture, especially in post-colonial relations such as global migration, are a second-degree extraction of formerly colonized people that should not be taken for granted in any intersecting analysis of data and colonization.

Le Ludec goes on to suggest that 'data colonialism' is a fundamental misconceptualization. Using the example of a fieldwork conducted in Madagascar, he argues that Internet users who are allegedly subject to data colonialism cannot be treated in the same way as Madagascar's data producers. Rather than talking about data colonialism, he argues, we should thus be talking about 'data extractivism' or 'digital colonialism'. He argues that these latter concepts better account for the inscription of the digital industry into old-fashioned colonial oppression mechanisms.

Brevini then joins the discussion to argue that, rather than a fundamental misconceptualization, there is a gap in the data colonialism concept with regard to the environment. She argues that an intervention is needed in the data colonialism discourse that introduces an environmental justice critique that takes into account the outstanding environmental harms of data-driven communication technologies development. In all their variety of forms, communication technologies rely on large swathes of land and sea, greenhouse gas-emitting machines and infrastructures that deplete scarce resources in their production, consumption, and disposal. They require increasing amounts of energy, water, and finite resources (Brevini, 2020, 2021). She draws the reader's attention to the diverse ways in which data technologies, including energy-intensive artificial intelligence (AI) models and data centres have environmental impacts that go far beyond their carbon

footprints and can only be grasped by exploring the global production/ supply chain and life cycle of communication technologies.

The chapter ends with Linaa Jensen returning to Fubara-Manuel's claim that 'before capture became computational, it was colonial'. He claims that colonial logics of exploitation can be found in the earlier historical context of medieval feudalism. Medieval society, created with strong hierarchies, was based not only on exploitation but also on regimes of visibility and public shaming. In line with his recent book, *The Medieval Internet* (2020), he argues that the contemporary public sphere, mediated by digital technologies, resembles the medieval village with mutual surveillance and mechanisms of public shaming such as pillories and other public punishments. Today, he argues, such regimes of shaming are digitally mediated but contribute to uphold the same regimes of dominance on a micro and macro level as found in medieval societies.

Through the insights in these four contributions, we therefore aim to make a contribution to existing debates about the colonial dynamics of the current era, opening up new pathways for understanding and conceptualizing contemporary logics of extraction and capture.

A second–degree extraction: on capture and colonial violence – Irene Fubara-Manuel

Early scholarship on the role of technology in reorganizing labour and occupational communities deployed a metaphor of computation as colonization (Agre, 1995). This metaphor has continued through to contemporary concepts such as 'data colonialism' that connect data capture to capitalist extraction (Thatcher et al, 2016; Couldry and Mejias, 2019). A broad application of this metaphor does not account for the tangible effects of colonization as an existing system of oppression that pre-dates modern computation. In reducing colonization to a metaphor, the violent history of colonization and its continuation in modern computing are not adequately interrogated.

Capture is one of the two models of data privacy according to Agre (1994). Instead of the older model of surveillance that relied on vision – for instance, state-sanctioned CCTV recording the lives of citizens – smart devices track and capture the everyday minutiae of workers and consumers. For Agre (1994: 112), 'capture is never purely technical but always socio-technical in nature' in that it incorporates pre-existing political ideology into categorizing activities and requires imposition of technologies that change systems of meaning for a group of people. It is this computational imposition that Agre (1995) defines as colonization. However, computational imposition and extraction on their own do not equate to colonization. The digitization of the workflow of a border agent, for instance, is not the same as the systemic

subjugation of a people and their culture. As a system of power, colonization has its own logic of imposition that is rooted in violence and dehumanization.

Colonization is a form of 'domination and submission' that instrumentalizes and appropriates indigenous people and their land – 'colonization = "thing-ification"', according to Césaire (2000: 42). Before computers could reduce human activities and 'thing-ify' them as profitable data, colonial apparatuses had mapped, dissected, and categorized indigenous people and their land as commodities. Recall the early development of biometric technologies with the handprint of the Bengalese road contractor, Rajyadhar Kōnāi. William Herschel captured Kōnāi's body, as alternative to a signature on a deed. Herschel then sent this handprint to Francis Galton (Pearson, 1914), who would later establish the science of fingerprinting. By placing Kōnāi's handprint in the archives of colonial violence, capture becomes a symbolic dissection of a colonial subject. Kōnāi's handprint becomes a synecdoche – biometric data stands in for the whole person. The capture of this print, not only reduces Kōnāi, but also commodifies his body and labour as property of the British Empire. In Kōnāi's handprint is a symbol of the colonial imperative of transparency and appropriation of indigenous people as raw data (Fubara-Manuel, 2019).

More than a century after Kōnāi's handprint, biometrics are widely applied in migration in the enforcement of borders and the categorization of border crossers. The pattern of thing-ification remains the same. In the aftermath of British colonization, racialized citizens from former colonies face more scrutiny at the UK border (Cellan-Jones, 2020; McDonald, 2020). They are required to produce more data to verify their identities. The heightened scrutiny can be traced back to visa impositions placed during the 1960s to 1980s, as a response to the problem of 'coloured immigration' (Fubara-Manuel, 2020; Patel, 2021). In the 1940s and 1950s, their parents might have been able to live and work in the UK using their British colonial passports that gave them rights to citizenship. The visa impositions of the 1960s and 1980s meant that migrants from former African and South Asian British colonies submitted more data to cross the UK border. As with Kōnāi's handprint, placing modern visas in the archives of colonial violence, there is a continuation of submission. In addition to the submission of bodies, land, and culture, there is a submission of data.

While it is important to place the capture of migrants in the aftermath of colonization and empire, it is also necessary to study the specific changes digital technology contributes to the oppression of formerly colonized people. Big data as a contemporary technological development, means that one body part does not only stand in for a person, but that one migrant stands in as a prediction for a group of 'suspect nationalities' or preferred travellers, as in the UK's streaming tool algorithm (Joint Council for the Welfare of Immigrants, 2020). International biometric databases such as the

EURODAC (European Asylum Dactyloscopy Database) store the captured fingerprints of asylum seekers in the EU to keep their bodies tethered to the border in spite of their physical location (Amoore and Hall, 2010; European Commission, 2022).

Before capture became computational, it was colonial. The key intervention in this chapter is a call to acknowledge and study the form of second-degree extraction that further reduces thing-ified colonial subjects to 'suspect nationalities', high-risk border crossers, or offenders. As with capitalist expansion, colonization reduces people to data and profit. Separate from capitalist data extraction, there is no feigning of consent, there are no terms of service, but only violence and dehumanization. Socio-technical analysis of the pre-existing thing-ification of indigenous people will further uncover the historical methods of commodification that are not only mimicked but deployed and perpetuated by state-owned and commercial technology companies. It would also contribute to an understanding of indigenous modes of resistance against colonial and computational capture.

Behind data production for AI: old-fashioned post-colonial processes – Clément Le Ludec

For Nick Couldry and Ulises Mejias (2019), today's 'datafication' – or turning the world into data – can be compared to the social quantification of the 19th century in that it allows individuals to be categorized (p 119). Similarly, today's cloud firms are extending yesterday's imperial logic in their quest to appropriate data. This extraction of value is also facilitated and extended by the possibilities of coordination made possible by digital technology. Finally, these players control the logistical infrastructure, which gives them some power over other firms: owning a large number of data centres enables them to offer data storage and computing power services to other, smaller firms.

However, Nick Couldry and Ulises Mejias's (2019) analysis is based on the implicit assumption of a digital specificity that would allow this colonialism of data, an assumption that is questionable insofar as (neo)colonial logics also persist in analogue or non-digital environments. Thus, the question remains whether data colonialism simply refers to the expansion and acceleration of data collection, or whether there are characteristics unique to data-driven colonialism.

Based on our own research on data workers working for French AI companies in French-speaking African countries, and on Latin-American data workers, we argue that it is questionable whether it is really possible to equate the extraction of value from data produced by consumers in Northern countries with the extraction of value from international outsourcing chains for the production of AI. Does the term 'colonialism' apply in the same way? Can we put low-paid workers in Madagascar who spend 40 hours a week

'processing' data on the same level as influencing or Instagram users who post pictures? Outsourced work thus represents a cost in human resources. The 'work' of social network users, on the other hand, enables big tech to earn money directly through advertising. Moreover, can we equate the work experience of these two categories of workers? The former, on the other side of the planet, would rather work to survive, while the latter would benefit from the economic security of the countries of the North.

As part of the DIPLab's project on data workers, we interviewed 147 workers, managers, and directors in ten Malagasy companies (Le Ludec and Cornet, 2023). At the same time, we sent a questionnaire to 296 data workers based in Madagascar. They usually earned between 96 and 126 euros a month, with a huge gap between their pay and that of the team leaders, who are also usually in Madagascar and work in the country, but take home eight to ten times as much (Le Ludec et al, 2023).

We think it might be more appropriate to talk about 'data extractivism'. In fact, the notion of extractivism is already well established in the literature in relation to other industries, such as textiles or mining. First, it refers to the appropriation of value by Northern countries and corporations and the alteration of the source of that value (mines, fields). It also refers to the centralization of economic and social power by the same actors in the countries of the North, based on inequalities that are sometimes pre-existing, especially post-colonial relations. These relationships are sometimes violent (in terms of working conditions) and non-reciprocal (little or no investment in the country).

The work of Aneesh Aneesh (2015) on the communicative and affective work of the employees of service centres, who, thanks to the neutralization of their accent and mimesis, make it possible to communicate with American consumers, is very fruitful to illustrate this post-colonial dynamic. They also show the impact on workers of the adaptation of these centres to the constraints of the American market, particularly in terms of working hours. Still in call centres, this time in the Philippines, Jan Padios (2018) speaks of a process of 'colonial recall', showing how post-colonial relations are being redefined as added value for American companies.

As well as cheap labour, the data industry benefits from a well-educated workforce – most have been to university and speak fluent French, learnt at school, online or in Institut Français classes. Founded in 1883, the Institut Français was originally set up to extend imperial power through language to colonized populations. Workers are thus encouraged to develop the colonial skills needed for the export-oriented data and outsourcing industry. This is also where we feel the 'extractive' dimension is more fruitful, as the remnants of the old colonial order are now being joined by new players who are shaping the acquisition of skills and access to the labour market for young Malagasy workers in Antanarivo.

Antonio Casilli (2017) proposes another framework for analysing digital labour in the Global South, based on the call for a 'digital decolonial turn' to make visible the invisible labour performed by marginalized workers. An interesting contribution is the way the author defines digital labour platforms: on-demand platforms such as Upwork or Airbnb, micro-work platforms such as Amazon Mechanical Turk, online social platforms such as Facebook, and 'smart' platforms such as connected object websites. Taking a post-colonial perspective, Casilli (2017: 3943) argues that 'platform capitalism operates by leveraging social constructs of race and gender'. With this in mind, he outlines a number of theories linked to the concept of 'digital colonialism', which focuses primarily on the data imperialism practised by large technology companies.

To extend this latter structural analysis, Casilli argues that the concept of 'coloniality', which refers to 'long-standing patterns of power that emerged as a result of colonialism, but which define culture, labour, intersubjective relations and knowledge production far beyond the strict confines of colonial administrations', is even more accurate than 'digital colonialism' as it also 'recounts the assumptions of dominant discourses while seeking emancipation for all marginalized identities at work' (Casilli, 2017: 3947).

Finally, rather than extending the features of colonialism to the digital oppression of large digital corporations (like in the 'data colonialism' concept), we need to go back to the existing relations of digital colonization. Our fieldwork in Madagascar thus illustrates the direct continuity between colonial institutions and the development of the outsourcing industry. Indeed, the AI sector benefits from a specific policy – the 'tax-free zones' created in 1989 for the textile industry. Since the early 1990s, French companies have been setting up satellites in Madagascar, particularly for the digital publishing industry. These special zones, which exist in many other developing countries, attract investment by offering highly attractive tax exemptions.

Today, of the 48 companies offering digital services in the tax-free zones, only nine are owned by Malagasy citizens, compared with 26 owned by French entrepreneurs. Apart from the situation with formal companies, the sector has developed a practice of cascade subcontracting, with informal companies and entrepreneurs at the bottom of the pecking order, treated poorly and called in when there are labour shortages elsewhere in the sector.

At the end of the AI value chain, workers are struggling with local living standards. Ultimately, they cope with the situation by developing 'colonial skills' which will reversely feed the AI industry with limited long-term benefit for Madagascar. Our fieldwork illustrates why we criticize Nick Couldry and Ulises Mejias's (2019) concept. Hence, we consider that 'digital coloniality' and 'data extractivism' are both concepts that take better

account of the institutional and economic mechanisms underpinning the development of the data industry than 'data colonialism' while restoring workers' voices.

An ecological critique of data colonialism: the hidden environmental costs of data-driven technologies – Benedetta Brevini

In this intervention I argue that data colonialism as a framework misses a crucial component of analysis. It needs an ecological critique that places at centre stage the outstanding environmental damage of data-driven communication technologies development. In all its variety of forms, communication technologies rely on large swathes of land and sea, greenhouse gas-emitting machines and infrastructures that deplete scarce resources in their production, consumption, and disposal. They require increasing amounts of energy, water, and finite resources (Brevini, 2020, 2021).

In order to develop an ecological critique of data colonialism, it is crucial to recognize its environmental costs by exploring the global production supply chain and life cycle of communication technologies, from the extractionism characterizing the mining of metals and minerals that are crucial to produce hardware, to the energy costs of hyper-consumption of data colonialism, to the ecological toll of disposal of the devices employed and its consequent electronic waste (e-waste) and e-pollution emergency.

In her research on digital advancements within humanitarian structures, Mirca Madianou (2019: 2) introduces the concept of 'technocolonialism' to examine how the convergence of digital developments, humanitarian structures, and market forces revitalize and reshape colonial legacies. This persistence of colonial genealogies and inequalities is also evident in the global supply chains' data-driven communication technologies, as the extractive nature of technocolonialism extends to the mining of minerals required for digital technologies. For instance, the increasing demand for mineral resources, driven by the adoption of AI, is projected to lead to a 3500 per cent rise in lithium demand in the European Union by 2050 (EC, 2022). Many mining projects in the Global South are located in indigenous territories. The expansion of mining activities can lead to the forced displacement of indigenous communities from their ancestral lands, often without adequate compensation or consent. Additionally, labour conditions in mines can be hazardous and exploitative, with workers facing unsafe working conditions, inadequate pay, and limited access to healthcare and social protections (Smart, 2017).

Moving along the global supply chain, the production of data-driven communication technologies also carries significant environmental costs. A study conducted by the College of Information and Computer Sciences at

the University of Massachusetts Amherst (Strubell et al, 2019) quantifies the energy consumption associated with running AI programs. In the examined case of a common AI training model in Linguistics, the study found that it could emit over 284 tonnes of carbon dioxide equivalent, equivalent to the lifetime emissions of five average American cars or approximately 100 return flights from London to New York City. Additionally, the convergence of communication systems that support AI generates its own set of environmental challenges, including energy consumption, emissions, material toxicity, and e-waste (Brevini and Murdock, 2017). To this calculation, it is crucial to add the increasingly worrying water consumption of applications such as Chat GPT, particularly alarming considering the growth and increasing complexity of AI models (George et al, 2023). Data centres, which play a crucial role in the global supply chain, consume an average of 200 terawatt hours (TWh) of energy annually, surpassing the national energy consumption of some countries, such as Iran (International Energy Agency, 2017; Jones, 2018). Furthermore, the information and communications technology (ICT) sector, encompassing mobile phone networks, digital devices, and television, accounts for 2 per cent of global emissions (Jones, 2018). Greenhouse gas emissions from the ICT industry could surpass 14 per cent worldwide by 2040, exceeding half of the current contribution from the entire transportation sector (Belkhir and Elmeligi, 2018). The cooling systems of data centres require substantial amounts of water, raising policy concerns in regions such as the US and Australia that have experienced prolonged droughts (Mosco, 2017). As explained on Google's DeepMind website (Evans and Gao, 2016), one of the primary energy-consuming elements in data centres is cooling. The excess heat generated by servers powering Google services such as Search, Gmail, and YouTube necessitates the use of large industrial equipment like pumps, chillers, and cooling towers. DeepMind proposes Machine Learning as a solution to this issue, which, ironically, is also highly energy-intensive and produces carbon emissions.

At the end of the global supply chain, the disposal of devices becomes a pressing concern. When communication devices are discarded, they contribute to e-waste, presenting challenges for local municipalities in terms of safe disposal. This challenge is so overwhelming that it is often outsourced, and numerous developing economies now find themselves serving as digital dumping sites for more affluent nations, as illustrated by the situation in Kenya (Brevini and Murdock, 2017; Naipanoi, 2021).

In sum, the framework of 'data colonialism' fruitfully highlights a series of alarming social justice concerns, denouncing the power inequalities between the entities collecting and controlling data, the violation of privacy rights and sovereignty rights of those from whom the data is extracted. However, it fails to address environmental justice considerations that are, in fact, a crucial

component of social justice as they aim to rectify historical and ongoing inequalities in the distribution of environmental benefits and costs.

The digital medieval village: social media as a new feudalism – Jakob Linaa Jensen

As mentioned by Irene Fubara-Manuel in the first section of this chapter, before capture became computational, it was colonial. As colonialism was a globalized system of inequalities and exploitation, such mechanisms were found much earlier than the colonial era. In the Middle Ages such an economic system was known as feudalism, a term defined by French historian Max Bloch (1961). It denotes a hierarchical system of exploitation where the serfs and peasants were at the bottom of the pyramid. Their work was exploited by the lords in exchange for some kind of protection and stability. On the other hand, the lords themselves were exploited by the kings and nobility through taxes and army conscriptions. In exchange they got land titles and the right to rule those below them. The central good in the medieval economy was labour, just as data is a central good in the platform economy (van Dijck et al, 2018).

At the core of the feudal system were strong logics of visibility and surveillance, organized through spatial practices. Medieval cities were characterized by close-knit communities organized within narrow city walls with ubiquitous surveillance and power hierarchies upheld by public punishments. The penal system was based on public shaming of the offenders, in stocks, wooden horses, or even in the gallows. Punishment was public and spectacular and served to uphold the existing societal order. The socially mediated public sphere similarly combines these two phenomena: visibility and mutual surveillance.

The medieval village forms a powerful metaphor for the logics of sociality and surveillance in the platform economy. Even if not actively commenting or liking, users keep an eye on each other, through what I have elsewhere dubbed as a social sixth sense (Jensen, 2009). Many people write and publish on social media based on anticipated reactions. It is a paradox that the Internet, the ultimate symbol of modernity, transparency, and enlightenment, facilitates logics of enclosure, censorship, and social control.

Social media, like the rest of the Internet, was intended to ease communication and encourage connection, but as it gathered momentum the consciousness of the 'omnopticon' (Jensen, 2007) intruded on the surface of all Internet activities, creating a norm that cripples rather than liberates, making people conform even more to existing norms of what is perceived as mass approved behaviour.

This is particularly striking in the case of what I call 'digital pillories'. In medieval society, the pillory was one of the great instruments of public

shaming. Here offenders were exhibited for shorter or longer times for the purpose of public censure. They received public ridicule, outright hatred, and physical acts such as hitting and spitting from the bypassers. The pillory was a strong instrument of social control.

According to French philosopher Michel Foucault (1975), public shaming and punishment disappeared in modern societies and were instead privatized and hidden within the walls of prisons and asylums. However, public shaming as rituals of power, discipline, and the sustenance of inequalities never really disappeared. It still has a presence in modern society, not least in the age of social media and the culture of visibility. As social media are used as a powerful spectacle for celebrities and ordinary citizens alike, they have also proved to be tools for public shaming and what I have elsewhere defined as digital witch hunts (Jensen, 2020).

Witch hunts are often seen as characteristic of medieval society, although they peaked in the sixteenth and seventeenth centuries (Jensen, 2020). Online witch hunts only deviate from their historic predecessors in the method of punishment. After all, the mob does not have formal judicial power. But the rapid and often wrong judgement of 'the social media court of justice' can do real harm to their victims anyway. There are numerous examples of teenagers who have committed suicides because of online bullying, or innocent people who have been driven to suicide by false accusations of crimes. The judgement by the crowd belongs in archaic contexts such as the medieval marketplace. A particularly striking case is the Boston Marathon terrorist attack, where the wrong persons were accused as a result of a social media-led private 'investigation'.

Not all digital pillories are created explicitly with the purpose of public shaming. In recent decades we have seen an explosion of online rating services, evaluating everything from travel agencies and restaurants to teachers and solicitors. The idea is in line with the touted notion that the middlemen – the tourist guide, the critic, the professional – were now unnecessary. Rather than relying on curated advice by experts, everybody can become a reviewer. The idea of trusting fellow consumers and citizens is immediately appealing. Further, in rating portals the expertise is crowdsourced, Wikipedia-style, while the number of reviews increases the reliability, making one odd review less important. The rating regime is closely connected to logics of appropriation and value creation through commodified crowd-sourcing, as discussed in the introduction.

In short, using the feudal logics of the Middle Ages and the Medieval village as an analytical prism for understanding the contemporary platform economy, we can go a step deeper and further back in understanding contemporary logics of platformization, data colonialism, and hierarchical and asymmetric distributions of power.

Conclusion

In conclusion, in our chapter we propose to re-theorize data colonialism by looking at the continuity of practices of appropriating data-related values with older forms of oppression. Whether we are talking about the roots of biometric technologies in Bengalore (India) or the data industry between France and Madagascar, everything suggests that we need to go back to truly colonial or post-colonial inequalities in order to analyse the development of big tech. Jensen's piece even suggests going deeper, using medieval feudal logic to make sense of contemporary logics of platforming, data colonialism, and hierarchical and asymmetric power relations. Our chapter also suggests a number of avenues for further research, for example the issue of the environment is identified by Brevini as a major gap in the work of Couldry and Mejias (2019), who advocates for analysis of environmental costs induced by digital and data-related activities to be integrated into the data colonialism framework.

DISCUSSANT RESPONSE

Unpacking the ambiguity of data colonialism – Andrea Jimenez

In this intervention, I will present my critique of the concept of data colonialism, which is around its vagueness and fuzziness. My contention is that the concept's ambiguity has aided its current popularity while also making it difficult to operationalize in current data practices and a further focus on transformational justice. This ambiguity is not just a conceptual problem, it has material implications.

Extensive literature demonstrates how digital and data-driven technologies perpetuate colonial legacies, often within distinct academic fields, including science and technology studies (Harding, 2011), computer science (Philip et al, 2012), philosophy of technology (Wong, 2012; Mohamed et al, 2020), critical data studies (Couldry and Mejias, 2023), communication studies (Casilli, 2017; Madianou, 2019), human–computer interaction (Reynolds-Cuellar, 2021), and AI (Mhlambi, 2020; Birhane, 2021).

This varied body of knowledge has produced concepts describing similar characteristics. Concepts like technocolonialism (Madianou, 2019) and digital colonialism (Kwet, 2019) also advance our understanding of how data-driven digital technologies are being deployed in ways that represent a distinct new stage of colonialism. As well as data colonialism, these concepts place attention on the larger societal power structures, considering who benefits and who suffers from the current applications of data practices (Jasanoff,

2007). For example, in how technology corporations collect data from marginalized populations to profit from them, providing big technological corporations with 'immense political, economic, and social power' (Kwet, 2019: 4). Or how state-led datafication can negatively impact vulnerable communities in post-colonial contexts (Krishna, 2021). Examples also include how humanitarian and international development organizations use digital technologies to improve the efficacy of their services (Madianou, 2019).

The question that arises to me is how such a diverse set of experiences all can be speaking of the aftermaths of colonialism (Quijano and Ennis, 2000; Maldonado-Torres, 2016). Although I understand that the span of time and geographic scope of colonialism has led to a wide range of diverging and overlapping perspectives (Jimenez et al, 2022b, 2022c), I believe it is important to recognize a distinction between 'decolonization' as the process of returning indigenous land, repatriation of resources, and rebalancing of power (Tuck and Yang, 2012); and 'decoloniality', as the work focused on the aftermaths of colonialism, which have resulted in asymmetrical logics, power, and knowledge structures (Maldonado-Torres, 2016). Although both views are interrelated, I see the importance of this distinction because it helps understand Fubara-Manuel's compelling argument in relation to Le Ludec's one. If we adopt Tuck and Yang's prevalent argument that 'decolonization is not a metaphor' as it refers to the repatriation of indigenous land, then we emphasize how data capture pertains to colonized subjects. This specificity implies that the capture and extraction of daily life might not neatly fit into this category. Correctly so. The subjugation and displacement of indigenous people must be viewed as the central legacy of colonialism today.

Nonetheless, we must recognize when data practices and their global supply chains are part of a larger structure of capitalism/modernity/racism/ heteropatriarchy with colonial roots (Grosfoguel, 2011). Colonialism was so extensive and pervasive that it is as much about dehumanization as it is about the oxygen we breathe. Not only is it ubiquitous, but it is also shaping our future, which is why it is also about its current and future environmental costs, as argued by Brevini. And so, in the absence of clear distinction between colonization or coloniality, data colonialism can be vague and fuzzy, a 'catch-all' ideas concept that raises more questions than answers.

The presence of ambiguity has moved our attention away from using the concept in our work. In our research around co-producing the Water Observatory in Lima, Peru (see Hoefsloot et al, 2022; Jimenez et al, 2022a), we noted how the existing data practices and water infrastructures date back to colonial times, when water data and distribution were only directed at more powerful and affluent people (Bell, 2022). We decided to centre the notion of data justice as a way to work towards fairer data practices. This meant that, beyond showing how government data practices were rendering

marginalized people invisible, we actively engaged in creating a platform where those who were not included in the official datasets had the option to be so. While a focus on data colonialism would have centred our discussions on what is colonial about this data practice and/or where coloniality begins and ends, a focus on justice required us to go further and actively attempt to eliminate the injustices we observed.

And so, as data colonialism becomes increasingly popular, a stronger push for clarity and conceptual focus is needed. Otherwise it risks becoming another buzzword in the decolonial bandwagon (Moosavi, 2020).

Notes
[1] Discussant
[2] Facilitator

References

Agre, P. E. (1994) Surveillance and capture: two models of privacy. *Information Society*, 10(2): 101–127. https://doi.org/10.1080/01972243.1994.9960162

Agre, P. E. (1995) From high tech to human tech: empowerment, measurement, and social studies of computing. *Computer Supported Cooperative Work*, 3(2): 167–195. https://doi.org/10.1007/BF00773446

Amoore, L. and Hall, A. (2010) Border theatre: on the arts of security and resistance. *Cultural Geographies*, 17(3): 299–319. https://doi.org/10.1177/1474474010368604

Aneesh, A. (2015) *Neutral Accent: How Language, Labor, and Life Become Global.* Durham, NC: Duke University Press.

Belkhir, L. and Elmeligi, A. (2018) Assessing ICT global emissions footprint: trends to 2040 & recommendations. *Journal of Cleaner Production*, 177: 448–463. doi:10.1016/j.jclepro.2017.12.239

Bell, M. G. (2022) Overlooked legacies: climate vulnerability and risk as incrementally constructed in the municipal drinking water system of Lima, Peru (1578–2017). *Geoforum*, 132: 205–218. https://doi.org/10.1016/j.geoforum.2021.02.016

Birhane, A. (2021) Algorithmic injustice: a relational ethics approach. *Patterns*, 2(2): 100205. https://doi.org/10.1016/j.patter.2021.100205

Bloch, M. (1961) *Feudal Society*. Chicago: University of Chicago Press.

Brevini, B. (2020) Black boxes, not green: mythologizing artificial intelligence and omitting the environment. *Big Data & Society*, 7(2). doi:10.1177/2053951720935141

Brevini, B. (2021) *Is AI Good for the Planet?* Cambridge: Polity.

Brevini, B. and Murdock, G. (2017) *Carbon Capitalism and Communication*. London: Palgrave Macmillan.

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

Cellan-Jones, R. (2020) 'Home office drops "racist" visa decision algorithm', BBC News, 4 August. Available at: www.bbc.com/news/technology-53650758

Césaire, A. (2000) *Discourse on Colonialism*. New York: New York University Press.

Couldry, N. and Mejias, U. A. (2019) *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. 1st edition. Stanford, CA: Stanford University Press.

Couldry, N. and Mejias, U. A. (2023) The decolonial turn in data and technology research: what is at stake and where is it heading? *Information, Communication & Society*, 26(4): 786–802. https://doi.org/10.1080/1369118X.2021.1986102

European Commission (EC) (2022) EURODAC (European Asylum Dactyloscopy Database) | Knowledge for policy, European Commission. Available at: https://knowledge4policy.ec.europa.eu/dataset/ds00008_en

Evans, R. and Gao, J. (2016) 'DeepMind AI reduces Google data centre cooling bill by 40%', DeepMind. Available at: https://deepmind.com/blog/article/deepmind-ai-reduces-google-data-centre-cooling-bill-40

Foucault, M. (1975) *Discipline and Punish*. Paris: Gallimard.

Fubara-Manuel, I. (2019) Animating opacity: race, borders, and biometric surveillance. PhD thesis, University of Sussex.

Fubara-Manuel, I. (2020) Biometric capture: disrupting the digital codification of black migrants in the UK. *African Diaspora*, 12(1,2): 117–141. https://doi.org/10.1163/18725465-01201002

Fuchs, C. (2019) 'Karl Marx in the Age of Big Data Capitalism'. In: C. Fuchs and D. Chandler (eds) *Digital Objects, Digital Subjects: Interdisciplinary Perspectives on Capitalism, Labour and Politics in the Age of Big Data* (pp 53–72). London: University of Westminster Press.

George, A. S., George, A. H., and Martin, A. G. (2023) The environmental impact of AI: a case study of water consumption by Chat GPT. *Partners Universal International Innovation Journal*, 1(2): 97–104.

Grosfoguel, R. (2011) Decolonizing post-colonial studies and paradigms of political-economy: transmodernity, decolonial thinking, and global coloniality. *TRANSMODERNITY: Journal of Peripheral Cultural Production of the Luso-Hispanic World*, 1(1). https://doi.org/10.5070/T411000004

Harding, S. (ed) (2011) *The Postcolonial Science and Technology Studies Reader*. Durham, NC: Duke University Press.

Hoefsloot, F. I., Jimenez, A., Martinez, J., Miranda Sara, L., and Pfeffer, K. (2022) Eliciting design principles using a data justice framework for participatory urban water governance observatories. *Information Technology for Development*, 28(3): 617–638. https://doi.org/10.1080/02681102.2022.2091505

International Energy Agency (2017) 'Digitalisation and Energy'. IEA: Paris. Available at: www.iea.org/reports/digitalisation-and-energy

Jasanoff, S. (2007) Technologies of humility. *Nature*, 450(7166), Article 7166. https://doi.org/10.1038/450033a

Jensen, J. L. (2007) 'The Internet Omnopticon: Surveillance or Counter-insurgency'. In: H. Bang and A. Esmark (eds) *New Publics With/out Democracy* (pp 351–380). Copenhagen: Samfundslitteratur.

Jensen, J. L. (2009) Fra onlinefællesskaber til onlinenetværk: Facebook som augmentering af den sociale virkelighed [From online communities to online networks: Facebook as an augmentation of social reality]. *MedieKultur: Journal of Media and Communication Research*, 25(46).

Jensen, J. L. (2020) *The Medieval Internet*. Bingley: Emerald Publishing.

Jimenez, A., Hoefsloot, F. I., and Miranda Sara, L. (2022a) 'The co-production of the Metropolitan Water Observatory (MWO) platform'. KNOW Working Paper No. 8. Available at: www.urban-know.com/no-8-know-working-paper

Jimenez, A., Vannini, S., and Cox, A. (2022b) A holistic decolonial lens for library and information studies. *Journal of Documentation*, 79(1): 224–244. https://doi.org/10.1108/JD-10-2021-0205

Jimenez, A., Delgado, D., Merino, R., and Argumedo, A. (2022c) A decolonial approach to innovation? Building paths towards buen vivir. *The Journal of Development Studies*, 58(9): 1633–1650. https://doi.org/10.1080/00220388.2022.2043281

Joint Council for the Welfare of Immigrants (2020) 'We won! Home office to stop using racist visa algorithm', Joint Council for the Welfare of Immigrants. Available at: www.jcwi.org.uk/news/we-won-home-office-to-stop-using-racist-visa-algorithm

Jones, N. (2018) 'How to stop data centres from gobbling up the world's electricity', *Nature*, 12 September. Available at: www.nature.com/articles/d41586-018-06610-y

Krishna, S. (2021) Digital identity, datafication and social justice: understanding Aadhaar use among informal workers in south India. *Information Technology for Development*, 27(1): 67–90. https://doi.org/10.1080/02681102.2020.1818544

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

Le Ludec, C. and Cornet, M. (2023) 'How low-paid workers in Madagascar power French tech's AI ambitions'. *The Conversation*, 30 March. Available at: http://theconversation.com/how-low-paid-workers-in-madagascar-power-french-techs-ai-ambitions-202421

Le Ludec, C., Cornet, M., and Casilli, A. A. (2023) The problem with annotation: human labour and outsourcing between France and Madagascar. *Big Data & Society*, 10(2): 20539517231188724. https://doi.org/10.1177/20539517231188723

Madianou, M. (2019) Technocolonialism: digital innovation and data practices in the humanitarian response to refugee crises. *Social Media + Society*, 5(3): 2056305119863146. https://doi.org/10.1177/2056305119863146

Maldonado-Torres, N. (2016) 'Outline of ten theses on coloniality and decoloniality'. Frantz Fanon Foundation. Available at: https://fondation-frantzfanon.com/outline-of-ten-theses-on-coloniality-and-decoloniality/

McDonald, H. (2020) 'Home office to scrap "racist algorithm" for UK visa applicants'. *The Guardian*, 4 August. Available at: www.theguardian.com/uk-news/2020/aug/04/home-office-to-scrap-racist-algorithm-for-uk-visa-applicants

Mhlambi, S. (2020) 'From rationality to relationality: Ubuntu as an ethical and human rights framework for artificial intelligence governance'. Carr Center Discussion Paper Series, 2020–009.

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

Moosavi, L. (2020) The decolonial bandwagon and the dangers of intellectual decolonisation. *International Review of Sociology*, 30(2): 332–354. https://doi.org/10.1080/03906701.2020.1776919

Mosco, V. (2017) 'The Next Internet'. In B. Brevini and G. Murdock (eds) *Carbon Capitalism and Communication* (pp 95–107). London: Palgrave Macmillan.

Naipanoi, L. (2021) 'Dumped e-waste threatens Kenyan lives, contributes to global warming'. The Elephant, 6 November. Available at: www.theelephant.info/features/2021/11/06/dumped-e-waste-threatens-kenyan-lives-contributes-to-global-warming/

Padios, Jan M. (2018) *A Nation on the Line: Call Centers As Postcolonial Predicaments in the Philippines*. Durham and London: Duke University Press.

Patel, I. S. (2021) *We're Here Because You Were There: Immigration and the End of Empire*. Brooklyn, NY: Verso Books.

Pearson, K. (1914) *The Life, Letters and Labours of Francis Galton*. Cambridge: Cambridge University Press.

Philip, K., Irani, L., and Dourish, P. (2012) Postcolonial computing: a tactical survey. *Science, Technology, & Human Values*, 37(1): 3–29. https://doi.org/10.1177/0162243910389594

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

Reynolds-Cuellar, P. (2021) 'Decolonial pathways: our manifesto for a decolonizing agenda in HCI research and design'. MIT Media Lab. Available at: www.media.mit.edu/publications/decolonial-pathways-our-manifesto-for-a-decolonizing-agenda-in-hci-research-and-design/

Scholz, T. (2017) *Uberworked and Underpaid: How Workers Are Disrupting the Digital Economy*. New York: John Wiley & Sons.

Smart, S. (2017) Resistance against mining extractivism in Chile. *Critical Planning*, 23.

Srnicek, N. (2017) *Platform Capitalism*. New York: John Wiley & Sons.

Strubell, E., Ganesh, A., and McCallum, A. (2019) 'Energy and policy considerations for deep learning in NLP'. Cornell University, arXiv:1906.02243.

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

Tuck, E. and Yang, K. W. (2012) Decolonization is not a metaphor. *Decolonization: Indigeneity, Education & Society*, 1(1), Article 1. https://jps.library.utoronto.ca/index.php/des/article/view/18630

van Dijck, J., Poell, T., and de Waal, M. (2018) *The Platform Society*. New York: Oxford University Press.

Wong, P.-H. (2012) Dao, harmony and personhood: towards a Confucian ethics of technology. *Philosophy & Technology*, 25(1): 67–86. https://doi.org/10.1007/s13347-011-0021-z

Zuboff, S. (2018) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Profile Books.