# CHAPTER 4. Data colonialism now: Harms and consequences

### Gabriel Pereira & Nick Couldry

## Summary

* Data colonialism is the latest stage of colonialism: instead of land, it grabs human life in the form of data. This data extraction is radically different in scale and depth from data extraction in the past.
* We get drawn into this data extraction in banal ways on platforms and on our devices, but this is only part of a much larger change in how business relates to human life through extraction.
* This can't be fixed by reforming a few rogue cases, because this is a new landgrab on a truly colonial scale, affecting many sectors: even education, agriculture and health, the fundamental sectors for ensuring the quality of human life, are now dominated by tech corporations that use them to extra data on an industrial scale.

## What is data colonialism?

Colonialism, and the appropriation of resources and knowledge, like all major historical phenomena, is not static: it goes on developing. Its latest form is data colonialism. Whereas historical colonialism grabbed land, the land’s resources and the bodies to mine them, the latest phase of colonialism acquires something new to appropriate and grab: human life, seized through the medium of data. The flow and texture of individual human lives are being seized by corporations and sometimes governments too. They are being seized in the form of data. That data generates value: economic value for corporations and the value that governments get from controlling us more effectively. Either way, a new source of power is being created at human beings’ expense.[[1]](#footnote-1)

It is not that tracking human life to extract value from it is wholly new. Some institutions such as prisons and schools have for centuries involved close surveillance. Some workers have been tracked much more intensively than others, the most extreme case being the constant eye of the plantation owner on enslaved people. Most often, it was through data that marginalized communities have been controlled and surveilled, for example through ‘race classification’ under Apartheid.

What is new today is the scope, scale and depth of how human life is tracked for the benefit of elites. Today’s forms of data extraction are more universal, more fine-grained, and more multi-layered than anything previously in history. And data extraction operates not just at particular moments, but cumulatively: the data taken from us at one moment can be combined with data collected from us and from other people at other times. Our lives are becoming part of a vast grid of continuous comparison and analysis by external institutions, which often remain opaque under the guise of company's Intellectual Property (IP) secrets. This represents a major shift in the power relations in contemporary life: knowledge is power, and the amount of knowledge that external institutions have about us is increasing massively.

We are often alert to power grabs of this sort. But the data power grab is happening in ways that are completely banal, through social relations of data extraction that can seem harmless, until their wider impacts are understood. For example, the ‘cookie’ started this: it was invented in 1994 to make interaction with websites easier. It is a file that gets added to your computer when you visit a web page, a file that enables your computer to be tracked every time it revisits that page, or even visits other pages. But the cookie was just the start of a vast shift in how advertisers operate, from mass marketing to individually targeted marketing, based on continuous surveillance.

## What is data colonialism’s specific importance today?

With the COVID-19 pandemic, the scale and pervasiveness of data colonialism has grown. As the pandemic progressed there was a growing need to monitor how the disease was spreading throughout the world and as people were under lockdown, many more processes moved onto digital platforms. As the scale and pace of digital transformation grew, so too did the reach of data colonialism. Data dispossession is particularly material when also looking at how that same data is being instrumentalised with the AI technologies of digital-era government.[[2]](#footnote-2)

In fact, it may be easy to not perceive data colonialism as a part of our lives. Data and algorithms are most often operating without any direct consent or permission, in the background of our lives. We often agree to this because companies say they just want to make things more seamless, connected, organized. We may also end up trapped within these extractive systems because there are no other options, or even due to the fact they are useful for us. For all of these reasons, it may be really hard to imagine an alternative way beyond the colonial framework that we see around us.

Recognizing the operation of data colonialism allows us to better respond to its harms and consequences. Data colonialism allows us to understand the oppressions of datafication today, including the material impacts data extraction has on people's wellbeing. This includes, for example, how it is used to further entrench racism through datafied systems (as previously discussed by Nai Lee Kalema in Chapter 2). Moreover, it serves as a stepping-stone towards a generative project that considers which future we actually want. A decolonial framework considers how people can have full control over their data, as well as any algorithms that operate from those. It considers not only which kinds of data collection we want, but also those we do *not* want because they are not necessary. What counts as ‘necessary’ (in terms of what data is to be collected and processed, and how) is, after all, strongly shaped by powerful interests: we need to move towards a situation where it is shaped more by actual social deliberation by citizens, that is, by those affected by data’s categorisations. As such, we suggest considering data colonialism as a call for thinking otherwise, for building alternative futures.

## How is data colonialism happening in practice?

Data is not neutral. The way data is collected and constructed defines what it will do in the future, and what gets privileged or not. Extractivism and dispossession undergird how data is located across our everyday life today: a key goal of the dominant system is taking control over people's data in order to profit from it.

The digital platforms we use, like Facebook, literally make no sense except as machines that process and extract value from the data we generate by using them. The same goes for our phones and search engines —except that, in those cases, a public backlash against tracking has led to some recent adjustments. Google has announced that from 2023 it will prohibit third party tracking cookies in its Chrome browser, while Apple has decided to require that apps on its phones must ask for consent to track users via Apple’s mobile advertising ID. But that Apple ID was invented to help those same third parties (and Apple) track us continuously!

But these recent changes will make only a small shift in the wider landscape. The reality is that we are being tracked continuously by websites, apps, and our devices; and countless third parties can track us via those routes. That’s the way the digital economy has come to work over the past three decades. And this is still only the beginning of a much bigger shift in how human beings relate, are governed, and exploited, a much bigger shift in power relations.

We often first hear of these developments through scandals, particularly in areas where governments are involved. We hear that, in the hope of saving money, governments in many countries are employing algorithms to automate their decision-making, or that of their key agencies. Very often, however, those shortcuts don't work well. For example, the commercially developed COMPAS algorithm, designed to influence a criminal court's legal sentence by ‘predicting’ whether the accused is likely to commit a future offense, was found to be no more accurate than a guess made by contributors to the Amazon Mechanical Turk platform.[[3]](#footnote-3) And yet, US judges had been asked to rely on it in deciding their sentences, without having any access to the details of how it worked (or didn't work).

Data colonialism has been particularly effective in changing how social welfare works. In the UK and US, algorithms have been used to 'automate' decisions about the protection of vulnerable children by social services departments, but those algorithms have been found to embed numerous errors which, very often, only compound existing inequalities and injustices. A similar project took place in Argentina, through the collaboration between Microsoft and the Salta Province. The Horus Project, as it was called, monitored poor women and children, using algorithms to supposedly prevent teenage pregnancy and school drop-out.[[4]](#footnote-4) Not only does it continue the longer trajectory of over-surveilling marginalized communities, the project further inscribes goals of prediction that ignore the lived experience of subjects in favor of data and computation.

Meanwhile, in India, a comprehensive biometric data system called Aadhaar is being developed for centralizing welfare, taxes, and all interactions with the State. Its key goal is transforming citizens into machine-readable data, which can be managed and controlled. As explained by Linnet Taylor, ‘the database started as a way to keep track of welfare payments and work programs, but has gradually morphed into a unique public-private configuration’.[[5]](#footnote-5) The consequences of this are manifold: although the project continues to gather data in ways that are unsafe and threaten privacy, it did not in fact help reduce the inefficiency issues it was proposed to solve. Moreover, its inflexible requirement may reinforce inequality for people who, for any reason, have their biometrics not recognized: they may, for example, be denied subsidies and grants and thus further marginalized.

Other examples of data colonialism in our everyday life abound. In education, large EdTech (Education Technology) companies such as Apple, Google, and Pearson are an increasingly significant force in the classroom. They design and run the platforms on which our children study, as well as shape curriculum management, evaluation and guidance of students, and school management. When these companies enter the classroom, not only do they earn by selling their technology, they also seek to gather huge amounts of data on student use of their tools and to get students ‘hooked’ on their proprietary platforms. And what's more: they can also use their newly-captured data for surveilling students. This can go as far as the use of eye-tracking in classrooms to detect ‘mind wandering’ or any student distraction. Not only do such surveillance technologies generate lots of errors, they also exacerbate the tendency of further data grabs, concentrated in the hands of few companies and allowing them a central role in the analysis of education – a crucial element of our society.

A very different sector which faces similar pressures is agriculture. As Alistair Fraser has said, 'the landgrab yields a data grab'.[[6]](#footnote-6) AgTech companies such as John Deere are producing 'smart' farm technologies that rely on the continuous gathering of data about the farming process at previously unimaginable levels of detail, but always under the control of the corporation. Other companies like Monsanto are similarly expanding towards data control, extracting data from sensors which can be sold back to farmers. Not only do these forms of data analysis further the monopoly of data in companies from the Global North, they also support a way of farming which ignores local knowledge and further introduces biodiversity-endangering monocultures. In resistance to such prospects is, for example, the work of the Indian Digital Ecosystem for Agriculture, which seeks ‘to move from a narrow app-centric approach toward an ecosystem model that accounts for a farmer-centric vision of value creation through digitalization’.[[7]](#footnote-7) Here, the goal is decentralization and ecosystem-thinking, rather than the monopolizing of data colonialism by large corporations, including so-called humanitarian organizations, as sponsored by ex-Microsoft CEO Bill Gates.[[8]](#footnote-8)

Meanwhile health services promise perhaps the largest data grab of all, driven by genuine concerns at the need for gathering more data about the spread of dangerous disease and poor health habits. In a clear colonial echo, the Wall Street Journal had called the health sector an 'open frontier' - not so much for scientific knowledge, as for the extraction of data and profit.[[9]](#footnote-9) Although this is now a reality in much of the Global North, with Amazon expanding into healthcare and companies merging to monopolize the field, much of the Global South is slowly creeping in the same direction.

Our cities aren't immune to the land grab of data colonialism. Supposedly ‘smart cities’ are being rolled out across the world, developed by big companies such as IBM and Google. Although they present themselves as cool solutions to real problems, the integration of such digital gadgets furthers surveillance and data capture. For example, smart streetlights deployed in San Diego were supposed to track traffic but were used instead to surveil citizens and share data with law enforcement.[[10]](#footnote-10) Meanwhile, corporations are seeking to fully control a city's data flows, with promises of optimizing them with algorithms. The impacts of this for urban infrastructure are many—including the way they can monopolize the future use of data that was generated. Furthermore, questions regarding infrastructural power abound, for example in the case of Chinese companies providing data infrastructure development in South Africa and Kenya.

Perhaps unsurprisingly, data colonialism is often promoted by, or ends up supporting the goals of, surveillance by law enforcement and the police. Though the police have historically sought to surveil and control, the scale of such surveillance can be increased through continuous data capture. In Rio de Janeiro (Brazil), for example, IBM has been tasked with building an ‘Integrated Command and Control Centre’, from which police and other services surveil and act on the city. However, such data analysis of the city ‘operate[s] without any consent and awareness of the population’, which also furthers future uses which may go against people's privacy.[[11]](#footnote-11) One particularly controversial use of big data gathered by law enforcement is for forecasting future crimes – also known as ‘predictive policing’. Different places across the world, for example India and Rio de Janeiro, have been experimenting with these flawed technologies, which further discriminate against oppressed and over-surveilled communities. Another issue is the use of surveillance on criminalized people, particularly those incarcerated. In Brazil, for example, the use of monitoring through ankle bracelets has significantly grown during the pandemic, signifying ‘a rite of passage to a “virtual prison”, a process that has visible ethical issues with regards to privacy and surveillance that need to be further discussed.[[12]](#footnote-12)

On the borders, data and algorithmic technology are a pressing concern for migrants, communities of people that are often already subject to multiple forms of oppression. The ‘smart border’ is becoming a reality, with Palantir among other companies seeking to datafy migrants and analyze "data from multiple databases run by DHS and law agencies, as well as from data streams linked to people’s internet and social media activity".[[13]](#footnote-13) The consequences of this are the continued arrests and killing of migrants, as well as their imprisonment by ICE and other law enforcement around the world –all while enriching companies with the public's money.

As all these different cases show, once data is collected, it can be used in many different ways —often in ways we don’t expect or don’t want. Data extracted by companies about our everyday lives may be used for their profit, rather than for the reason why we originally ceded the data. Without you even realizing it, the same information may be shared across companies to give you recommendations of new music albums, suggest what you should buy next, or even to increase the price of your health insurance. This is particularly problematic because mass data collection will always support surveillance from those with most power in society. The Police or other agencies may, for instance, use data that was created with a different goal in mind for prosecuting people, a form of violence which disparately impacts those already marginalized and powerless.

At the same time, a lot of what we think about technology comes from the myths perpetuated by the companies who profit from them. For example, companies that want to surveil people will try to frame their technology as positive and efficient. In reality, however, data colonialism is currently a dominant form of operation, one that we must actively resist. There is nothing inevitable about the way things are right now, though they may want us to accept so.

It is tempting to believe that, if only those bad cases can be corrected and principles learned from them, the role of data in decision-making can be improved. But to believe this is to ignore the wider game that is at work. In numerous areas of everyday life, large corporations plan to make money from rearranging them around the extraction and management of data streams, processes that lie firmly under the control of those same corporations, rather than the professionals that previously had the leading expertise in those areas.

1. Mejias and Couldry, Data Grab: the New Colonialism and how to Resist It. [↑](#footnote-ref-1)
2. Killian Clarke, ‘When Do the Dispossessed Protest? Informal Leadership and Mobilization in Syrian Refugee Camps’, Perspectives on Politics 16.3 (2018): 617–33. doi:10.1017/S1537592718001020. [↑](#footnote-ref-2)
3. Julia Dressel and Hany Farid, ‘The accuracy, fairness, and limits of predicting recidivism’, Science advances 4.1 eaao5580 (2018). https://doi.org/10.1126/sciadv.aao5580 [↑](#footnote-ref-3)
4. João Carlos Magalhães and Nick Couldry, ‘Giving by Taking Away: Big Tech, Data Colonialism, and the Reconfiguration of Social Good’, International Journal of Communication 15 (2021). [↑](#footnote-ref-4)
5. Linnet Taylor, ‘Why Today’s Aadhaar Judgement Matters for Data Justice’, Global Data Justice, 26 September 2018. https://globaldatajustice.org/gdj/1859/ [↑](#footnote-ref-5)
6. Alistar Fraser, ‘Land Grab/Data Grab: Precision agriculture and its new horizons’, The Journal of Peasant Studies 46.5 (2019): 893–912. [↑](#footnote-ref-6)
7. Sakhi Shah and Ranjitha Kumar, ‘The Digital Ecosystem Opportunity for Indian Agriculture: Making the Right Choice’, IT for Change, 24 August 2022. https://itforchange.net/node/2196 [↑](#footnote-ref-7)
8. Navdanya International, ‘Gates Ag One: The Recolonisation Of Agriculture’, 16 November 2020. https://navdanyainternational.org/publications/gates-ag-one-the-recolonisation-of-agriculture/ [↑](#footnote-ref-8)
9. Sarah E. Needleman and Rob Copeland, ‘Google’s “Project Nightingale” Triggers Federal Inquiry’, *Wall Street Journal,* 12 November 2019. https://www.wsj.com/articles/behind-googles-project-nightingale-a-health-data-gold-mine-of-50-million-patients-11573571867?reflink=desktopwebshare\_permalink [↑](#footnote-ref-9)
10. Tekla S. Perry, ‘Cops Tap Smart Streetlights Sparking Controversy and Legislation’, *IEEE Spectrum,* 8 August 2020. https://spectrum.ieee.org/cops-smart-street-lights [↑](#footnote-ref-10)
11. Lalita Kraus, Fabiola de Cássia Freitas Neves, and Aldenilson dos Santos Vitorino Costa, ‘Unequal smart spaces: The Command and Control Centre of Rio de Janeiro’, *Espaço e Economia. Revista Brasileira de Geografia Econômica* 23 (2022). https://doi.org/10.4000/espacoeconomia.21619 [↑](#footnote-ref-11)
12. Maria Rita Pereira Xavier, Ana Paula Ferreira Felizardo, and Fábio Wellington Ataíde Alves, ‘Smart Prisoners: Uses of Electronic Monitoring in Brazilian Prisons during the COVID-19 Pandemic’, *Surveillance & Society* 19.2 (2021): 216–227. https://doi.org/10.24908/ss.v19i2.14303 [↑](#footnote-ref-12)
13. Mizue Aizeki, Geoffrey Boyce, Todd Miller, Joseph Nevins, and Miriam Ticktin, ‘Smart Borders or a Humane World? - Immigrant Defense Project*’, Immigrant Defense Project’s Surveillance, Tech & Immigration Policing Project, and the Transnational Institute* (2021). https://www.immigrantdefenseproject.org/smart-borders-or-a-humane-world/ [↑](#footnote-ref-13)