



## **Making data colonialism liveable: how might data's social order be regulated?**

**LSE Research Online URL for this paper:** <http://eprints.lse.ac.uk/101339/>

Version: Published Version

---

### **Article:**

Couldry, Nick and Mejias, Ulises (2019) Making data colonialism liveable: how might data's social order be regulated? *Internet Policy Review*, 8 (2).

10.14763/2019.2.1411

---

### **Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>



# Making data colonialism liveable: how might data's social order be regulated?

**Nick Couldry**

*Media, Communications and Social Theory, London School of Economics & Political Science, United Kingdom*

**Ulises Mejias**

*Communication Studies, State University of New York at Oswego, United States*

Published on 30 Jun 2019 | DOI: 10.14763/2019.2.1411

**Abstract:** Humanity is currently undergoing a large-scale social, economic and legal transformation based on the massive appropriation of social life through data extraction. This quantification of the social represents a new colonial move. While the modes, intensities, scales and contexts of dispossession have changed, the underlying drive of today's data colonialism remains the same: to acquire "territory" and resources from which economic value can be extracted by capital. The injustices embedded in this system need to be made "liveable" through a new legal and regulatory order.

**Keywords:** Data relations, Capitalism, Colonialism, Appropriation

## Article information

**Received:** 01 Feb 2019 **Reviewed:** 28 May 2019 **Published:** 30 Jun 2019

**Licence:** Creative Commons Attribution 3.0 Germany

**Competing interests:** The author has declared that no competing interests exist that have influenced the text.

**URL:**

<http://policyreview.info/articles/analysis/making-data-colonialism-liveable-how-might-datas-social-order-be-regulated>

**Citation:** Couldry, N. & Mejias, U. (2019). Making data colonialism liveable: how might data's social order be regulated?. *Internet Policy Review*, 8(2). DOI: 10.14763/2019.2.1411

*This paper is part of Transnational materialities, a special issue of Internet Policy Review guest-edited by José van Dijck and Bernhard Rieder.*

## INTRODUCTION

A new order is being constructed through the continuous extraction of data from our social lives. This new order, optimised for the creation of economic value, may well become the social order on which the next phase of capitalism depends for its viability. As part of that emerging order,

calls for the regulation of data processing have intensified in the past two years, unsurprisingly perhaps given that capitalism has shown that it needs to be regulated if it is to be made liveable (Polanyi, 2001). But this push for regulation has been framed entirely in terms of taming certain rogue forms of contemporary capitalism. This article argues, however, that to frame data issues solely in terms of a “bad” form of capitalism misses the full scope, scale and nature of what is happening with data. Legal, social and civic responses to what is underway need to be grounded in a broader argument about what we will call “data colonialism”.

There is no doubt of course that what is happening with data today is inextricably linked to the development of capitalism. But is something even larger going on? We argue here that today's quantification of the social—also known as datafication (Mayer-Schönberger and Cukier, 2013; Van Dijck, 2014)—represents the first step in a new form of colonialism. This emerging order has long-term consequences that may be as far-reaching as were the appropriations carried out by historic colonialism for the benefit of the capitalist economies and international legal order that subsequently developed.

Recognising what is happening with data as a *colonial* move means acknowledging the full scope of the resource appropriation under way today through datafication: it is human life itself that is being appropriated so that it can be annexed directly to capital as part of a reconstruction of the very spaces of social experience. In arguing this, we share some common ground with Shoshana Zuboff's well-known argument on “surveillance capitalism”, but there are also crucial differences, which we briefly summarise in three points here (and further unpack later). <sup>1</sup>

First, the transformation of what can be considered an input to capital actually goes well beyond what has been observed in the social media sector to include, for example, the rise of logistics, the new methods of control in the workplace, the emergence of platforms as new structures for profit extraction (for instance, in transportation and tourism), and most generally the reformulation of capitalism's *default business model* around the extraction and management of data (Davenport, 2014). <sup>2</sup> What is going on with data, in other words, is much wider than a problem with a limited number of rogue surveillance capitalists who have gone astray, a problem that can be corrected by their reform. There is only one historic precedent for such a shift in the resources available for economic exploitation, and that is the emergence of colonialism in the late 15th and early 16th centuries. <sup>3</sup>

Second, rethinking data processes on this longer 500-year time-scale allows us to see their implications for capitalism's future in a broader way, too. Here we must recall that industrial capitalism *itself* was only made possible by the profits and socioeconomic reconfigurations that came with historic colonialism.

Third, a colonial framing highlights two central aspects of today's transformations that would otherwise seem like mere collateral: the *subjugation* of human beings that is necessary to a resource appropriation on this scale (relations of subjection to external powers were central to historic colonialism), and the grounding of this entire transformation in a general *rationality* which imposes upon the world a very singular vision of *Big Data's* superior claim on knowledge (just as colonisers justified their appropriation on the ground of the West's superior rationality).

Our argument will consider the long-term historical relations between capitalism and colonialism in the first part of this article, and in the second part offer a discussion—informed by decolonial theory—of Carl Schmitt's classic interpretation of historic colonialism's relation to international law. We hope to give more substance to general calls to recognise the fight against “dataism” (Van Dijck, 2014) as “the most urgent political and economic project” of the 21<sup>st</sup>

century (Harari, 2016, p. 459). This article, written from the intersection of social theory, decolonial theory, and critical data studies rather than policy studies, will hopefully be useful to those who wish to develop a more robust starting-point for critical work on data policy.

## A DECOLONIAL READING OF DATAFICATION

In this first section, we summarise our arguments for analysing contemporary practices of data extraction and data processing as replicating colonial modes of exploitation (see Couldry and Mejias, 2018; Couldry and Mejias, 2019). This will allow us to provide the starting-point for our policy-related discussion later on.

The public is often told that "data is the new oil" (Economist, 2017). A recent article in the *Harvard Business Review* goes further and argues not only that "data is the fuel of the new economy, and even more so of the economy to come," but also that:

Algorithms trained by all these digital traces will be globally transformational. It's possible that new world order will emerge from it, along with a new "GDP" – gross data product – that captures an emerging measure of wealth and power of nations (Chakravorti, Bhalli and Chaturvedi, 2019).

While the evocative idea of "new oil" might recall the benefits (for some) of historic colonialism, it obscures precisely the most important level at which data colonialism must be empirically studied. The most fundamental fact about data is that it is *not* like oil, but rather a social construct operating at a specific moment in history (Gitelman, 2014; Scholz, 2018), driven by much wider economic and social forces. The concept of data colonialism, therefore, highlights the reconfiguration of human life around the maximisation of data *collection* for profit. Without the resulting data flow, there would be no substance related to human life that could, even potentially, be called "oil". The claim that data is like oil is thus an attempt to naturalise the *outcome* of data's collection, and so make data extraction (and the categories it embeds in daily life) part of a social landscape whose contestability is hidden from view (Bowker and Star, 1999). Since regulating data depends, fundamentally, on opening up that contestability, it is essential to understand how the naturalisation of data collection occurs.

To do this, we draw on critical political economy and decolonial theory to trace continuities from colonialism's historic appropriation of territories and natural resources to the datafication of everyday life today. While the modes, intensities, scales and contexts of dispossession have changed, the underlying drive of today's data processes remains the same: to acquire "territory" and resources from which economic value can be extracted. To do so in no way diverts us from an analysis of capitalism. On the contrary, it places datafication squarely within the centuries-long relations between colonialism *and* capitalism, whose separation is now widely contested (Williams, 1994; Beckert and Rockman, 2016). Far from being disconnected from capitalism, the current phase of colonialism (data colonialism) is understood as preparing the way for a new, still undefined stage of capitalism, just as historic colonialism paved the way gradually for industrial capitalism. The medium for this long-term transformation are the interdependencies and rationalities through which social relations, conducted and organised via processes of data extraction, become a normal part of everyday life.

We therefore use the term "colonialism" not as a metaphor, <sup>4</sup> but to name an actual reality. In

this non-metaphorical usage, however, our focus is on colonialism's longer-term historical function: the dispossession of resources and the normalisation of that dispossession so as to generate a new fuel for capitalism's global growth. Distinctive to data colonialism are the subjection of human beings to new types of relations configured around the extraction of data, and, even more broadly, the imposition on human life of a new vision of knowledge and rationality tailored to data extraction (the vision of Big Data). Each generates fundamental questions, in turn, about legal values such as freedom and autonomy, and challenges for existing systems of commercial regulation (we return to those challenges in the next section).

Underlying our argument are two forms of analysis: an analysis of the political economy of the data industry, or what we call the *social quantification sector*; and an analysis of the multimodal forms of exploitation that unfold through our participation in digital platforms and data-processing infrastructures, or what we call *data relations*. These two terms deserve more explanation.

The social quantification sector can be broken down into various sub-groups, starting with the manufacturers of digital devices and personal assistants: well-known media brands such as Amazon, Apple, Microsoft and Samsung, and less well-known makers of devices operating in the fast-expanding 'Internet of Things'. Another group in the social quantification sector includes the builders of the computer-based environments and tools by means of which we connect: household names such as Alibaba, Baidu, Facebook, Google, TenCent and WeChat. Yet another group comprises the growing field of data brokers and data processing organisations such as Acxiom, Equifax, and (in China) TalkingData that collect, aggregate, process, repackage, sell and make decisions based on data of all sorts, while also supporting other organisations in their uses of data. In addition, the social quantification sector also includes the vast domain of organisations that increasingly depend for their basic functions on processing data from social life, whether to customise their services (like Netflix and Spotify), to link sellers and buyers (like Airbnb, Uber, and Didi), or to exploit data in areas of government or security, such as Palantir and Axon (formerly Taser). Finally, analytical consideration of the social impact of the social quantification sector needs to take into account the vast areas of economic life where internal data collection has become normalised as corporations' basic mode of operation, for example in logistics (Cowan, 2014). Corporations such as IBM are key supporters of this wider infrastructure of business data collection (Davenport, 2014), even though they are not associated with either social media platforms or specialised data brokerage.

By data relations we do not mean relations between data, but the new types of human/institutional relations through which data becomes extractable and available for conversion into economic value. When fully established in daily life, data relations will become as naturalised as labour relations, and together comprise a second pillar of the social order on which capitalism is based. <sup>5</sup> This transformation—we propose—goes much further even than the shaping of social relations around the extraction of "surveillance capital" that Zuboff describes. Under data colonialism, human life becomes, as it were, present to capital without obstruction, although this "presence" is based on many levels of technosocial mediation. Data relations give corporations a privileged "window" onto the world of social relations, and a privileged "handle" on the levers of social differentiation. More generally, human life itself, including its relations to technology, becomes a direct input to capital and potentially exploitable for profit. Data relations make the social world *readable to and manageable by* corporations in ways that allow not just the optimisation of profit, but also new models of social governance, what legal scholars Niva Elkin-Koren and Eldar Haber (2016) call "governance by proxy".

In this context, digital spaces for social life and economic transactions called “platforms” (Gillespie, 2010; compare Bucher, 2016; Gerlitz and Helmond, 2013) have significance beyond their convenience for individuals and corporations. *Platforms become software-constructed spaces that produce the social for capital*. Social life is thereby transformed into an open resource for extraction that is somehow “just there” for exploitation. For sure, capitalism has always sought to commodify everything and control all inputs to its production process. But how “everything” is defined at specific historical moments varies. What is unique about this historical moment is that human life is becoming *organised through data relations* so that it can be a direct input to capital. This transformation depends on many things: shifts in daily habits and conventions, software architectures that shape human life through, as Lessig famously argued, “code” (Lessig, 2001), and explicit legal frameworks that legitimate, sanction and regulate such arrangements. In this article, we focus on the last, but including the underlying *legal rationalities* that, as Julie Cohen (2017) argues, work to frame data as ownerless, redefining notions of privacy and property in order to establish a new moral order that justifies the appropriation of data.

To summarise the argument so far: humanity is currently undergoing a large-scale transformation of a social, economic and legal order, based on the massively expanded appropriation by capital of human life itself through the medium of data extraction. The long-term sustainability of this transformation depends, however, on the regulation or harmonising of various factors: the weight of habit and convenience in daily life; various social pressures on consumers, producers and workers towards datafication, which amount to something like a life force (Grewal, 2008); and, crucially, an emerging legal infrastructure. As a result, larger questions arise as to how to regulate this transformation and its emerging institutions. The answers depend on what approach we take to the question of what sort of transformation this is. We have argued, in condensed form, that this transformation can only be fully understood bifocally, that is, through the double lens of capitalism and colonialism. In the second part of the article, we extend this discussion into a brief review of current approaches to regulating personal data processing, and their limitations.

## THINKING BEYOND EXISTING LEGAL APPROACHES TO DATAFICATION

The building of a new social and economic order based on the extraction of value from human life through data relations is not something that individuals can resist, or even manage, by themselves. It matters little whether I delete an app from my phone or withdraw from a platform. Nor, incidentally, can we expect much from the possibility that some players in data markets might act more ethically than others. Society-wide responses are needed to such society-wide transformations. If—to return once again to Polanyi (2001)—large-scale economic change requires a double regulatory movement (first, the transformation of social relations so as to fit the new economic organisation, and then the emergence of a social counter-movement to make the transformation actually *liveable*), then the project of socially managing datafication is likely to be long and complex, and legal reform must play some part in that.

We have little interest here in proposed legal reforms that make partial adjustments to how social media platforms manage aspects of their operations (for example, the algorithms that organise personal news feeds). Our concern instead is with the prospects for large-scale regulation of the extraction of economic value from personal data, and what might currently be

blocking this regulation (by “personal data” we mean not just data which explicitly relates to an individual person, but any data whose collection and processing can generate decisions relating to that person).

There is no doubt that important legal reforms concerning data practices have been advanced recently. Five years ago, North American market rhetoric went largely uncontested, arguing that the wholesale collection and processing of data, whether *about* a person (personal data, in a narrow sense) or otherwise, was essential to the development of the global economy. It is easy to find examples of such discourse, for example, from the World Economic Forum or from business consultants (Letouzé, 2012; World Economic Forum, 2011; McKinsey, 2011). But the balance has been disturbed by one particular legislative intervention, the European General Data Protection Regulation (GDPR), which came into effect in May 2018.

The GDPR's very first sentence announces a normative challenge to market rhetoric about data: “the protection of natural persons in relation to the processing of personal data is a fundamental right” (GDPR, recital 1). Thus, one of the GDPR's basic ideas is that whether or not she is likely to consent to it, the “data subject” must be informed “of the existence of a [data] processing operation” which affects her, and “its purposes.” Indeed, she should be informed of the “consequences of any data profiling” (Recital 60). This challenged the until-then dominant idea that personal data processing is just what corporations and markets do, and has been going on for *so long and on such a scale* that it cannot be challenged (an argument Helen Nissenbaum (2017) calls Big Data exceptionalism). Without going into the GDPR in detail, its importance as a symbolic challenge to the ideology of ‘dataism’ (Van Dijck, 2014) cannot be denied. The GDPR is being used as a model for legislative proposals in a number of countries across the world, including Brazil and the UK, and compliance with the GDPR has become a major feature of recent business practice.

While it is still unclear how effective the GDPR's challenge to data practices from the perspective of human rights such as privacy will be, there is no doubt of the influence its publication has had on the climate of a global debate around data issues. Consider two UN reports from 2014 and 2018, both called “The Right to Privacy in the Digital Age” (UN High Commissioner for Human Rights, 2014, 2018). The 2014 report is almost entirely concerned with state surveillance; when it mentions corporations (paragraphs 42-46), it focuses on whether they should accede to *state* requests for access to their data. The question of whether corporations themselves should be more responsive to human rights concerns regarding how they collect data—arguably the key issues revealed, if not debated, in the 2013 Snowden revelations—is not even mentioned. By 2018 however, the emphasis had shifted to include a discussion of the growth in corporations' data collection practices and their “analytic power” (paragraphs 15 and 16). The later report mentions “a growing global consensus on minimum standards that should govern the processing of personal data by state, business enterprises and other private actors” (paragraph 28), and insists that the resulting human rights protection “should also apply to information derived, inferred, and predicted by automated means, to the extent that the information qualifies as personal data” (paragraph 30). In effect, the 2018 UN report encourages states to adopt something like the GDPR. Yet there are still important gaps in its recommendations: at no point does the report challenge corporate data *collection* as such, or recognise how the continuous collection of data from and about persons might *in itself* undermine values such as freedom and autonomy, even though the report references the fundamental European law principle that the “individual should have an area of autonomous development, interaction and liberty” (para 5), a point to which we shall return.

These legal principles, if pursued, might have the potential to disrupt datafication. But so far it is not legislation but the work of critical legal scholars which has articulated these principles more fully. Scholars of privacy law have often noted that traditional notions of privacy are inadequate to deal with the vast amount of data which flows without being specifically attached to a particular named person, yet, which in combination with even small amounts of other information related to that person can lead to their identification. The result is, as Solon Barocas and Helen Nissenbaum put it in the language of American football, “Big Data’s end run around anonymity and consent” (Barocas and Nissenbaum, 2014). In other words, the scale of data processing that generates decisions affecting the algorithmically produced entities or “data doubles” (Haggerty and Ericson, 2000) to which actual individuals are tethered makes old style privacy regulation by individual consent almost impossible to practice. And yet “consent” is the basic principle on which the GDPR relies.

In response to this problem, Julie Cohen (2013, p. 1931-1932) has proposed an important meta-principle for regulating data practices, that of “semantic discontinuity”. This is designed to limit the possibility of separate data sets being combined so as to generate inferences of a sort that data subjects did not consent to being made. Recently Frischmann and Selinger (2018, p. 275-276) have endorsed this proposal, which radicalises the older principle of “contextual integrity” (Nissenbaum, 2010). But we do not know yet if this proposal has any chance of being translated into law in some form. It runs directly contrary to the purpose of corporate data collection, which is precisely to *combine* data streams without limit, so as to maximise the algorithmic inferences that can be generated from them. How can semantic discontinuity be made effective as a legal principle when it contradicts the stated purposes of countless corporations who seek access to personal data? Would the injunction of the 2018 UN report that “personal data processing should be necessary and proportionate to a *legitimate* purpose that should be *specified by* the processing entity” be sufficient to ground the principle of semantic discontinuity? Presumably not, if a business had a legitimate purpose which depended on semantic *continuity*, and that purpose was in broad terms disclosed to, and consented on by, a data subject. The same question could be asked of non-commercial organisations which might be protected *prima facie* by the “public interest exception” written into the GDPR (Article 21 (6)). On what ground could a “higher” principle of semantic discontinuity override that exception?

What becomes clear here is that a far-reaching challenge to the expanding rationalities of continuous data collection and value extraction runs against the basic organisation of power in contemporary economies and societies, issues which have not yet been broached by even the most enlightened legislation. This potential conflict between critical legal thinking and capitalism’s investment in datafication was anticipated in a remarkable article two decades ago by Paul Schwartz (1999). Schwartz foresaw that the emerging data collection practices made possible by the internet’s new infrastructure of connection would generate “a new structure of power over individuals” with “significant implications for democracy” (1999, p. 815). Schwartz also predicted that individualist liberal notions of autonomy would prove inadequate to counter this development, because they ignore the “constitutive value” (1999, p. 816) that protecting individuals from regular privacy violations and their consequences have for democratic culture itself. Schwartz’s implicitly relational (and post-liberal: Cohen, 2013) understanding of autonomy/freedom connects with more recent accounts of the social costs of datafication and algorithmic decision-making (Eubanks, 2017; Noble, 2018). But the way forward for building effective opposition to the changes under way requires us to move beyond the domain of contemporary legal theory and introduce a decolonial perspective on what is going on with datafication. We turn to this in the next section.



## SCHMITT AND COLONIALISM'S RELATION TO LAW

At this juncture our argument finds support in a surprising source, someone who was certainly not an opponent of historic colonialism: the controversial German legal and political theorist Carl Schmitt. Schmitt (2006 [o.p. 1950]) offered the most clear-sighted account of the relation between law and the appropriation of territory and natural resources within historic colonialism, an account which has implications, we suggest, for grasping the regulatory implications of today's data colonialism. <sup>6</sup> In discussing Schmitt as an exemplary case, we will admittedly be abstracting from the centuries-long debates about the possible legal justifications for the domination by some humans of others. Choosing Schmitt however is justified because of the clarity with which he makes explicit the underlying links between law, force and rationality within historic colonialism.

Schmitt analysed law's relation to historic colonialism, and therefore to the industrial capitalism which colonialism made possible (Schmitt, 2006, p. 4), at a nostalgic moment. Looking back at colonialism, he found it to be an essential underpinning of a eurocentric international legal order which he believed had been shattered by Germany's defeat in World War II. This context does not, however, diminish the importance of Schmitt's remarkably direct portrayal of colonialism and its relation to law.

For Schmitt, controversially, the very idea of law (*nomos*) is based on the seizure of land (2006, p. 42). He interprets the international law of property and nations that dominated the world from the 17<sup>th</sup> to mid-20<sup>th</sup> centuries as emerging from the demise of an earlier order, the "medieval spatial order of the *respublica Christiana*" whose legitimacy was fading by the 16<sup>th</sup> century. According to Schmitt, what enabled a new international legal order to be built was the discovery of "previously unknown (i.e., by Christian sovereigns) oceans, islands, and territories" (2006, p. 131).

Two things are remarkable about the analysis Schmitt develops. First, he makes no pretence that colonial conquests were legal in a conventional sense; rather he distinguishes two types of land-appropriation, those which proceed in accordance with international law, and those (of which historic colonialism was an example) "which uproot an existing spatial order and establish a new *nomos*" of property entitlement (2006, p. 82). In this initially law-less, but ultimately lawful move of historic colonialism, "law and order are one . . . they cannot be separated" (2006, p. 81). Order, that is, *makes* law. Second, Schmitt regards the extra-legal seizure of territory by colonial powers as justified by a higher principle of rationality, or rather a legitimate hierarchy in relation to rationality itself. As he writes (2006, p. 131), "the means of the legal title 'discovery' lay in an appeal to the historically higher point of the discoverer vis-à-vis the discovered." For Schmitt, the conqueror's "scientific cartographic survey *was* a true title to a *terra incognita*," because it embodied a superior rationality, generating a "completely different type of legal title . . . 'effective occupation'" (2006, p. 133).

For Schmitt, the history of colonial appropriation represented the legitimate fusion of effective force (order) into law, justified by a claim to higher knowledge or rationality. Here is Schmitt's fullest statement of the relations between law, force and a certain "modern" reading of rationality: "European discovery of a new world in the 15<sup>th</sup> and 16<sup>th</sup> centuries thus did not occur by chance . . . it was an achievement of newly awakened Occidental rationalism . . . The Indians lacked the scientific power of Christian-European rationality. The intellectual advantage was entirely on the European side, so much so that the New World could simply be 'taken'" (2006:

132). This unapologetic argument for colonialism's *rationality* offers some interesting parallels with the contemporary justification and rationalisation of Big Data practices, parallels that we can only notice within the bifocal approach to capitalism and colonialism that we are proposing. Within this perspective, we also see more clearly the significance of the failure so far of even the boldest legislation on datafication to challenge its basic practice: the banal, almost universal *collection* of personal and non-personal data, and, through this, the creation from the flow of human existence of an *informational terrain* from which extraction for economic value is possible, indeed increasingly seamless. What are the parallels between the legal status of contemporary datafication (understood as a new type of colonial enterprise) and Schmitt's reading of the legal status of historic colonialism?

First, datafication involves a *de facto* appropriation of resources, a domain of connectible information that, through processing, can be attached reliably to entities that are proxies for actual individuals ("data doubles") and thus provide a basis for judgements that effectively discriminate between real individuals. That appropriation depends on the prior collection of data, that is, on the multi-dimensional monitoring of as much of these individuals' online activity as possible, regardless of the device they are using. Granted, there is a legal debate and potential conflict at present (for example via the GDPR) around the legality of some of the consequences of this appropriation, just as there was early on in relation to the Spanish conquests of the "New" World. But, as we saw, these legal debates tend *never* to challenge the fundamental fact of continuous monitoring itself, even if it is in tension with established values such as autonomy (for example the "right to full development of the personality" under German constitutional law: Hornung and Schnabel, 2009).

Second, although it is as yet only in the early stages of development, a justificatory ideology of data appropriation is emerging that parallels Schmitt's version of colonial ideology: the vision that only through the superior calculating power of Big Data and machine learning can a higher state of human knowledge be achieved, thereby *justifying* corporate access to data that can be extracted from the flow of individuals' daily lives. The core issue here is the imposing on the whole domain of human life a very specific version of rationality which requires all life to be tracked continuously in the interests, simultaneously, of capital and of a certain version of human knowledge (the vision of Big Data or dataism).

It follows, thirdly, —and here we move from parallels onto implications— that the more fundamental challenge to processes of datafication to which critical legal scholars such as Cohen and Frischmann are committed requires a challenge to the underlying legitimacy of acquiring data through data relations, which is today a feature of most platforms, apps, and mechanisms for knowledge production and daily organisation (think of the Internet of Things). Cohen's principle of "semantic discontinuity" is important, but only goes so far as challenging the *transferability* of data, when it is the very act of collecting data that must above all be challenged.

There are indeed good reasons (which Cohen in her work has noted) for arguing that the continuous collection of data from and about individuals conflicts with the principle of autonomy on which democracies, fundamentally, rely. Continuous surveillance or monitoring by the state is, after all, generally regarded as "chilling" of individual agency (Cohen, 2013, p. 1911-1912). The same is true of surveillance when it is conducted by private corporations, particularly if those corporations often have both capacity and need to yield up data to the state. What so far has been difficult to assert is the primacy of these concerns against the opposing rationality of the social quantification sector, which relies on its "effective occupation" of human life (to use

Schmitt's chilling phrase) as the starting-point for defending its practices of data collection against interference by the state. *What is needed is to reject precisely this act of "effective occupation"*. What cuts through all the rationalities which mask the dynamics of datafication is precisely the realisation that the social quantification sector's "right" to hold what they gather is no more legally justifiable than (and just as legally contentious as) the effective occupation of overseas territory by colonial states once was.

If so, the existence (or not) of "consent" to continuous monitoring is beside the point. What matters are the implications of this occupation for what we call *the space of the self*, that is, the basic idea of selfhood on which most notions of democracy and even legal authority rely. <sup>7</sup> We are drawing here on a relational notion of freedom which assumes that "individual" freedom can only emerge through a web of social relations (Elias, 1978), but also more specifically on the idea that, underlying all notions of freedom and autonomy (some of which no doubt are today unsatisfactory) and underlying also all culturally relative formulations of personal privacy, is a basic notion of the "space of the self": that is, "the socially grounded integrity without which we cannot recognize ourselves or others as selves at all" (Couldry and Mejias, 2019, p. 155). This is the space that Hegel captured in his relational definition of freedom as "the freedom to be *with oneself* in the other" and that Dussel terms the "natural substantivity of the person." <sup>8</sup>

Our approach to reframing legal challenges to datafication is, we acknowledge, expansive. It cuts across the detailed debates of policy and law in particular contexts. But it usefully sidesteps the confusion caused by the anomalous notion of "personal data". As many critics of traditional notions of privacy have noted, much of the data that makes a difference to how we are treated by corporations is not *personal* data, because it is not exactly "about" us. Rather, it is *relational* data, in which patterns emerge across myriad comparisons within much larger data sets, patterns that predict particular outcomes for a data double to which as a real individual each of us is tethered. The protection of "personal data" in a more straightforward sense—data about individuals and data files such as photos that an individual claims to own—is therefore likely only to protect people from part of the harms that can be done to them through data. Our approach challenges the very validity of continuous data collection, regardless of what entities happen to be affected by any one particular decision or practice. It challenges, in other words, the multiple practices which construct the new "territory" of human life from which something like "personal data" emerges as potentially extractable, a territory which is steadily supplanting the space of social interaction and social governance that was taken for granted before datafication through a process that started centuries before the advent of digital data. In other words, it makes this challenge in response to processes of human subjection that only a colonial perspective can fully recognise.

There is one last and crucial respect in which legal and civic challenges to datafication require the frame of colonialism. This regards the underlying rationality of Big Data itself which works as a reference-point for and legitimation of data collection in all its breadth and depth. Underlying all the specific and important issues under discussion about algorithmic injustice lies a deeper injustice that, following decolonial thinker Boaventura de Sousa Santos (2014), we can call "cognitive injustice". Put simply, this is the assumption that there is *only one path* to human knowledge and that it lies through the progressive extraction, collection, processing and evaluation of data from the flow of human life, and indeed life more generally. <sup>9</sup> The characteristics of this rationality have been expressed not by an analyst of capitalism or even modernity, but a decolonial thinker, the Peruvian sociologist Aníbal

Quijano, reflecting on the relations between capitalism, modernity and the longer process of not

just historic colonialism but *coloniality*:

Outside the 'West', virtually in all known cultures... all systematic production of knowledge is associated with a perspective of totality. But in those cultures, the perspective of totality in knowledge includes *the acknowledgement of the heterogeneity of all reality*; of the irreducible, contradictory character of the latter; of the legitimacy, i.e., the desirability of the *diverse* character of the components of all reality — *and therefore, of the social*. The [better, alternative] idea of social totality, then, not only does not deny, but depends on the historical diversity and heterogeneity of society, of every society. In other words, it not only does not deny, but it *requires* the idea of an 'other' — diverse, different. (Quijano, 2007, p. 177, added emphasis).

Through the quantification of the social, we risk installing a new version of this exclusive notion of rationality, via what Jose van Dijck (2014) has called “dataism”. Only legal proposals which challenge rationales of data collection in this more fundamental way can hope, effectively, to challenge the direction of data colonialism.

Our approach therefore stands firmly against other recent proposals for individuals to own “their” data, be free to manage access to it, and perhaps even to be paid in return for such access (Lanier, 2013; Arrieta-Ibarra et al, 2018; for a recent popular argument in *The Economist*, see will.i.am (2019)). Such proposals risk legitimating precisely the underlying practices of data collection, and ignoring completely the rationality of appropriation which underlies data colonialism.

## CONCLUSION

Our goal in this article has been to develop the starting-points of a more radical and potentially more comprehensive approach to framing critical legal and policy responses to ongoing processes of datafication. We began by reframing what is currently going on with data not just within the continuing expansion of capitalism, but as a new and epochal renewal of colonialism itself, which, in time, may pave the way for a stage of capitalism whose full outline we cannot yet predict.

By placing datafication within the longer history of colonial appropriations of territory and natural resources on a global scale, we seek to address more effectively the fundamental unease across wide sectors of the population at today's practices of expanding surveillance via marketing, artificial intelligence and the Internet of Things. Existing legal approaches, and even critical legal theory, fall short of providing an adequate starting-point for wider critique. So too do accounts of capitalism which frame what is going on with data principally in terms of recent developments (surveillance capitalism, platform capitalism, and the like), rather than the longer term relations between colonialism and capitalism.

By contrast, legal approaches which take seriously Carl Schmitt's reading of the role of historic colonialism in making law through effective force (that is, what becomes an *order*) offer a warning of the underlying direction of change. Unless we grasp this, policy debate regarding the challenges of datafication is always likely to fall short of the mark.

A postscript: One day after being fined for privacy violations, Google announced that “data is more like sunlight than oil” (Ghosh and Kanter, 2019). In other words, instead of a resource that is being appropriated from someone’s territory, Google would like us to believe that data is a replenishable, inexhaustible, owner-less resource that can be harvested sustainably for the benefit of humanity. This illusion, once again, conveniently bypasses the questions about privacy and protecting the individual that any attempt at “regulation” would normally want to raise. Instead, this “regulation” attempts to establish data colonialism as the *status quo*. It is time for a more radical grounding of established regulatory discourse that enables it to challenge datafication’s social order. This must involve more than regulatory adjustments to certain aspects of contemporary capitalism. What is required is a fundamental challenge to the direction and rationale of capitalism as a whole in the emerging era of data colonialism.

## REFERENCES

- Arrieta-Ibarra, I., Goff, L., Hernandez, D., Lanier, J., & Weyl, G. (2018). Should We Treat Data as Labor? Moving Beyond “Free”. *AEA Papers and Proceedings*, 108, 38–42.  
doi:10.1257/pandp.20181003
- Barocas, S. & Nissenbaum, H. (2014). Big Data's End Run Around Anonymity and Consent (pp. 44-75). In J. Lane, V. Stodden, S. Bendo, & H. Nissenbaum (Eds.), *Privacy, Big Data and the Public Good*. New York: Cambridge University Press.
- Beckert, S. & Rockman, S. (Eds). (2016). *Slavery's Capitalism*. Philadelphia: University of Pennsylvania Press.
- Bowker, G. & Leigh Star, S. (1999). *Sorting Things Out*. Cambridge, MA: The MIT Press.
- Bratton, B. (2016). *The Stack: On Software and Sovereignty*. Cambridge, MA: The MIT Press.
- Bucher, T. (2017). The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms. *Information Communication and Society*, 20(1),30–44.  
doi:10.1080/1369118X.2016.1154086
- Chakravorti, B., Bhalli, A., & Chaturvedi, R. S. (2019, January 24). Which Countries are Leading the Data Economy? *Harvard Business Review*. Retrieved from <https://hbr.org/2019/01/which-countries-are-leading-the-data-economy>
- Cohen, J. (2013). What Privacy Is for. *Harvard Law Review*, 126(7), 1904–1933. Retrieved from <https://harvardlawreview.org/2013/05/what-privacy-is-for/>
- Cohen, J. (2018). The Biopolitical Public Domain: The Legal Construction of the Surveillance Economy. *Philosophy & Technology*, 31(2), 213–233. doi:10.1007/s13347-017-0258-2
- Cohen, J. (2019). *Between Truth and Power*. Oxford: Oxford University Press.
- Couldry, N., & Mejjias, U. A. (2018). Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. *Television & New Media*, 20(4): 336-349.  
doi:10.1177/1527476418796632
- Couldry, N., & Mejjias, U. A. (2019). *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capitalism*. Redwood City, CA: Stanford University Press.
- Cowen, D. (2014). *The Deadly Life of Logistics*. Minneapolis: University of Minnesota Press.
- Davenport, T. (2014). *Big Data @ Work*. Cambridge, MA: Harvard Business Review Press.
- Dussel, E. (1985). *Philosophy of Liberation*. Oregon: Wipf and Stock.
- The Economist. (2017, May 6). The World's Most Valuable Resource Is No Longer Oil, but Data. Retrieved from <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
- Elias, N. (1978). *What is Sociology?* London: Hutchinson.
- Elkin-Koren, N., & Haber, E. (2016). Governance by Proxy: Cyber Challenges to Civil Liberties. *Brooklyn Law Review*, 82(1), 105–162. Retrieved from

<https://brooklynworks.brooklaw.edu/blr/vol82/iss1/3/>

Eubanks, V. (2018). *Automating Inequality*. New York: St. Martin's Press.

Frischmann, B. & Selinger, E. (2018). *Reengineering Humanity*. Cambridge: Cambridge University Press.

Gerlitz, C., & Helmond, A. (2013). The Like Economy: Social Buttons and the Data-intensive Web. *New Media & Society* 15 (8), 1348-1365. doi:10.1177/1461444812472322

Gillespie, T. (2010). The Politics of 'Platforms'. *New Media & Society* 12(3): 347-364. doi:10.1177/1461444809342738

Gitelman, L. (Ed). (2013). *"Raw Data" is an Oxymoron*. Cambridge, MA: The MIT Press.

Gosh, S., & Kanter, J. (2019, January 22). Google says data is more like sunlight than oil, one day after being fined \$57 million over its privacy and consent practices. *Business Insider*. Retrieved February 1, 2019, from <https://www.businessinsider.com/google-data-is-more-like-sunlight-than-oil-france-gdpr-fine-57-million-2019-1>.

Grewal, D. (2008). *Network Power*. New Haven, CT: Yale University Press.

Haggerty, K., & Ericson, R. (2000). The Surveillant Assemblage. *British Journal of Sociology*, 51(4): 605–622. doi:10.1080/00071310020015280

Hegel, G. W. F. (1991). *Elements of the philosophy of right* (A. W. Wood, Ed.; H. B. Nisbet, Trans.). Cambridge: Cambridge Univ. Press.

Hildebrandt, M. (2015). *Smart Technologies and the End(s) of Law*. Cheltenham: Edward Elgar Publishing.

Hornung, G., & Schnabel, C. (2009). Data Protection in Germany I: The Population Census Decision and The Right to Informational Self-determination. *Computer Law & Security Review*, 25(1): 84–88. doi:10.1016/j.clsr.2008.11.002

Lanier, J. (2014) *Who Owns the Future?* London: Allen Lane.

Lessig, L. (2000). *Code and Other Laws of Cyberspace*. New York: Basic Books.

Letouzé, E. (2012). *Big Data for Development: Challenges & Opportunities* [Report]. New York: UN Global Pulse. Retrieved from <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf>

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data*. London: John Murray.

McKinsey (2011). *Big Data: The next frontier for innovation, competition, and productivity* [Report]. McKinsey Global Institute.

Nissenbaum, H. (2010). *Privacy in Context*. Stanford, CA: Stanford University Press.

Nissenbaum, H. (2017). Deregulating Collection: Must Privacy Give Way to Use Regulation? doi:10.2139/ssrn.3092282

- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.
- Pippin, R. (2008) *Hegel's Practical Philosophy*. Cambridge: Cambridge University Press.
- Polanyi, K. (2001). *The Great Transformation*. Boston: Beacon Press.
- Postone, M. (1998). Rethinking Marx (In a Post-Marxist World) (pp. 45-80). In C. Camic (Ed.), *Reclaiming the Sociological Classics*. Oxford: Wiley-Blackwell.
- Quijano, A. (2007). Coloniality and Modernity/Rationality. *Cultural Studies* 21(2-3): 168-178. doi:10.1080/09502380601164353
- Santos, B. de S. (2016). *Epistemologies of the South: Justice Against Epistemicide*. London: Routledge. doi:10.4324/9781315634876
- Schmitt, C. (2006) *The Nomos of the Earth*. Candor, NY: Telos Press.
- Scholz, L. (2018). *Big Data is not Big Oil: The Role of Analogy in the Law of New Technologies* [Research paper No. 895]. Tallahassee, FL: FSU College of Law. Retrieved from [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3252543](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3252543)
- Schwartz, P. (1999). Internet Privacy and the State. *Connecticut Law Review*, 32, 815-859. Available at <https://scholarship.law.berkeley.edu/facpubs/766/>
- Sen, A. (2002). *Rationality and Freedom*. Cambridge, MA: Harvard University Press.
- Shepherd, T. (2015). Mapped, Measured and Mined: The Social Graph and Colonial Visuality. *Social Media + Society*, 1(1). doi:10.1177/2056305115578671
- Thatcher, J., O'Sullivan, D. & Mahmoudi, D. (2017). Data Colonialism Through Accumulation by Dispossession: New Metaphors for Daily Data. *Environment and Planning D: Society and Space*, 34 (6), 990-1006. doi:10.1177/0263775816633195
- UN High Commissioner for Human Rights. (2014). *The Right to Privacy in the Digital Age*. Retrieved from <http://www.justsecurity.org/wp-content/uploads/2014/07/HRC-Right-to-Privacy-Report.pdf>
- UN High Commissioner for Human Rights. (2018). *The Right to Privacy in the Digital Age*. Retrieved from <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/239/58/PDF/G1823958.pdf>
- Van Dijck, J. (2014). Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology. *Surveillance & Society*, 12(2), 197-208. doi:10.24908/ss.v12i2.4776
- will.i.am. (2019, January 21). We Need to Own Data as a human right – and be compensated for it. *The Economist*. Retrieved from <https://www.economist.com/open-future/2019/01/21/we-need-to-own-our-data-as-a-human-right-and-be-compensated-for-it>
- Williams, E. (1994). *Capitalism and Slavery*. Chapel Hill: University of North Carolina Press.
- World Economic Forum. (2011). *Personal Data: The Emergence of a New Asset Class*. Retrieved from



[http://www3.weforum.org/docs/WEF\\_ITTC\\_PersonalDataNewAsset\\_Report\\_2011.pdf](http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf).

Zuboff, S. (2015). Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology*, 30(1),75–89. doi:10.1057/jit.2015.5

Zuboff, S. (2019) *The Age of Surveillance Capitalism*. London, UK: Profile Books.

## FOOTNOTES

1. The full context of our argument is provided in Couldry and Mejias (2019). We have developed it since 2016, and first presented it publicly at the Big Data in the Global South network at IAMCR, Cartagena, Colombia, in July 2017 (<https://data-activism.net/2017/07/dataactive-presents-big-data-from-the-south-in-cartagena-july-15/>). For a summary version of our book's argument, see Couldry and Mejias (2018).

2. We therefore question the boundary between “capitalism” and “surveillance capitalism” (sometimes called “raw surveillance capitalism”) on which Zuboff relies, when she writes: “When a firm collects behavioral data with permission solely as a means to product or service improvement, it is committing capitalism but not surveillance capitalism” (2019, p. 22). But this assumes a world where “permission” is clearly delineated, and the purposes of data use and scope of data collection are neatly delineated too: the purpose of data colonialism is to blur those boundaries in the service of a broader appropriation of human life itself.

3. Interestingly Zuboff notes the colonial precedent at certain points (e.g., Chapter 6), but without either theorising data processes as a new type of colonialism, or explaining the implications of the colonial precedent for her framing of what's going on with data exclusively in terms of capitalism.

4. For recent valuable discussions of the colonial in relation to data, Thatcher et al. (2017) see ‘data colonialism’ explicitly as a metaphor, while Cohen (2017) and Shepherd (2015) emphasise neo-colonial continuities in data practices. None proposes, as we do, that data practices constitute literally a new phase of colonialism.

5. This unorthodox extension of Marx's critical theory of capitalism is inspired by Moishe Postone's reading of Marx and the importance of abstraction, rather than labour as such, as the fundamental driver of creating a capitalist social order (Postone, 1998). There is no space to discuss this in detail here, but see Couldry and Mejias, 2018; Couldry and Mejias, 2019, chapter 1.

6. For an earlier discussion of Schmitt's discussion of law and colonialism in relation to the internet, see Bratton (2016: 19-40).

7. Our larger argument here draws on the philosophy of G. W. F. Hegel and Enrique Dussel: for more detail, see Couldry and Mejias (2019, chapter 5). On the question of legal authority, see Hildebrandt (2015).

8. See Hegel's *Encyclopedia* quoted by Pippin (2008, p. 186); Dussel (1985, p.158).

9. On data extraction from physical nature, see Gabryz (2016).