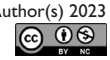
Critical Sociology

Article

On Digital Fetishism: A Critique of the Big Data Paradigm

Critical Sociology
1–14



DOI: 10.1177/08969205231202873
journals.sagepub.com/home/crs


Andrea Miconi
IULM University, Italy

Abstract
The article takes into exam the current literature about Big Data and data capitalism, from the perspective of the critical Internet theory. Particular attention will be placed to the ideas of data exploitation and raw data, which will prove to betray a form of digital fetishism: in short, the focus on the final results of the production process, rather than on the social relations by which the very same process is fueled.

Keywords
media sociology, Big Data, data capitalism, fetishism, Karl Marx, critical Internet theory, platform economy

Introduction: What Platforms Say

In his study on six major players in platform economy—Google, Apple, Microsoft, Facebook, Amazon, and Netflix—Nicholas Petit showcases a quantitative content analysis of 10,000 reports released by each of these companies (what he calls 'the 10-Ks'). All in all, Petit's goal is to figure out whether the tech giants are in a monopolistic position over the market, and the research questions are thereby derived. Of particular interest, thus, is the frequency of the references to the competition with other firms, or the way the so-called MAGNAF assess their own condition—for instance, how often they state to be threatened to be taken over by new start-ups, or by their commercial antagonists (Petit, 2020: 37–60). Petit's (2020) answer, at that, is not straightforward, as no real trace of industrial trust would be present, at least from this point of observation: on the opposite, the '10-Ks are filled with claims of significant risks of competitive entry' (p. 123). Rather than a as monopoly in the classical sense, the explanation goes, the nature of these platforms should be thought of as a unique, and unprecedented combination of different factors. First, it is the 'infrastructural monopoly' of the hubs occupying the strategic nodes, and also equipping the other

Corresponding author:
Andrea Miconi, Arts and Media Department, IULM University, via Carlo Bo 1, 20143 Milan, Italy.
Email: andrea.miconi@iulm.it

players with the basic tools and resources for thriving online: an idea which replicates the features of the 'infrastructural platforms', as opposed to the sectorial ones, as laid out by Van Dijck et al. (2018: 12–22). Second, it has to be considered the 'cognitive oligopoly' of the major commercial services, due to the sedimentation of the cultural and social capital accumulated over the years.[1] By taking together these two aspects, Petit creates the neologism 'moligopoly', named after the crasis of the two original and more common words (Petit, 2020: 33–35). What is finally underestimated, according to Petit, is the platforms' actual exposition to the risk of being put out of business, which would be caused by the proverbial dynamism of the high-tech compart (Petit, 2020: 72–91). This last leg of the reasoning, from our standpoint, is possibly the less convincing. In short, Petit draws on Schumpeter's notion of destructive creation, fueled by the never-ending competition among investors and venture capitalists. It makes sense to object, though, that this state of competitiveness and rivalry is specific to the first stage of technological revolutions, as described by Tim Wu in his analysis of the cyclical evolution of media systems. 'It is here that the Cycle begins', Wu (2010) observes, 'in a lonely room where one or two men are trying to solve a concrete problem. So many revolutionary innovations start small, with outsiders, amateurs, and idealists in attic or garage' (p. 18). But this is only half of the story:

> Each of these inventions, in time, passed through a phase of revolutionary novelty and youthful utopianism; each would change our lives. For whatever social transformation any of them might have effected, each would take its place to uphold the social structure that has been with us since the Industrial Revolution. Each became, that is, a highly centralized and integrated new industry. (Wu, 2010: 6)

In some way, Schumpeter himself allows for this cyclical interpretation, as it is 'the possibility of retention, at least temporarily, of above-normal profits from innovations', that 'may well stimulate a higher rate of innovation and technological improvement': to the point that in some periods, 'the monopoly position of the innovator [. . .] may be the price society pays for technological progress and thereby a higher growth rate over the long run' (Schumpeter, 1934: xxxvii). It is of theoretical importance, additionally, that Schumpeter (1934) ascribes the 'monopoly revenue' to the modality of the 'rent' (p. 168)—important, as his position is here close to Marx's (1863–1883) reading of the *rent* as the resurgence of a pre-capitalist formula, conflicting with more advanced exploitation methods, and namely with waged labor (pp. 557–562). Even though the Marxist perspective is the farthest to Petit's, moreover, the market tendency toward the monopoly should not to be intended as a static state-of-being: it being a *tendency*, by definition, it produces variable outputs and different 'digital monopolies', through a 'dynamic process' affected by several factors, including policies and regulatory measures (Bilić et al., 2021: 60). In a more general stance, we can say that it is the very consolidation of global platforms, which acts as a proof of the concentration process aptly described by Wu. This notwithstanding, on the margins of Petit's investigation evidence emerges, which leads us back to the contents of the 10-Ks:

> Observe that we do not discuss here the tech giants' diversification through the prism of «data». The reason is that big tech firms do not specifically discuss «data» in 10-Ks. [. . .] Moreover, in our current state of knowledge, we lack a theory of value or of externalities that considers 'data' as an explanatory variable for economic profits amenable to estimation. (Petit, 2020: 121)

As accidental as it may be, this is a finding on which our discourse can be usefully premised: if it is all about data, why there is no trace of that in the documents released by the most dominant players?

The Myth of Data Capitalism

That the big data has reached its hype, can easily be measured with such simple tools as the Google Ngram Viewer. Besides all quantitative implications of that, though, the aim of the article is to reflect on our common understanding of this phenomenon, and on its theoretical implications. In particular, we will focus on how the topic has been elaborated on in three relevant, albeit different approaches: Nick Couldry and Ulisses Mejias' idea of data colonialism; Viktor Mayer-Schönberger's contributions on the big data paradigm; and Shoshana Zuboff's interpretation of the categories of 'big other' and digital surveillance. We will not indulge in a detailed review of these books, while taking into exam a facet that is common to all of them, which is the over-emphasizing of the role of big data—or, to paraphrase Ian van Dijk's critique to Manuel Castells' classical theory of the network society (Van Dijk, 1999: 220–223), their contribution to the 'reification' of the data. While the importance of the data in digital economy could hardly be exaggerated, in general terms (for an updated bibliographical account, see Bolin, 2022), here we will single out the aspects that the big data category cannot explain—and what this category is possibly hiding from our view.

Our historical moment is 'distinctively colonial', Couldry and Mejias (2019) argue, as the unearthing of a new kind of 'raw material', the data, is opening the way to a new form of economic exploitation—exactly as the discovery and the predation of the overseas colonies have prepared the ground for the industrial revolution, and therefore for modern capitalism (p. 5). The definition of the data as 'raw material' is actually nothing new, per se: it had been already advanced by Nick Srnicek (2017), among the others, and adopted by Shoshana Zuboff (2019: 65) as well. 'In the twenty-first century', Srnicek (2017) writes, 'advanced capitalism came to be centered upon extracting and using a particular kind of raw material: data' (p. 28). As to Couldry and Mejias (2019), and as repeatedly stated in their book, the authors are not suggesting a simple analogy between two different eras: modern 'colonialism and data colonialism' are in *historical and material continuity*, as 'they share some fundamental structures that ground the resource appropriation and the social relations of each' (p. 85; see also pp. 188–190).

Whether we take it as an analogy or as a historical recourse, the category of data colonialism comes short in explaining an aspect of pivotal importance, at least in the perspective of critical theory. In actuality, the data cannot be equated to raw materials, and their well-rounded objectivity should not divert our attention from the fact that they are themselves the result of a socially bound process of production. As Marx wrote in the *Fragment on Machines*, whose rediscovery in the 1980s paved the way to the theorization of the general intellect,

> the labour process was simply included within capital, and, as regards its physical conditions, its material presence, capital appeared as the totality of the conditions of this process, and correspondingly sorted itself out into certain qualitatively different parts, *material of labour* (this, not raw material, is the correct expression of the concept), *means of labour* and *living labour*. (Marx, 1857–1861)

To put it in a simpler way, data do not exist as a given entity, unlike the natural resources plundered by the colonial empires—so that the very expression 'raw data is an oxymoron', to quote Lisa Gitelman (2013), or 'the raw data' are actually 'cooked', as it emerges from a study on the 'social construction of the data' in the military sector (Räsänen and Nyce, 2013). If we look at the historical semantics of the concept, in fact, the term 'likely dates informally to the 1990s' (Boellstorff, 2013: 2), and it is told to have appeared in academic discourse in 2003 (Lohr, 2013). The big data locution originates in the language of business administration,[2] and over the years it has taken on different meanings: among which, for what is of our interest, exactly those of 'natural force' or 'natural resource', as it will be in Couldry and Mejias (Puschmann and Burgess, 2014). Upon this historically contingent metaphor,

> the givenness of data is analogized through the givenness of natural resources, which can be mined or grown and which can act as a form of capital with no persistent ties to their creator. Data is not a natural resource that replenishes itself, but in social media platforms it is created by users with intentions entirely unrelated to its use as a valued commodity. (Puschmann and Burgess, 2014: 1699)

We may observe that Couldry and Mejias (2019), despite the stand taken against the ideology of 'dataism' (p. 17), somehow contribute to the already cited reification of the data, by attributing them the status of a raw material—which, by definition, would pre-exist the stage of social division of labor. In other words, we would suggest that platforms do not only capitalize on the data, more importantly, they 'take advantage of the time the users spend' online consuming and producing data (Bilić, 2016: 1). We will get back to this point in the final section of the article.

Mayer-Schönberger's work on big data starts with the acknowledging its revolutionary impact. We need to get used to a new epistemological paradigm, Mayer-Schönberger and Cukier (2013) argue, which will replace the 'why' with the 'what': the statistical associations made visible by digital archives 'may not tell us precisely *why* something is happening', in fact, 'but they alert us *that* it is happening' (p. 14). It follows their call for a cognitive adjustment in favor of 'correlation', rather than explanation (pp. 50–72): an idea that is quite popular among data scientists (see Barabási, 2010; Rouvroy, 2012), and has been notoriously vulgarized by Chris Anderson (2008) into announcing the death of the theory. By and large, Mayer-Schönberger's stance is that we should get rid of our classical concepts, as 'eventually most value will be in the data itself' (Mayer-Schönberger and Cukier, 2013: 134). Mayer-Schönberger and Ramge's more recent book stresses this point to the limits, when hypothesizing the transition from 'money-rich' to 'data-rich' markets: with information quanta (allegedly) becoming more important than economic transactions, we 'may no longer label the future 'capitalist' in the sense of power concentrated by the holders of money', therefore 'prov[ing] Karl Marx wrong' (Mayer-Schönberger and Ramge, 2018: 138–142). As to Marx himself, it is pretty clear that the authors are mistaking *value*—the exploitation of plus-labor, made possible by the split between the human work and the control of the material means necessary to it—with 'money', which comes last, as its very final objectification. Moreover, when we consider the money as the terminal abstraction of social activities—following the commodification of labor, the appropriation of value, and its transformation—we may reach very different conclusions. According to Slavoj Žižek, we should even 'reformulate the standard Marxist topic' of reification, and for the very opposite reason: as it 'relies on a notion of the fetish as a solid object whose stable presence obfuscates its social mediation'.

> Paradoxically, fetishism reaches its acme precisely when the fetish itself is 'dematerialized', turned into a fluid 'immaterial' virtual entity; money fetishism will culminate with the passage to its electronic form when the last traces of its materiality will disappear. Electronic money is the third form, after 'real' money which directly embodies its value (gold, silver) and paper money which, though a 'mere sign' with no intrinsic value, still clings to its material existence. And it is only at this stage, when money becomes a purely virtual point of reference, that it finally assumes the form of an indestructible spectral presence: I owe you $1000, and no matter how many material notes I burn, I still owe you $1000, the debt is inscribed somewhere in virtual digital space. (Žižek, 2012: 246)

We do not agree on this 'dematerialized' stage necessarily evading Marx's categories, but still, Žižek addresses a very relevant point: the replacement of material money with electronic payments and exchanges is itself a fetishist process, and it marks a step forward into reinforcing the capitalist abstraction of reality—and overall, 'the movement of industrial capital is this abstraction *in actu*' (Marx, 1863–1878: 61). For sure, nowhere in his essays does Marx say that capitalism is based on the 'power concentrated by the holders of money'. As we have already addressed this issue in

detail (Miconi, 2022: 112–116), we will limit ourselves to notice that Mayer-Schönberger and Ramge, in such a way, in their turn run the risk of isolating the data from the general process they are part of—'the movement of industrial capital'.

The impact of Zuboff's book on the international debate can hardly be overestimated, and it is not our intention to diminish the importance of her recognition on the exploitation strategies put in place by the digital companies. In particular, *The Age of Surveillance Capitalism* is precious in providing an overall view of the whole arsenal of technological tools, legal agreements, and lobbying activities deployed by the most influential players. As far as this article is concerned, an epochal transition is stated from the industrial regime, which used to transform the 'nature's raw materials into commodities', to 'surveillance capitalism', that rather 'lays its claims to the stuff of human nature for a new commodity invention' (Zuboff, 2019: 94). A transition, in short, from the realization of material commodities to the conversion of people's life into 'behavioral data'—what she refers to as 'rendition' (Zuboff, 2019: 232–242)[3]—by which the market trade of 'future behavior' prediction and modification is propelled. In this respect, Zuboff takes a position paradoxically comparable to that of post-industrial literature, according to which the discontinuity with traditional manufacturing is marked by *what* is produced, more than by the *how* it is produced, or the overall restructuring of the system toward delocalization and flexibility (which is the post-Fordist correction of the discourse). This is totally acceptable per se, we would argue, whereas it can be questioned its compatibility with the concepts of *surplus* and *surveillance capitalism*, that Zuboff does endorse, and in explicit fashion. In this respect, *Critical Inquiry* similarly took the stance that 'it is no longer enough to say that data is big': 'data is now in a state of surplus', as 'we have progressed from the megabyte, the terabyte, the petabyte, and now in 2022 debatably to the zettabyte era' (Halpern et al., 2022: 197). That a huge quantitative growth would also result in a qualitative change is indubitable, and it is part of the Marxist critical scheme as well—at least since Friedrich Engels' *Dialectics of Nature*. That such an increase in the dimensions of information archives would generate a *surplus*—a data, a cognitive, or an economic surplus——is instead questionable. According to Marx, surplus is not premised on the accumulation of the resources: on the opposite, *it is the accumulation that requires the very existence of the surplus*, which in its turn is due to the implementation of a specific relation of production.

> We have seen how money is changed into capital; how through capital surplus-value is made, and from surplus-value more capital. But *the accumulation of capital presupposes surplus-value*; surplus-value presupposes capitalistic production; capitalistic production presupposes the pre-existence of considerable masses of capital and of labour power in the hands of producers of commodities. (Marx, 1867: 507, italics added)

In the Marxist perspective evoked and advocated by the above-cited concepts, going back to Zuboff, industrial capitalism is not based on the conversion of 'raw materials into commodities', which is a structural epiphenomenon (though historically relevant, at that), but on a social fact: namely, the invention of a new relation of production. 'Relative surplus-value', Marx notes, is 'directly proportional' to productiveness, and therefore to the amount of time necessary to the realization of a given commodity. Hence,

> there is immanent in capital an inclination and constant tendency, to heighten the productiveness of labour, in order to cheapen commodities, and by such cheapening to cheapen the labourer himself.
>
> *The value of a commodity is, in itself, of no interest to the capitalist*. What alone interests him, is the surplus-value that dwells in it, and is realisable by sale. Realisation of the surplus-value necessarily carries with it the refunding of the value that was advanced. (Marx, 1867: 224, italics added)

This theoretical incongruency is symmetrical to that of José van Dijck, which tried to combine the definition of platform as a multi-sided market with the Marxist notion of commodification (see van Dijck et al., 2018: 37, 59). This exercise will unlikely be productive, we opine: according to the multi-sided model, the market is the place where offer and demand meet, the free competition triggers the productions of wealth, and counter-balances are in place. Following Marx, at the opposite, the appropriation of value takes place well before the players access the market stage, 'in the background' of capitalist economy (Marx, 1867: 112–115 in particular). Indeed, van Dijck, Poell and de Waal do not refer to the data as raw material, and they focus on their exploitation: which nonetheless is based on the asymmetrical relation between different players, which can hardly be explained upon the principles of network and multi-sided economy (see Miconi, 2022).

In the case of Zuboff, we may advance that such incongruency is not totally surprising, from our view. When compared to the world acclamation of the book, it is even awkward to remark upon some theoretical shortcuts, when not shortcomings of Zuboff's work, which are going unnoticed. A similar nonchalant use of the concept of 'Big Other' can be observed, which is even equaled to a 'puppet master' (Zuboff, 2019: 376), without Slavoj Žižek's interpretation being dealt with (not to mention the original Lacan's category[4]), which had already been brought into the Internet studies by Jodi Dean (2005: 162–167 in particular), and has been recently revived by Jamil Khader (2022) and Matthew Flisfeder (2022). It is a shame, as Žižek—through the 'immanence of resistance to Power' attributed to Foucault—points to the big other as a governmental entity both codifying the norms of obedience, and subsuming and prescribing the possible practices of subversion (Žižek, 1999: 261–264 in particular). It may be suggested that the two-way communication of the social media, in this respect, materially embodies the objective spirit that Žižek refers to as big other, with people's feedback being incorporated and looping back into the strategies of the media companies—which in the end is, beyond recognition, a radical realization of the cybernetics' control principles. In addition, Zuboff hardly problematizes the distinction between the two semantic areas covered by the notion of surveillance[5]: the *disciplinary* strategy, evolving from the centralized methods to the reciprocal control among web users (Lyon, 2018: 49), if not to its embodiment in people's life (Andrejevic, 2020: 75); and the *exploitation* strategy, as a way to watch over the workers, in order to augment the productivity or, in Marxist terms, to increase the rate of relative surplus extracted from their activity. This can be inferred from the pages that Marx (1867) dedicated to the optimization of production practices, by means of the imposition 'on the workman [of] increased expenditure of labour in a given time' (p. 279), which is made possible by the fact that the 'labourer works *under the control* of the capitalist to whom his labour belongs' (p. 131, italics added). It is finally in the section related to the *Progressive production of a relative Surplus Population or Industrial Reserve Army*, that Marx's prose takes on the most unsettling of the tones:

> [. . .] within the capitalist system all methods for raising the social productiveness of labour are brought about at the cost of the individual labourer; all means for the development of production transform themselves into means of domination over, and exploitation of, the producers; they mutilate the labourer into a fragment of a man, degrade him to the level of an appendage of a machine, destroy every remnant of charm in his work and turn it into a hated toil; they estrange from him the intellectual potentialities of the labour process in the same proportion as science is incorporated in it as an independent power; they distort the conditions under which he works, subject him during the labour process to a despotism the more hateful for its meanness; they transform his life-time into working-time, and drag his wife and child beneath the wheels of the Juggernaut of capital. (Marx, 1867: 451)

The centrality of digital surveillance to the exploitation of people's labor has been accounted by Mark Andrejevic (2012), Thomas Allmer (2015: 138–139), and Sebastian Sevignani (2016: 71),

in all cases by individuating a hiatus between the perceived importance of political surveillance and the unawareness or under-estimation of the control exercised by the employers and the companies. Speaking of critical theory, it is our impression that Zuboff overly prioritizes the first dimension (*critique*) over the second (*theory*): which is totally legitimate, as well as it has to be laid out. It is probably no accident that the paradigm orientating her work is based on Skinner's psychology (Zuboff, 2019: 295, 370–371 in particular), on the experiments in behavior modification, when not on the onset of subliminal messages (pp. 293–295, 300–301). Zuboff's (2019) insistence on the imposition of behavioral changes, by means of the fabrication of social cues (pp. 299–302 in particular), is possibly another step backward in our understanding of social media. In particular, we have in mind Jodi Dean, who made a theoretical move away from the ancient idea of the 'false consciousness', so as to examine the attention loop between the user and the platform 'in terms of the beliefs underlying practice', following Žižek: and therefore, as a grammaticalization of people's spontaneous perceptions and desires, rather than as the induction of artificial needs (Dean, 2010: 5).

Moreover, it remains quite unclear how to harmonize Zuboff's approach with the whole legacy of communication research, according to which the media are very strong in reinforcing pre-existing attitudes, with little to none empirically proven effect, when it comes to changing people's mind and opinion. As we know, her position is that technological innovation has eventually provided the manipulators with the material tools only dreamt of by Skinner—basically, the tools for mastering the 'social physics', as codified by Alex Pentland (Zuboff, 2019: 360–366).[6] It therefore makes sense, finally, to notice how Zuboff's analysis also conflicts with contemporary literature, which is all about the algorithms customizing the offer based on the users' preferences and on the already contracted habits—the 'amplification of users' behavior', more than its shaping (Chun, 2008: 58)—in terms of web clusterization, filter bubbles, or balkanization of the Internet.

'The Fantastic Form of a Relation Between Things'

The last vague in the big data discourse is certainly that of automated or synthetic data: basically, all those information streams that are artificially created by the software, without the humans taking any part in it. Among the examples of its application, we can cite the building of forecasting models, or the financial speculation based on the existing, or underlying models—all cases in which the data are used to generate data about the future state of things. It is in this respect that Adam Arvidsson—which has been using concepts from *post-operaism*o to address digital capitalism—speaks about an equivalent 'social logic of derivative', with data being used as a new biopolitical means of control, through 'probabilistic inferences' about people's 'movement and activities' (Arvidsson, 2016).

In a general stance, the category would encompass several variants, ranging from machine learning to generative algorithms, to the omnipresent ChatGPT, and the more. This strain of technological advancement is expected to be a real game-changer, and for reasons that appear to be evident, at a first look. Both the data uploaded by the users and those automatically extracted, this approach argues, require the participation of a human being[7]: either in terms of actively producing contents or exposing oneself to the tracking, spying, geo-localization, and to the quantification of one's own self (see Till, 2014). In the other way, James Steinhoff writes, 'the appeal of synthetic data for data-intensive capital' is something well beyond the regime of 'surveillance capitalism' we have discussed in the previous section: it would even affirm 'a historical tendency within capitalism toward the autonomization of the circuit of capital' (Steinhoff, 2022: 2). A tendency toward the overcoming of people's exploitation: it may well be, and here we will not take any position, as it is our belief that scientific research cannot foresee what is still to happen. In the meantime, it is of

advantage to narrow down the scope to the three versions of synthetic data production listed out by Steinhoff: data augmentation, generative techniques, and simulated environments. Data augmentation, to start with, can be defined as 'the automated application of minor modifications to a dataset in order to enlarge or diversify it'. In the case of the 'generative model approach', in a similar vein, the 'researchers created synthetic datasets' from 'publicly available datasets', expected to allow for more advanced predictive analyses. The most disruptive innovation is ignited by the 'simulated environments': 'once they are constructed', in fact, they would create an artificial world which is 'not extrapolated from an existing dataset' (Steinhoff, 2022: 5–9).

While we do not have the necessary expertise to discuss the technical features of these advanced solutions, we will observe them through the prism of the critical theory, rooted in the notion of value. As we saw, data augmentation and generative models both require the input of a pre-existing set of data, and even the synthetic environments are simulated by the machine, once the machine itself has been built and set-up. As we can read in another report, it is about 'the generation of synthetic data sets where some or all of the observed data have been replaced by synthetic values *generated from models based on the original data*' (Snoke et al., 2018: 664, italics added). It is possible that the automated production of data will become increasingly central to digital economy: this notwithstanding, we do not see how it would create value anew, as in all cases the process starts from something *that was already there*. This reminds us of what Marx discussed in the Draft Chapter 6, where he takes into account the case of labor, when it is 'absorbed in so far as capital takes the shape of the specific means of production'.

> But in the labor process which is at the same time a capitalist production process the means of production make use of the worker, with the result that labor appears only as a means whereby a certain *amount of value*, hence a certain amount of *objectified labor*, absorbs living labor in order to preserve and increase itself. Thus, the labor process appears as a *process of the self-valorisation of* objectified labor by means of living labor. (Marx, 1864)

What appears with the likes of the 'dead capital'—the machine—is actually the crystallization of a given quota of 'objectified labor': and no matter how many transformations the commodity will go through, it will ultimately echo that original moment (which is, in Marxist terms, the never-ending recourse of the primitive accumulation[8]). In our perspective, this consideration can also apply to the positions taken by Mayer-Schönberger, Couldry and Mejias, and Zuboff. For sure, these authors have to be credited with an advanced reflection of the role played by the big data: and it is notable that, for all them, the exploitation of the stored information goes beyond their simple trading or selling (see Couldry and Mejias, 2019: 7; Zuboff, 2019: 152), to which it is often, and simplistically reduced. What the same authors have rather in common with current literature is the concession, so to speak, to *data-centrism*: with attention placed on the data themselves, and less on the upstream moment of the data generation, and on the downstream moment of data assemblage, sharing and reuse. About this downstream stage, which is not the core topic of the article, we will limit ourselves to a short consideration. The way people concretely engage with the data—what Barbara Lupton (2017) has come to define the 'lively data'—has been studied in ethnographic inquiries, which show—if anything—the variety of motivations and practices in which even the users of quantified-self services actually engage. Here 'self-tracking is also about exploring new forms of expression that do not privilege numbers a priori', as a 'means of meaningfully communicating in and navigating a world that speaks' both languages of numbers and human contents (Sharon and Zandbergen, 2017: 10).

Similarly, the point we need to highlight is not simply that there are other sources of profit in the platform economy, along with data exploitation. This has been largely investigated in scientific

literature, ranging from Christian Fuchs' (2014) study of the global value chain, to Antonio Casilli's (2019) and Sarah Roberts' (2019) investigations on the material dimensions of the work required by the information economy. We are already accustomed to think at this in terms of a 'division of labor between humans and machines', as even the allegedly automated services, such as Google translator, still involve human contributors (Ekbia, 2017: 21 and 219). As Sarah Roberts (2019) observes, 'there are still many things that' the humans 'can do much more effectively than computers, such as identifying objects in a photo or video' (p. 35)—and last but least, they still serve as 'a cheap pools of laborers' (p. 48), when compared to the costs of technological innovation.

More specifically than that, though, it is the very discourse about the data to be flawed, as it hides the social relations of production beneath the guise of a purely objective outcome: in a way that reminds us of one of the most powerful concepts we owe to Karl Marx. 'A commodity is therefore a mysterious thing', the idea goes, as the social dimension of human's work appears as an 'objective character', with the relation between workers and capital no longer being visible.

> There it is a definite social relation between men, that assumes, in their eyes, the fantastic form of a relation between things. In order, therefore, to find an analogy, we must have recourse to the mist-enveloped regions of the religious world. In that world the productions of the human brain appear as independent beings endowed with life, and entering into relation both with one another and the human race. So it is in the world of commodities with the products of men's hands. This I call the Fetishism which attaches itself to the products of labour, so soon as they are produced as commodities, and which is therefore inseparable from the production of commodities. This Fetishism of commodities has its origin, as the foregoing analysis has already shown, in the peculiar social character of the labour that produces them. (Marx, 1867: 47–48)

Our brief literature review revealed something analogous: the isolation of the data from their social milieu, in the likes of raw material, self-explaining statistical patterns, synthetic simulations, or needs imposed by the big other. What we can easily label, in the end, as a form of digital *fetishism*: the hypostatizing of the final output, which takes out of the discourse the social process by which this very same result is produced. As Nick Couldry and Andreas Hepp acutely observed, 'big data approaches to social knowledge do not depend on any theory of communication', while 'they do depend on ignoring the contextual nature of information' (Couldry and Hepp, 2016: 194). Starting from a Marxist perspective, Christian Fuchs (2017) similarly asserts that the exaltation of big data 'disregards the societal contest, contradictions and power structures, into which computer technologies are embedded' (p. 588).

A final note about this tendency toward the reification of the data, which is eclipsing the memory of the social entanglement of any production process. We observed that the big data paradigm would lead to a transition—and this is possibly the main argument taken as a proof of its radical novelty—from the capitalist appropriation of human work, to that of the human being as a whole. In Couldry and Mejias (2019), it is 'the discovery of new forms of raw material', that we have already discussed, which allows 'data colonialism' to expand 'by appropriating for exploitation even more layers of *the human life itself*'. In the same paragraph, and consistently, the authors remark upon the 'excessive focus on whether digital labor is being exploited', which is 'not the most important feature of today's transformations' (p. 5). The same aspect is occasionally touched on by Mayer-Schönberger and Cukier (2013)—'datafication is not just about rendering attitudes and sentiments into an analyzable form, but human behavior as well' (p. 93)—while rising to the forefront in Zuboff's work. In what she calls the 'behavioral value reinvestment cycle' (p. 69), people's attitudes and practices become the oil of the XXI century: in terms of 'instrumentalization of behavior for the purposes of modification, prediction, monetization, and control' (Zuboff, 2019: 352).

In all this talking about the colonization of people's behavior, though, it strikes the lack of reckoning of a plain fact: that capitalism exploiting human life is *nothing new*, from the standpoint of critical theory. In his 1975 essay on *Proletarians and the State*, Antonio Negri wrote,

> [. . .] for this new proletariat, it is no longer exclusion from the extortion of capitalist work that is specific, but on the contrary inclusion within the totality of the productive social process and withing the twists and turns of its conditions that is fundamental. We have seen the mass worker (the first massive concretization of capitalist abstraction of labor) produce the crisis. Now we see restructuring, far from overcoming the crisis, unfolding and lengthening its shadows over the whole of society, not on its surface but in its heart. (Negri, 1975: 126)

Negri's thesis is based on a cornerstone of the Italian *operaismo*: the idea, dating back to Mario Trionti's *Operai e capitale*, that the extraction of relative surplus-value, as it requires the reorganization of the human relations rather than their formal subsumption, is destined per se to break the barriers between the work activities and the rest of society (Tronti, 1962: 24). In another way, Negri's approach is very peculiar, as it turns upside down the classical Marxist interpretation: here contemporary history is not shaped by the industrial forces but driven by the evolution of the social body—from the mass to the multitude—with the capital *successively* fighting back, by adapting itself to the various incarnations of its class enemy. We cannot say if Negri is totally right, in this respect, and in any case, this would be too broad an issue, for the limits of our article. What we do know is that Tronti's and Negri's radical reading of Marx, a few decades ago, put on the agenda an element destined to occupy the international debate: that capitalist exploitation is not limited to the appropriation of the working time, while enlarging its domain to the 'whole of society' and vampirizing people's life, energy, emotions, and attention (the so-called 'social factory').

Conclusion

In this article, we argued that current literature on the big data paradigm is flawed for two major, and deeply intertwined reasons: the over-emphasizing of the novelty with the previous stages in the evolution of capitalism and the under-estimation of the role played by human labor. In the first case, it is to be recalled not only the importance of industrial production in the Internet society, but also the fact that platform economy, for its very spatial ambition and configuration, appears as a specific conjuncture in the long-durée evolution of the world system, as Peck and Phillips (2020) brilliantly observed in a Braudelian perspective (pp. 86–87). In the second case, we refer to a growing a body of research—the above-cited works by Fuchs, Casilli, and Roberts—which is unraveling the most disparate job activities, that post-Fordist accumulation tends to hide from consumers' view. The same expression digital fetishism, in this direction, has been already used by Fuchs (2016), in his analysis of the reification induced by the social media (pp. 86–87).

There is no reasonable doubt that technological innovation has recently equipped the private corporations with powerful tools for the exploitation of users' activities. Nonetheless, it is significant that much literature on the big data companies' appropriation of human life hardly recognizes any credit to the post-1970s Marxist theory—in actuality, their most prominent authors are not quoted whatsoever, in the reviewed studies.[9] Here, the removal of the social dimensions of the data generation process, as stated, is coupled with another intellectual tendency: the over-emphasizing of the discontinuity between the age of industrial exploitation and that of data exploitation. Further investigation will be needed, to better understand whether the big data is establishing a totally new paradigm, or it results from the mid-range effects of the reorganization of international capitalism after the 1970s downturn: what Giovanni Arrighi would define the historical interval between the

'signal crisis' and the 'terminal crisis' of the accumulation regime (Arrighi, 2005b: 111–112; see also Arrighi, 2005a). As a matter of fact, it is in the 1970s that the colonization of human life on the part of capital has been observed, way before technological innovation allowed for wide-scale information collection and tracking.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This publication is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004488. The information and views in this publication are those of the authors and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

ORCID iD

Andrea Miconi https://orcid.org/0000-0002-1960-6243

Notes

1. Petit's 'cognitive holigopoly' is close to the condition we commonly define as de facto monopoly, and it also resembles to the concept of 'monopsony' introduced by Couldry and Mejias (2019): for which, as simple as that, despite the exponential growth of 'digital tools', in practice 'there are limited options for choosing which platform to participate in' (p. 43). For a critical perspective on digital monopolies, see Bilić et al. (2021: 59–74) in particular.
2. In a different vein, Couldry and Mejias (2019) observe that data management has been originally monopolized by public institutions, while 'it was not until the 1970s and the 1980s in the United States that large private companies (marketers) acquired a data-collection capacity to rival that of the state' (p. 121; same finding in Halpern et al., 2022: 98). This is not necessarily in contradiction with Puschmann and Burgess' analysis, in our opinion, as the codification of the big data ideology might well be the historical watershed between the two stages. This overtaking of the sector on the part of private companies would be consistent with Hugo Moreira's (2010) explanation: which highlights the backwardness of public regulation when compared to the pace of technological innovation, with the big data era only starting in 2010, for what concerns the role of national and European policy-makers.
3. The tendency to rename the already existing concepts, in the academic discourse, would deserve some serious attention. It is our belief that *rendition* is basically the same as *datafication*—while *heteromation*, well, is in substance the same as *exploitation* (see Ekbia, 2017).
4. As much as the social media can be imagined as the living manifestation of the big other—or better, the only space in which it can possibly materialize—the reference to Lacan is quite problematic, especially as the connection between mega-platforms and the big other can hardly be explained in terms of desire (see Flisfeder, 2022: 416–417 in particular; and Dean, 2010: 39–41 and 93–95 in particular).
5. In terms of critical theory, it is also interesting how Zuboff reflects on the post-9/11 'state of exception', without dealing at all with Carl Schmitt and Giorgio Agamben, whose contributions could have enriched the discourse.
6. Zuboff (2019) is simply right in observing how Pentland 'transforms Skinner's fusty, odd utopia into something that sounds sophisticated, magical and plausible, largely because it resonates with the waves of applied utopistic that wash over our lives every day' (p. 430). She is right, because the proposal of using the data *for a greater good*—such as 'design[ing] a human centric society' (Pentland, 2014: 193)—is as dangerous as, and possibly more dangerous than their commercial exploitation.
7. The difference between uploaded data and extracted data is relevant, as these two modalities—respectively, active and passive—also imply two diverse exploitation strategies. This notwithstanding, this distinction it is quite often overlooked in scientific literature (see, for instance, Andrejevic (2009) and Cockaine (2016)—which, in other verses, provide relevant insights into the state of digital capitalism).

8. See Marx (1867: 506–509). We have no alternative but to use the word 'primitive', which nonetheless is an imprecise adaptation of the German locution: 'Unsprüngliche Akkumulation', which might better translate into 'originary accumulation'. This is not simply a philological question, while bearing theoretical consequences: the 'primitive accumulation' concept would picture a specific moment, lost in remote history, while the second solution allows for the understanding of the recursive return of this 'originary' moment, over the centuries, in correspondence to the rise of new exploitation methods.
9. If we go into the details, it is emblematic that Negri is never quoted in Zuboff's *The Age of Surveillance Capitalism*, where Marx himself is mentioned three times, and very shortly: as he is alluded to in Hannah Arendt's (2019: 99) work; as the author of the *Communist Manifesto* (p. 221); and as the inspirator of 'Lenin, Stalin and Mao' (p. 406)—so much, for the man who invented the scientific critique to capitalism. Couldy and Mejias, on their part, only cite Negri once, and namely his occasional reference to the big data (Couldy and Mejias, 2019: 35; the reference is to Hardt and Negri, 2017: 169).

References

Allmer T (2015) *Critical Theory and Social Media: Between Emancipation and Commodification*. London: Routledge.

Anderson C (2008) The end of theory: the data deluge makes the scientific method obsolete. *WIRED*, 24 June. Available at: https://www.wired.com/2008/06/pb-theory/

Andrejevic M (2009) Privacy, exploitation, and the digital enclosure. *Amsterdam Law Forum* 1(4): 47–62.

Andrejevic M (2012) Exploitation in the data mine. In Fuchs C (ed.) *Internet and Surveillance: The Challenges of Web 2.0 and Social Media*. London: Routledge, pp. 71–88.

Andrejevic M (2020) *Automated Media*. New York and London: Routledge.

Arrighi G (2005a) Hegemony unravelling—I. *New Left Review* 32: 32–80.

Arrighi G (2005b) Hegemony unravelling—II. *New Left Review* 33: 83–116.

Arvidsson A (2016) Facebook and finance: on the social logic of the derivative. *Theory, Culture & Society* 33(6): 3–23.

Barabási A-L (2010) *Bursts: The Hidden Patterns behind Everything We Do, from Your E-mail to Bloody Crusades*. New York: Dutton.

Bilić P (2016) Search algorithms, hidden labour and information control. *Big Data & Society* 3: 1–9.

Bilić P, Prug P and Žitko M (2021) *The Political Economy of Digital Monopolies. Contradictions and Alternatives to Data Commodification*. Bristol: Bristol University Press.

Boellstorff T (2013) Making big data, in theory. *First Monday* 18(10). DOI: 10.5210/fm.v18i10.4869.

Bolin G (2022) The value dynamics of data capitalism: cultural production and consumption in a datafied world. In Hepp A, Jarke J and Grampf L (eds) *New Perspectives in Critical Data Studies*. London: Palgrave Macmillan, pp. 167–186.

Casilli A (2019) *En attendant les robots. Enquête sur le travail du clic*. Paris: Seuil.

Chun WHK (2008) *Updating to Remain the Same: Habitual New Media*. Cambridge: MIT Press.

Cockaine DG (2016) Affect and value in critical examinations of the production and 'prosumption' of big data. *Big Data & Society* 2: 1–11.

Couldry N and Hepp A (2016) *The Mediated Construction of Reality*. London: Polity Press.

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

Dean J (2005) Žižek against democracy. *Law, Culture and the Humanities* 1: 154–177.

Ekbia HR and Nardi B (2017) *Heteromation, and Other Stories of Computing and Capitalism*. Cambridge, MA: The MIT Press.

Dean J (2010) *Blog Theory: Feedback and Capture in the Circuits of Drive*. London: Polity Press.

Flisfeder M (2022) Whither symbolic efficiency? Social media, new structuralism, and algorithmic desire. *Rethinking Marxism* 34(3): 413–432.

Fuchs C (2014) Theorising and analysing digital labour: from global value chains to modes of productions. *The Political Economy of Communication* 21(1): 3–27.

Fuchs C (2016) The Internet, social media and Axel Honneth's interpretation of Georg Lukács' theory of reification and alienation. In: Fuchs C (ed.) *Critical Theory of Communication: New Readings of Lukács,*

Adorno, Marcuse, Honneth and Habermas in the Age of the Internet. London: University of Westminster Press, pp. 153–175.
Fuchs C (2017) Günther Anders' undiscovered critical theory of technology in the age of big data capitalism. *tripleC* 15(2): 582–611.
Gitelman L (ed.) (2013) *'Raw Data' Is an Oxymoron*. London and Cambridge: MIT Press.
Hardt M and Negri A (2017) *Assembly*. New York: Oxford University Press.
Halpern O, Jagoda P, West Kirkwood J, et al. (2022) Surplus data: an introduction. *Critical Inquiry* 48(2): 197–210.
Khader J (2022) Welcome to the metaverse: social media, the phantasmatic big other, and the anxiety of the prosthetic Gods. *Rethinking Marxism* 34(3): 397–405.
Lohr S (2013) The origins of 'big data': an etymological detective story. *The New York Times*, 1 February. Availableat:https://archive.nytimes.com/bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/#:~:text=The%20unruly%20digital%20data%20of,technical%20Web%20sites%2C%20appropriately%20enough
Lupton B (2017) Personal data practices in the age of lively data. In: Daniels J, Gregory K and Mcmillan Cottom T (eds) *Digital Sociologies*. Bristol: Policy Press, pp. 339–352.
Lyon D (2018) *The Culture of Surveillance*. London: Polity Press.
Marx K (1857–1861) *Fragment on Machines* (trans. M Nicolaus. Annotated by B Fowkes). Available at: https://www.marxists.org/
Marx K (1863–1878) *Capital. A Critique of Political Economy. Volume II: the Process of Circulation of Capital* (ed. F Engels and Proofed by A Blunden, C Clayton and M Harris). Available at: https://www.marxists.org/
Marx K (1863–1883) *Capital. A Critique of Political Economy. Volume III: Process of Capitalist Production as a Whole* (ed. F Engels, proofed and corrected by C Clayton and M Harris). Available at: https://www.marxists.org/
Marx K (1864) *Draft Chapter 6 Capital* (trans. B Fowkes). Available at: https://www.marxists.org/
Marx K (1867) *Capital. A Critique of Political Economy. Volume I: the Process of Production of Capital* (trans. S Moore and E Aveling, ed. F Engels and Proofed by A Blundend, C Clayton, M Harris and D Allinson). Available at: https://www.marxists.org/
Mayer-Schönberger V and Cukier V (2013) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. London: John Murray.
Mayer-Schönberger V and Ramge T (2018) *Reinventing Capitalism in the Age of Big Data*. London: John Murray.
Miconi A (2022) On value and labor in the age of platforms. In: Armano E, Risi E and Briziarelli F (eds) *Digital Platforms and Algorithmic Subjectivities*. London: Westminster University Press, pp. 107–119.
Moreira H (2010) *Governing Knowledge and Technology: Technological Pressure for Convergence in EU, California, and China Data Protection Regulation*. Lisbon: ISCTE-IUL. Available at: https://osf.io/preprints/socarxiv/v6uf3/
Negri A (1975) Proletarians and the state: toward a discussion of workers' autonomy and the historical compromise. In: Negri A (ed.) *Books for Burning: Between Civil War and Democracy in 1970s Italy*. London: Verso Books, pp. 118–178.
Peck J and Phillips R (2020) The platform conjuncture. *Sociologica* 14(3): 73–99.
Pentland A (2014) *Social Physics. How Good Ideas Spread—The Lessons from a New Science*. New York: Penguin Books.
Petit N (2020) *Big Tech and the Digital Economy: The Moligopoly Scenario*. New York and Oxford: Oxford University Press.
Puschmann C and Burgess J (2014) Metaphors of big data. *International Journal of Communication* 8: 1690–1709.
Räsänen M and Nyce JM (2013) The raw is cooked: data in intelligence practice. *Science, Technology, & Human Values* 38(5): 655–677.
Roberts S (2019) *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven, CT: Yale University Press.

Rouvroy A (2012) The end(s) of critique: data behaviourism versus data process. In: Hildebrandt M and De Vries K (eds) *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology*. London: Routledge, pp. 143–167.

Schumpeter JA (1934) *Theory of the Economic Development: An Inquiry into Profits, Capital Credit, Interest, and the Business Cycle*. New Brunswick and London: Transaction Publishers (1983).

Sevignani S (2016) *Privacy and Capitalism in the Age of Social Media*. London: Routledge.

Sharon T and Zandbergen D (2017) From data fetishism to quantifying selves: self-tracking practices and the other values of data. *New Media & Society* 19(11): 1695–1709.

Snoke J, Raab GM, Nowok B, et al. (2018) General and specific utility measures for synthetic data. *Journal of the Royal Statistical Society* 181(3): 663–688.

Srnicek N (2017) *Platform Capitalism*. London: John Wiley & Sons.

Steinhoff J (2022) Toward a political economy of synthetic data: a data-intensive capitalism that is not a surveillance capitalism? *New Media & Society*. Epub ahead of print 11 June. DOI: 10.1177/14614448221099217.

Till C (2014) Exercise as labour: quantified self and the transformation of exercise into labour. *Societies* 43(4): 446–462.

Tronti A (1962) Factory and society. In: *Workers and Capital* (trans. D Broder). London and New York: Verso Books: pp. 12–35.

Van Dijck J, Poell T and de Waal M (2018) *The Platform Society: Public Values in a Connective World*. New York: Oxford University Press.

Van Dijk I (1999) The one-dimensional society of manuel castells. *New Media & Society* 1(1): 127–138.

Wu T (2010) *The Master Switch: The Rise and Fall of Information Empires*. New York: Knopf.

Žižek Z (1999) *The Ticklish Subject: The Absent Centre of Political Ontology*. London: Verso Books.

Žižek Z (2012) *Less Than Nothing: Hegel and the Shadow of Dialectical Materialism*. London: Verso Books.

Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future and the New Frontier of Power*. London: Profile Books.