UvA-DARE (Digital Academic Repository)

For a heterodox computational social science

Törnberg, P.; Uitermark, J.

DOI

Publication date
2021

Document Version
Final published version

Published in
Big Data and Society

Link to publication

Citation for published version (APA):

Original Research Article

For a heterodox computational social science

Petter Törnberg (iD) **and Justus Uitermark** (iD)

Abstract

The proliferation of digital data has been the impetus for the emergence of a new discipline for the study of social life: 'computational social science'. Much research in this field is founded on the premise that society is a complex system with emergent structures that can be modeled or reconstructed through digital data. This paper suggests that computational social science serves practical and legitimizing functions for digital capitalism in much the same way that neoclassical economics does for neoliberalism. In recognition of this homology, this paper develops a critique of the complexity perspective of computational social science and argues for a *heterodox* computational social science founded on the meta-theory of critical realism that is critical, methodological pluralist, interpretative and explanative. This implies diverting computational social science' computational methods and digital data so as to not be aimed at identifying invariant laws of social life, or optimizing state and corporate practices, but to instead be used as part of broader research strategies to identify contingent patterns, develop conjunctural explanations, and propose qualitatively different ways of organizing social life.

Introduction

Over the past decade, digital data and methods have become the foundation of an emerging paradigm for studying social behavior. Computational approaches to social scientific research have begun to cluster under the label of 'computational social science' (CSS), an interdisciplinary subfield that 'includes analysis of web-scale observational data, virtual lab-style experiments, and computational modeling' (Watts, 2013: 6). Bringing powerful new methods to bear, this field has opened up exciting new research opportunities and enabled unprecedented examination of social phenomena. CSS is one of the most rapidly growing fields in academia, involving a wide range of scientific disciplines as well as strong links to the private sector, and is gaining recognition from funding agencies, governments, and the media (Edelmann et al., 2020).

Although still diverse and in rapid change, CSS is gradually beginning to establish itself as a paradigm through a growing number of dedicated conferences and journals as well as textbooks and manifestos. Leading scholars often characterize CSS as not only united by more than the use of digital data and methods, but also by a set of epistemic perspectives – a 'Weltanschauung' founded in a 'computational paradigm of society' (Cioffi-Revilla, 2017: 1). This perspective draws praxeomorphically (Bauman, 2013: 56)

on the interactional structure of digital data, contrasted with 'rows of cases and columns of variables' (Lazer et al., 2020: 1060) of traditional quantitative methods by emphasizing complexity and interaction, networks, and flows. Researchers with this understanding of CSS view the advent of digital data as an opportunity to bring to the study of social life the rigor, scope, and certainty of the natural sciences. To these scholars, the expansion of metaphors from the natural sciences appears as a promise of a coming data-driven 'hard' social science; a 'sociology of the 21st century' enabled by data giving us 'the chance to view society in all its complexity, through the millions of networks of person-to-person exchanges' (Pentland, cited in Manovich, 2011: 464). At the base of this promise is the conviction that society is a complex system that can be understood through theories and methods that have been developed to understand other complex systems

Amsterdam Institute for Social Science Research, University of Amsterdam, Amsterdam, the Netherlands

Corresponding author:
Petter Törnberg, Amsterdam Institute for Social Science Research, University of Amsterdam, Postbus 15629, 1001 NC, Amsterdam, the Netherlands.
Email: p.tornberg@uva.nl

such as brains, ecosystems, or ant colonies. Not all scholars identifying as computational social scientists operate under this framework, but the perspective constitutes the implicit or explicit foundation for a dominant subcurrent of CSS. It is this mainstream that we focus on and engage with in this paper.

The advent of digital data has, however, also separately spawned another major literature, which takes a political economy lens to examine the role and nature of digital data in contemporary society. Although CSS tends to view digital data as accidental byproducts of social life – often referred to as 'digital traces' or 'footprints' – that can be used to reconstruct interactions, this literature views digital data as commodities or a form of capital. Under labels such as 'informational capitalism' (Fuchs, 2010), 'communicative capitalism' (Dean, 2005), 'platform capitalism' (Srnicek, 2017), 'digital capitalism' (Sadowski, 2019), or 'surveillance capitalism' (Foster and McChesney, 2014; Zuboff, 2019), this literature views platforms as designed to produce and valorize data; they are deployed to monetize social interactions by quantifying social life to allow prediction and manipulation (Couldry and Mejias, 2020; Cukier and Mayer-Schoenberger, 2013). The value instilled in data relies on (belief in) the capacity to capture, analyze, predict, and control the social world (Van Dijck, 2014; see also Lohr, 2015). The data are, in other words, not an accidental trace of social interaction – *au contraire*, it would be more apt to say that the sociality is the afterthought of the extraction of digital data.

Placing CSS in light of the literature on digital capitalism, this paper suggests that the relation between CSS and contemporary capitalism is homologous to the relation between neoclassical economics and neoliberal capitalism, in the sense that that it constitutes a scientific paradigm which provides tools and scientific legitimacy to a mode of capitalist accumulation. CSS is located at the interstices between academia and industry, as the field provides the training, methods, theories, and legitimacy that help instill digital data with economic value, with many leading scholars being affiliated to companies such as Microsoft or Google. In light of this liaison between CSS and digital capitalism, many scholars have distanced themselves from CSS, instead developing alternative approaches to understanding the role of digital data, establishing subdisciplinary fields such as new media studies (Rogers, 2013), digital geography (Ash et al., 2018; Leszczynski, 2015), digital sociology (Marres, 2017), and critical data studies (Dalton et al., 2016).

This paper attempts instead to salvage the potential of methods and metaphors developed within CSS, by re-embedding the 'methodological approach of data-driven science within a different epistemological framing that enables social scientists to draw valuable insights from Big Data that are situated and reflexive' (Kitchin, 2014: 9). In recognition of the homology between CSS and

neoclassical economics, we propose a *heterodox* CSS (HCSS) that parallels the response of heterodox economics to neoclassical economics. This results in an approach that, similar to heterodox economics, is grounded in critical realism, and that is *critical*, *methodological pluralist*, *interpretative*, and *explanative*. It does not do away with the concept of complexity but rather situates complexity within a broader social ontology, premised on the conviction that there are no immutable laws of social life or any one optimal solution to social problems. The purpose of HCSS, then, is not to identify laws or provide optimal solutions but to develop conjunctural explanations, offer different ways of perceiving and addressing issues, and to propose qualitatively different ways of organizing social life.

CSS as an emerging paradigm

Although the roots of CSS can be traced back to decades of computational experiments and explorations (e.g. Conway, 1970; Langton, 1997), its recent explosive growth has been driven largely by the new availability of digital data. These data have brought computer scientists, physicists, and social scientists to coalesce under the flag of CSS, which is emerging as the foremost label for social scientific data analytics (Lazer et al., 2020). Although CSS remains in some ways a 'loosely connected intellectual community of social scientists, computer scientists, statistical physicists, and others' (Lazer et al., 2020: 1060), it is coming to formulate a distinct approach and metatheory, presented in a growing number of introductory textbooks and manifestos (Christakis, 2012). At the same time, the field is establishing itself institutionally through annual conferences (e.g. IC2S2) and dedicated journals (e.g. JCSS). CSS makes use of a toolkit of sophisticated computational methods, including natural language processing (Hirschberg and Manning, 2015), complex/social network analysis (Watts, 2004), machine learning (Molina and Garip, 2019), and agent-based modeling (Squazzoni, 2012).

Some CSS scholars argue that the field can be described as part of a new paradigm – the 'computational paradigm of society', as Cioffi-Revilla (2017: 2) refers to it in his introductory CSS textbook, characterized by a distinct worldview and epistemological perspective. The worldview described by Cioffi-Revilla draws strongly on physics and computer science, expressed in the oft-mentioned prediction that data will 'revolutionize' the social sciences. In this view, Big Data will enable a 'hard science' approach to the social world, in essence bringing it into the domain of physics. This hope can be observed in scholars such as Lev Manovich (2016) arguing that 'Digital is what gave culture the scale of physics, chemistry or neuroscience. Now we have enough data and fast enough computers to actually study the "physics" of culture'. The physicists Caldarelli et al. (2018: 870) similarly argue that the

proliferation of digital data 'provide the opportunity to build a "physics of society": describing a society—composed of many interacting heterogeneous entities (people, businesses, institutions)—as a physical system'. For these scholars, the promise of CSS is, thus, not only to contribute new methods to traditional social science research, but rather also to supplant it with 'an entirely new scientific approach for social analysis' (Conte et al., 2012: 327) aiming to 'uncover the laws of the society' (Conte et al., 2012). Nicholas Christakis similarly describes CSS as 'a new kind of social science' (Christakis, 2012), which answers to the crisis of the old approach of empirical sociology (Savage and Burrows, 2007) by supplanting surveys and interviews with data mining and simulation (Conte et al., 2012; Lazer et al., 2009). As Watts (2011: 266) puts it, 'just as the invention of the telescope revolutionized the study of the heavens, so too by rendering the unmeasurable measurable, the technological revolution in mobile, Web, and Internet communications has the potential to revolutionize our understanding of ourselves … we have finally found our telescope. Let the revolution begin'.

In some ways, this call goes beyond even making social science a 'hard' science in the sense of 'replicable, cumulative, and coherent' (Lazer et al., 2020: 1062). Watts (2017) proposes a 'solution-oriented' social science (see also Lazer et al., 2020: 1062), expressing frustration with the prevalence of a multitude of immensurable theories, and arguing that it is indicative of the dismal state of the social sciences that Microsoft's CEO would not be able to find a definitive answer in the scholarly literature to the question of how to optimally reorganize the corporation (Watts, cited in Van den Berg, 2017). Pentland (2015) similarly envisions a society in which massive data and new methods allow not only deeper understanding, but also make it possible to engineer 'better' social systems. Pentland points to social media platforms to show the possibilities to use social pressure to control and direct social life, treating society as a control problem which can be 'tuned' to produce 'better outcomes'. These scholars are in other words not merely aiming to understand social systems with the precision of a 'hard' science, but also to treat society as a system of engineering, to be tuned and optimized.

Society as a complex system

The conceptualization of society as a complex system is expressive of an underlying 'naturalism' in CSS: the belief that there is a continuity between the natural and social world that makes it unnecessary to appeal to qualities such as conscience, intentionality or meaning to account for social behavior. Although social physics may at times be presented as 'a new science' (e.g. Pentland, 2015), it harks back to the social physics of the 19th century, with for instance Auguste Comte's *Le Producteur* describing

social physics as a science that treats 'social phenomena […] as being subject to natural and invariable laws' (Iggers, 1959, cf. Conte et al., 2012). The naturalism of contemporary CSS is, however, of a different variety than that which informed 19th-century social physics. As Cioffi-Revilla (2017: 7) argues, CSS' naturalism draws on complexity science, with its historical roots in dynamical system theory and early scientific computing at the Los Alamos National Laboratory in the 1970s, and linked to earlier traditions of cybernetics (Galison, 1997). This ontological stance subscribes to naturalism but brings a fundamental criticism against the linearity and equilibria of traditional quantitative approaches (Cilliers, 2002). As Ball (2012: ix) puts it, in the complexity science perspective, the traditional conceptualization of social physics:

> … remains valid but it often drew on the wrong analogies. Society does not run along the same predictable, 'clockwork' lines as the Newtonian universe. It is closer to the kind of complex systems that typically preoccupy statistical physicists today: avalanches and granular flows, flocks of birds and fish, networks of interaction in neurology, cell biology and technology. (Ball, 2012: IX)

This notion of the social world being a 'complex system' in the same way as complex systems within the natural world is a recurring notion in CSS (e.g. Conte et al., 2012; Watts, 2013: 6), providing impetus to the notion that society is amenable to the same methods and approaches. This complexity perspective stems from physics and views macro-patterns as the emergent and aggregate outcome of micro-interactions (Conte et al., 2012). Such an approach allows powerful insight into phenomena in which the reflexivity and meaning-making of individuals can be bracketed, and captures well situations where these play limited roles, such as traffic congestions or pedestrian flows (Andersson et al., 2014). Many of the optimization algorithms that lay the foundation of data analysis are designed according to the same type of nature-inspired complex system metaphor, including machine learning and stochastic optimization. These algorithms are often built on simulating biological distributed computation systems, such as ants or swarms looking for food, neural networks, or biological evolution (Mitchell, 2009; Wahde, 2008). In this sense, the complexity perspective permeates CSS both methodologically and epistemologically.

The ontology of complexity aligns with the relational and interactive nature of digital data produced by digital platforms, contrasted against surveys, which are argued to slot reality into fixed categories, variables, and variances, concealing its interactional elements (Conte et al., 2012; Lazer et al., 2020). The epistemic features of digital data are, thus, taken to represent the true characteristics of the social world – heterogeneous, interactional, and emergent (Conte et al., 2012). Although census data are seen as

produced for scientific analysis, digital data are a 'naturally occurring by-product' (Edwards et al., 2013) of social processes, rather than something produced for scientific consumption. The data culled from platforms such as Facebook, Twitter, and Instagram are described as 'imprints' or 'traces' of people's actual behavior or moods (Lazer et al., 2020), informing us of what people actually *do*, instead of merely what they say they do (Pentland, 2015).

This type of uncritical praise of data as 'raw' and freely available has calmed somewhat in recent years, as questions of ethical issues and the role of platforms in shaping data have become topics of discussion also within the discipline. These are, however, not seen as fundamental epistemological issues, but rather as institutional and technical problems, to be resolved by, for instance, further consolidating CSS as a discipline, drawing out clear guidelines and setting ethical rules, and establishing models for collaboration and data sharing with the private sector (see Lazer et al., 2020).

Patterns and mechanisms

CSS has elaborated three main approaches to studying the social world: pattern-identification, simulation, and experimentation. These strands overlap and interlink, but also express internal tensions. The first strand comprises the identification of large-scale patterns in data. Informed by universalities in complexity science, this is a pursuit of regularities which apply in a range of social and natural domains, epitomized by the Feigenbaum constants in chaos theory (Cvitanovic et al., 2005), and by scaling and power laws (West, 2017). Such laws are often taken to point to underlying properties of the system, indicating the corresponding universal mechanisms at play (e.g. Watts, 2004; West, 2017). For instance, identifying a power-law distribution in social networks may be taken to imply preferential attachment in the formation of network ties (Barabási and Albert, 1999). The pattern-finding approach is also central to more descriptive research that provides a quantitative characterization of large data.

A second strand is the linking of macro-dynamics to micro-mechanisms through the use of simulations and computational modeling, such as agent-based models. Conte et al. (2012) refer to these models as aimed at 'generative explanation', in reference to Epstein's (2006) 'generative social science': 'computer code, that reproduce some key features of societies' (Conte et al., 2012; see also Cioffi-Revilla, 2016). Such models view social institutions, organizations, and behavior as 'emerging' from individual behavior, analogously to how the behavior of an ant colony emerges from the interactions of individual ants. This tends to take the form of formulating hypotheses about individual interactions, and then using computational methods to simulate a large number of causal steps, to see what type of macro-dynamics these generate, thus constituting a form of computational hypo-deductive modeling. An example is the Barabási–Albert model, which shows that when creation of a network tie at the micro-level depends on the node's previous number of connections, the outcome at the macro-level is a power-law distribution of ties (Barabási and Albert, 1999).

A final important strand of CSS research is large-scale online experiments, often aimed at identifying micro-mechanisms that produce emergent phenomena. This is epitomized by the use of large-scale online experiments, drawing from the commercial use of the so-called A/B testing to optimize digital technologies (Huang et al., 2018). In these experiments, researchers construct or modify online platforms and observe how user behavior is affected by a given treatment (Centola, 2010; Kramer et al., 2014). This experimental approach also extends to identifying natural experiments and examining their effect through always-on data sources (Aral and Nicolaides, 2017; Mas and Moretti, 2009), as well as artificial intelligence-powered methods for causal inference from observational data, for instance by automatically identifying most-similar cases in large data sets (Legewie, 2016; Pearl, 2019). Online experiments often focus on nudging or modifying the experience for individual users, to see how this brings about certain macro-level consequences. For instance, de Rijt et al. (2014) show through an online experiment that posts with many likes receive more likes, resulting in power-law distributions.

In summary, CSS is an emerging paradigm for social scientific research with strong links to the data analytics industry that methodologically brings powerful computational methods to bear within the social science domain, and epistemologically highlights the interactional complexity of the social world which traditional quantitative methods have tended to bracket. With the growth of CSS, 'the study of social phenomena has increasingly become the province of computer scientists, physicists, and other "hard" scientists' (Watts, 2013: 5), and digital social research is thus coming to be approached as a form of data analytics. As the field establishes itself through canonical publications, manifestos, and textbooks, mainstream CSS is developing a worldview congruent with the networked and interactional nature of the social world found in online platforms and digital data. This has brought leading scholars to characterize the social world as a complex system – expressing hopes of finding universal patterns and underlying mechanisms, to develop social science into not merely a form of physics but of engineering, by learning to control and tune its underlying social machinery.

Data in digital capitalism

The advent of digital data has resulted not only in the use of data for social scientific research, but also in a substantial

literature examining the political economy of digital data. This literature describes a period of capitalism in which data have become the defining commodity (e.g. Dean, 2005; Foster and McChesney, 2014; Fuchs, 2010; Sadowski, 2020; Srnicek, 2017; Zuboff, 2019). Although diverse, this literature shares some common conclusions: data collection has become an important motivation for businesses and governments (Zuboff, 2019); data are valuable and value-creating (Arvidsson, 2016; Srnicek, 2017); and data systems shot through with inequities and designed for extraction and exploitation (Andrejevic, 2014; Couldry and Mejias, 2020; Dalton et al., 2016).

Sadowski (2019) argues that data have become a form of capital, just like money and machinery, and that the production of data, similar to financialization before it, has become 'a new frontier of accumulation and next step in capitalism' (Sadowski, 2019: 9). Just as finance capitalism is characterized by the subordination of processes of production to financialization, so brings contemporary capitalism the subordination of production to data accumulation. Digital capitalism is taking shape as a political economic regime driven by the logic of accumulation, circulation, and manipulation of digital data.

Central to understanding the relationship between data and capitalism is that digital data is not 'naturally occurring', but actively extracted and inscribed in such a way as to become susceptible to evaluation, calculation, and intervention (Cukier and Mayer-Schoenberger, 2013). Couldry and Mejias (2020) compare the talk of data being 'just there' to historical colonialism and the legal doctrine of terra nullius: the idea that land such as the territory now known as Australia supposedly belonged to 'no one' and was 'for the taking' (Couldry and Mejias, 2020).

Seen in this light, the structure of digital data is the expression of a particular way of probing and representing the world's features and dynamics for the sake of manipulating and monetizing of human behavior (Sadowski, 2019). Digital data have to be extracted, in a process that reduces and abstracts, stripping context and including only certain aspects of the world. The process of generating data, thus, constitutes a way of exercising power over the world, by defining what counts as knowledge, who has access to it, and how it can be processed. Thatcher et al. (2016: 994) argue that these extractive practices 'mirror processes of primitive accumulation or accumulation by dispossession that occur as capitalism colonizes previously noncommodified, private times and places'. The value of data lies in their power to capture, predict, and control the social world, enabling every layer of the human experience to become the target of profitable extraction (Couldry and Mejias, 2020).

Platform data can, thus, not be considered traces of pre-existing social interaction, but should rather be seen as capturing the interactions that take place on the quantified playing social fields provided by platform infrastructures (Dalton et al., 2016). These platforms are sufficiently open-ended to allow social interaction and individual expression (Marres, 2017), but sufficiently controlled to structure and format social life in ways that render it amenable for large-scale monitoring, data analysis, and intervention (Couldry and Hepp, 2018). Although users are, to an extent, free to choose how they interact with these interfaces, they are not free to choose the context and conditions of this interaction; they cannot choose the menu of options from which they make their choice (Törnberg and Uitermark, 2020). This context is provided by the platforms, acting to pursue their own goals, such as extracting user data and maximizing platform engagement, through algorithms designed 'to nudge, coax, tune, and herd behavior toward profitable outcomes' (Zuboff, 2019: 8).

This logic has come to also shape the relationship between citizens and governments, as governments employ sophisticated methods for 'nudging' and directing users through nimble forms of control (Törnberg and Uitermark, 2020), for instance driving the emergence of forms of 'predictive policing' that use methods developed for analyzing consumer behavior to predict criminal behavior (Perry, 2013). More broadly, the extraction, distribution, and use of data are situated within an emerging political economy that has wide-ranging implications across society (Dalton et al., 2016; Sadowski, 2020): from cities (Ash et al., 2018; Leszczynski, 2015) and electric infrastructure (Levenda et al., 2015) to labor (Van Doorn, 2017) and media (Van Dijck et al., 2018).

In summary, the literature on digital capitalism suggests that we should understand data not as 'naturally occurring' or 'traces' of social processes, but as valuable commodities or a form of capital that are extracted and constructed in ways that reflect and perpetuate power inequities, while reshaping social processes into forms that best allow their analysis and manipulation through data analysis. This radical shift in perspective has epistemological and political implications for CSS.

CSS and digital capitalism

In epistemological terms, the digital capitalism literature suggests that the complexity and relationality of digital data is no less imposed or limiting, and no more the 'real' structure of the social world, than the old structure of 'rows of cases and columns of variables' (Lazer et al., 2020: 1060). This is not a question of 'bias' or 'artifacts', as CSS scholars might suggest, but better understood through Foucault's concept 'episteme' (Foucault, 2018): a way of imposing a certain structure on the world to make sense of it (Couldry and Mejias, 2020). When we study the structure of data, we are, thus, studying a structure imposed on social reality with the aim to produce data amenable to the same type of data analytics which CSS employs. Just as survey data are produced for scientific

inquiry, so are digital data shaped by certain models in such a way as to facilitate analysis, prediction, and control.

In political terms, if digital data are not traces of social reality, but the product of an abstraction created by power interests, constituting valuable commodities and means of production within contemporary capitalism, then building a science on its epistemic features appears not as bringing deeper understanding into the nature and structure of social reality, but as perpetuating and lending credence and providing methods to the current regime of accumulation (Couldry and Mejias, 2020; Van Dijck, 2014). This is not merely a question of 'ethics' that can be addressed through more detailed guidelines, but an issue of CSS being fundamentally implicated in digital capitalism, providing it tools and ideological backing. Rather than ethics guidelines, this begs for an answer to the question 'which side are you on?' (Byrne and Callaghan, 2013: 176). Leading CSS scholars such as Alex Pentland have given their unambiguous answers through manifestos that promote digital capitalist ideology, viewing society as an engineering problem, and suggesting a social science that supports supplanting political life with computation as the foundation of governance (Zuboff, 2019).

These epistemological and political issues are interlinked, as the episteme through which reality is perceived shapes it in turn. This is expressed in digital data and methods often reflecting and reinforcing racism, sexism, and other hierarchies of difference (e.g. Eubanks, 2018; Noble, 2018; O'neil, 2016). Examples include Google suggesting pornography for the search 'black girls' (Noble, 2018), targeting science and technology (STEM) job ads to men (Lambrecht and Tucker, 2019), showing advertisements for arrest records when searching for African-American sounding names (Sweeney, 2013), and algorithms systematically ranking women lower when analyzing resume (Dastin, 2018). Although corporations are addressing these sorts of issues through the notion of 'biases', the trouble runs deeper, as the algorithms are fundamentally founded on a logic of categorization and optimization (Finn, 2017). When algorithms are employed to optimize for the interests of corporations and governments by categorizing customers, employees, and citizens according to purchasing power, productivity, credit scores, health risks, and liability, then discrimination and inequities inevitably result.

A new neoclassicism

The role of CSS in relation to digital capitalism, thus, appears to be homologous to the relation between neoclassical economics and neoliberalism, in the sense of constituting a scientific paradigm which provides essential tools and legitimacy to a mode of capitalist accumulation. This is not to imply an equivalence between neoliberalism and digital capitalism. Instead, just as neoliberal elites adopted and instrumentalized certain ideas from neoclassical economics,

the elites of digital capitalism selectively adopt and instrumentalized the ideology of self-organization and complexity in CSS. And just as the 'marginal revolution' became the foundation not only of a scientific paradigm, but also of an ideology, so has the 'complexity revolution' provided the epistemology for an emerging form of capital accumulation: an epistemology where the social is fundamentally computational, suggesting that data have the potential to fully capture – and thereby commodify – the social world (Finn, 2017; Hayles, 2010). The particular relationships between methodology and epistemology of the scientific paradigms feed usefully the ideology and interests of the capitalist forms.

As noted above, this ideological and epistemological baggage of CSS has driven many social scientists to distance themselves from the field in favor of more critical approaches. These critical approaches focus less on data analytics, and more on how the procurement, processing, and mobilization of data emanates from and reshapes power relations, emphasizing methods such digital ethnography, political economy, discourse analysis, and the study of affordances and infrastructures. The digital methods (DM) approach, with an emphasis on what Marres and Gerlitz (2016) refer to as 'interface methods', has become in particular influential here, focused on examining the mutual shaping of sociality and media technology by re-purposing the 'social research methods [that] are already built into digital infrastructures, devices and practices, even if they currently tend to serve other-than-sociological ends' (Marres, 2017: 13), resulting in a more local and medium-specific understanding (Rogers, 2013). This constitutes an explicit distancing from the large-scale and quantitative 'backend'-oriented methods of CSS and its epistemological emphasis on underlying mechanisms and patterns.

Such a distancing is not without costs although. As Wyly (2009: 316) suggests through his 'strategic positivism', when we give up certain methods, we also 'give up the opportunity to shape and mobilize these constructions for progressive purposes'. This would imply instead attempting what Schuurman and Pratt (2002) refer to as an 'internal' critique: a critique that has a stake in the discipline, and which attempts to reshape it in a constructive manner. This follows Wyly's (2009: 310) suggestion for statistics that 'the presumed linkages between epistemology, methodology, and politics were never fundamental or immutable'. Rather than an outright rejection of CSS, we, therefore, propose a careful dissection to extract its problematic epistemic features, re-embedding its methods and metaphors in an alternative framework which would allow their mobilization for different scholarly and political aims.

A heterodox computational social science

We call for an HCSS: an attempt at reshaping CSS by re-embedding its methods in an alternative metatheoretical

framework. Drawing on the homology with neoclassical economics, we turn to the lessons learned in heterodox economics when it comes to reemploying methods in service for a different epistemic and ideological purpose. Lawson (2006) argues in an influential paper that the lynchpin that unites heterodox economics is a common foundation in critical realism. With its focus on the limits of naturalism, challenge of the quantitative-qualitative divide, emphasis of causal mechanisms, and conceptualization of social complexity, this metatheory provides a powerful means of challenging the linkages between epistemology, methodology, and politics within CSS, and reemploying its methods and metaphors for new purposes (Archer et al., 2013; Bhaskar, 2005; Danermark et al., 2001).

Centrally, critical realism allows the complexity which lies at the foundation of CSS to be constructively brought into a larger realist ontology, by drawing on an existing strand of research which integrates complexity and critical realism (Byrne and Callaghan, 2013; Harvey and Reed, 1996; Reed and Harvey, 1992). Reed and Harvey (1992) argue that complexity science provides a 'scientific ontology' consistent with a critical realist 'philosophical ontology', together forming a 'social ontology'. As Reed and Harvey (1992: 359) put it, such a 'complex realist' approach 'treats nature and society as if they were ontologically open and historically constituted; hierarchically structured, yet interactively complex; non-reductive and indeterminate, yet amenable to rational explanation', thus allowing us to 'steer a course midway between those positivists who would use chaos theory to revivify an exhausted scientism and those postmodernists who reject quantification on principle' (Harvey and Reed, 1996: 296). Drawing on research in critical realism, we briefly outline what such an HCSS entails by highlighting four of its qualities: *critical*, *methodological pluralist*, *interpretative*, and *explanative*.

Critical

Critical realism emphasizes that social science is an inseparable part of its own object of study, meaning that theory becomes a form of practice: we change the world by understanding it, and we understand it by changing it (Byrne, 2002; Danermark et al., 2001). In this paper, we have seen this play out in relation to CSS, and seen the costs of its neglect: when CSS are studying digital data, it studies data whose economic value and structure is shaped by the discipline's own models and theories. Critical realism suggests for us to acknowledge this and take responsibility for the consequences of our research, that is, to decide which side we are on, and what difference we want to make (Byrne and Callaghan, 2013). This suggests employing CSS as a powerful tool for alternative political purposes than those that it has today been made to serve, by bringing in emancipatory goals, and aiming for

justice and equality. As Wyly (2009) argues, even approaches born out of violence, colonial thought, and continuing oppression can provide valuable strategic bases for mobilization and organizing to challenge oppression. Data are necessarily built on abstractions, but when consciously employed, abstractions can be powerful tools for progressive aims (Rydin, 2007).

In contrast to Watts and Pentland's view of social science as a tool for 'tuning society' or 'solving problems', this calls for a social-scientific practice emphasizing creativity, conflicts, and negotiations. Rather than treating society as an engineering problem for which there can be single solutions (Andersson et al., 2014), the goal becomes to explore different possible pathways through simulation and experiment, both in silo and in empirical reality. This suggests thinking of computational analysis as a form of critique. The study of digital data and methods can be employed to critique the way algorithms reshape social life by embodying certain interests, and unveil the ideology embodied in digital data, rather than perpetuating or supporting this ideology (Noble, 2018). Central to such critical research is to go beyond merely describing patterns, to interrogate their limits, or criticize the structures that generate them (Danermark et al., 2001). As Carr (2014) puts it, 'A statistical model of society that ignores issues of class, that takes patterns of influence as givens rather than as historical contingencies, will tend to perpetuate existing social structures and dynamics'. This calls for constantly rejecting naturalization of social phenomena by revealing their contingency – that things could be otherwise. The resulting CSS finds its lineage not in physics, medicine, or engineering but in attempts of imagine and construe alternative futures in and through digitization (e.g. Medina, 2011; Pickering, 2010).

Methodological pluralist

As we have shown in this paper, CSS has tended to treat digital data and methods as granting unmediated access to the world as it 'really exists'. This is what Bhaskar (2013) refers to as an 'epistemic fallacy': confusing the abstract and the real by treating an abstraction as a description of concrete entity. Moving beyond this fallacy implies acknowledging that any modeling of a social system will involve not only a simplification *per se*, but necessarily also an ontological mismatch between model and system that cannot be resolved through any amount of empirical detail (Hayles, 2008). Critical realists have taken the methodological implications of this to be what Danermark et al. (2001) call 'critical methodological pluralism': the suggestion to move beyond methodological dichotomies – quantitative *versus* qualitative, positivism *versus* hermeneutics, universalism *versus* particularism – and often opting for combinations of methods. This mode of combining must, however, be based on ontological considerations, in

which we aim to identify the partial perspectives that the data and methods reveal (Zachariadis et al., 2013).

Since 'no data, big or small, can be interpreted without an understanding of the process that generated them' (Shaw, 2015: 1), rather than claiming neutrality, context is brought in by a combination between quantitative and qualitative perspectives, combining computational, ethnographic, and statistical approaches; 'quantitative data beg for qualitative interrogation' (Giglietto et al., 2012: 155). Instead of triangulation, where the purpose is validating a finding through different methods, methodological pluralism serves to bring into view different dimensions of social reality. Moving between different tools and perspectives, while remaining conscious of their limitations and biases, thus seeing a whole world by catching various 'glimpses of reality' (Byrne, 1998). Seen in this light, simulations, experiments, and models all provide valuable perspectives that can be further buttressed if they are complemented with other methods. We should not expect such an exercise to result in ever greater certainty about immutable regularities but instead view the methods as ways of bringing out different aspects of reality and establishing contingent patterns. An example is Hepp et al.'s (2016) combination of qualitative interviews with network analysis, aiming to bring out both the structure and meaning of communication networks across different platforms, showing differences among individuals and groups in how they establish and maintain relations.

A central critical realism strategy here is retroduction – a form of logical inference which charts a middle course between the 'data-driven' CSS and more traditional research. Retroduction places data in a central role, using guided knowledge discovery techniques to identify surprising or unanticipated observations worthy of examination and testing (see Kitchin, 2014: 6). This enables providing knowledge of transfactual conditions, structures, and mechanisms that cannot be directly observed in the domain of the empirical, thus moving from surface-level observations to deeper causal tendencies (Danermark et al., 2001). Retroduction as a strategy for logical inference fits well the pluralist approach suggested by, for example, Nelson (2020), which starts from methods of pattern-finding aimed to identify surprising observations and produce hypotheses, to then explain these observations via in-depth and qualitative analysis, and testing using quantitative methods. This situates the respective strengths of different methods within an approach that enables their application toward identifying deeper causal tendencies.

Interpretative

Many CSS scholars are either implicitly or, as in the case of Watts (e.g. 2011, 2014), explicitly critical to interpretation within social science, arguing that it tends to confirm preconceptions rather than contribute knowledge. Although Watts acknowledges that interpretation is occasionally helpful, he stresses that it should have no role in assessing the validity of theories (Watts, 2017; cf. Turco and Zuckerman, 2017). For critical realists, in contrast, interpretation and the identification of causal mechanisms are inseparable. Since causation in social systems occurs through symbols and meaning, causal analysis by definition involves interpretation (Collier, 1994). Interpretations are both a precondition and an aspect of the causality (Carter and New, 2004). Social life, as Paul Ricoeur wrote, has its very foundation in 'substituting signs for things' (Ricoeur, 1980: 219): that is, signs that embody interpretations (Couldry and Hepp, 2018). Social relations are inherently meaningful: nations, corporations, religions, or families do not exist independently of our interpretation of them.

This points to an approach that seeks an understanding of the actors' interpretations while recognizing that they are conditioned by history and embedded in context; mental states which bring about action can be complex, stratified, conflict-ridden, and more or less available to reflection. In short, in the social world, reasons can be causes (Byrne, 2002; Byrne and Callaghan, 2013). To disregard the role of meaning-making in the study of social life constitutes, as Hannah Arendt said of neoclassical economics, 'nothing less than the willful obliteration of their very subject matter' (Arendt, 2013: 57). As Couldry and Hepp (2018: 5) argue, 'whatever its appearance of complexity, even of opacity, the social world remains something accessible to interpretation and understanding by human actors, indeed a structure built up, in part, through those interpretations and understandings'.

This suggests an HCSS that employs computational methods to support, rather than supplant, interpretation. We believe that CSS includes methods that can be powerful for this purpose, employed as part of a larger interpretative framework, such as Nelson (2020) 'computational grounded theory'. Such an interpretative approach can also be found in cultural analysis, where scholars such as Christopher Bail (2014) use computational methods to access meaning in large datasets, not by 'measuring' meaning, but by supporting interpretation.

Explanative

Both CSS and critical realism aim to explain social phenomena by identifying the causal mechanisms that produce them (Danermark et al., 2001). However, there are important differences in how explanation is understood, leading to radically different frameworks for how to assess and employ CSS methods in its pursuit. The 'generative explanation' of CSS is singularly focused on explaining phenomena in terms of mechanisms in deeper strata (Epstein, 2006); 'Ultimately, the goal would be to find relevant quantities for describing societal and sociotechnical dynamics at a microscopic and macroscopic level and to

connect them, similar to the way thermodynamics works, going from the smallest scale—the individual—to the largest —society' (Caldarelli et al., 2018). From a critical realist perspective, such an approach provides important insights, but describes only one form of causality, while failing to account for the agency and meaning-making involved in emergence in the social world (Andersson and Törnberg, 2018) as well as the role of power and collective action (Uitermark, 2015).

Critical realism suggests that rather than viewing individuals as the foundation of social systems and structures as emergent, it is more fruitful to think of them as mutually implicated (Bourdieu, 1979; Byrne and Callaghan, 2013). Social mechanisms – including such widespread mechanisms as homophily or preferential attachment – are always shaped, facilitated, or activated by the context in which they occur, meaning that they are not the ground zero of social life but rather contingent and conditional (Uitermark and Van Meeteren, 2021). As Fuchs (2007: 27) puts it, 'the self-organization of society is not something that happens only blindly and unconsciously but depends on conscious, knowledgeable agents and creative social relationships'. Some agents and relationships are more powerful than others although and this is patently clear for digital environments that are purposefully created to condition users and establish patterns. Such interests and strategies should take center stage in the analysis of digital social life but they disappear from view when only generative explanation and upward causation is declared epistemologically acceptable. It is ironic that the computational social scientists, who are at the epicenters of digital power, embrace an ontology that blinds them to how such epicenters come about.

Critical realism suggests pursuing social explanation by tracing causal processes through the structures and elements involved, moving back and forth between upward and downward causation. Life in digital capitalism cannot be fully understood without accounting for all possible causal directions, as social emergence is in constant and mutual interaction with the structures that constrain and structure it. For instance, the practice of 'retweeting' on Twitter first emerged spontaneously and informally among users, was then discovered by the platform and incorporated into its design, which in turn brought about new emergent social dynamics as messages began spreading virally in the social network (Halavais, 2014). Similarly, social movements mobilize on social platforms against digital capitalism, as both themselves shaped by and aiming to reshape platform technologies (e.g. Milan, 2013). As these examples show, 'causality, in virtue of its transitivity, gives aid and comfort neither to the holist nor to the individualist. The causal chain just keeps rolling along' (Sober, 1980: 95).

There is, however, a place also for the more reductionist methodologies of CSS within a critical realist approach. For instance, Miller (2015: 188) calls for agent-based modeling to be embedded within critical realism, arguing for a 'creative, developmental, experimental, and iterative' practice of modeling that acknowledges that the conjecturing of mechanisms involves abduction. Through a critical realist lens, agent-based modeling appears more as computationally enhanced thought-experiments, allowing us to think through complex causality and emergence, rather than attempts at capturing social reality (Törnberg, 2018). Although CSS has focused primarily on individuals, these mechanisms play out at all levels of society. The mechanisms revealed through experiments and simulations are not the foundation of social life but express the workings of a particular, and inherently contingent and provisional, social context.

Example: social networks

We finally focus on a brief example to clarify the distinctiveness of HCSS through a comparison with orthodox CSS on one side, and DM on the other. We focus on the example of the network representations that are central in the backend of social media platforms, and expressed in interface patterns such as Twitter's following, mentioning, or retweeting.

CSS has seen an explosion of network research in recent years, driven by the growing availability of the network data employed by platforms. Most CSS research treats networks exclusively in graph theoretical terms (Fuhse, 2015). This approach has proven both powerful and parsimonious, enabling the application of sophisticated algorithms to identify and quantify structural patterns, but comes at the cost of the abstracting away what cannot be brought into the graph formulation, often implying the systematic disregard of cultural, interpretative, and intersubjective contexts. CSS employs the quantified output of these network algorithms to identify what are seen as universal mechanisms and patterns of human behavior: 'we want to be able to state principles that hold for all groups, all organizations, all societies' (Hanneman and Riddle, 2005: 196). Studying 'following' on Twitter, for instance, CSS may focus on questions such as the universal laws of social relationships, or the self-organized emergence of highly unequal power-law distribution of network influence (Gonçalves et al., 2011; Sadri et al., 2018).

In comparison, DM emphasizes the way technology actively participates in the enactment of relations, seeing following, mentioning, or retweeting as co-creating social relations (Marres, 2017: 140). This can be studied by repurposing of these platform functionalities toward research purposes (Marres, 2017: 15). DM tends to emphasize interpretative and qualitative aspects in these studies, employing numbers sparingly and primarily for illustrative purposes. In sharp contrast to the focus on universal mechanisms' characteristic of CSS, the result is a medium-specific and

local understanding, giving insights into the dynamics and culture of a given platform (Rogers, 2013). Focusing on Twitter, DM may focus on how Twitter's 'follow' and 'retweets' functionalities underpin specific cultures and vernaculars, playing part in shaping practices of influential content producers (Rogers, 2014; Schmidt, 2014).

The heterodox approach shares CSS' emphasis on causal mechanisms, complexity, and the obsolescence of the quantitative–qualitative divide, but emphasizes the role of culture and meaning, and views mechanisms as situational and contingent. If DM focuses on local and platform-specific culture, and CSS on universal network patterns, HCSS aims to leverage medium-specific and surface-level observations toward the critical examination of larger forces that structure mediatized social life – such as corporate power, capitalism, and racism (see e.g. Babic et al., 2020). CSS, thus, repurposes the 'backend' network methods, employed as one of the multiple methods to a retroductive approach (Buch-Hansen, 2014; Törnberg and Törnberg, 2019), aimed at critical examination of deeper mechanisms. Focusing on Twitter's following functions, this may consist of starting from CSS' identification of uneven distribution of network influence – but moving from this to asking what design elements produce this distribution, what social implications it has, and whose interests it serves. This acknowledges both CSS' emphasis on self-organized social structures, as well as DM's focus on the agency of technology – which together implicate the role of platforms in designing interface to produce this particular form of self-organization (Törnberg and Uitermark, 2020): for example, network representations are ubiquitous because their data are valuable for their capacity to identify consumer preferences, and micro-celebrities operate in mutualistic economic symbiosis with the platforms.

Conclusion

CSS has, over the past decade, emerged as one of the fastest-growing disciplines in academia, and the dominant field for the study of social behavior through digital data, located in the intersection of academia and industry. Viewing digital data from platforms as natural occurring by-product of digital social life, this field brings powerful new methods and approaches to bear by approaching social science as a form of data analytics. A dominant strand within CSS is coming to formulate an ontology commensurate with this approach, drawing on the networked and interactional nature of digital data to characterize social life as a complex system – that is, a pattern which emerges bottom-up from individual interactions. Through this complexity lens, social phenomena appear fundamentally computational in nature, making the quest for knowledge a quest for computation, and the social world may be the domain of the hard sciences.

However, the literature on digital capitalism suggests that data are less by-products of digital social life than the primary product which these platforms are geared to extract. Data are valuable commodities, and their complex and interactive structure has been imposed on the social world to make it amenable to analysis, prediction, and control – and therefore commodification – through precisely the data analytical tools that CSS applies and is part of developing. Through this lens, CSS appears as a paradigm which provides tools and scientific legitimacy to a mode of capitalist accumulation, while naturalizing and effacing conflict, power, and meaning-making from its subject matter. CSS is, in this sense, to digital capitalism what neoclassical economics is to neoliberalism.

These types of epistemological and ideological issues have brought critically oriented social scientific scholars to reject or distance themselves from the field, in favor for alternative means of studying digital social life. We have instead proposed that the problematic linkages between methodology, epistemology, and politics are not fundamental or immutable, but possible to challenge and revise. Rather than rejecting CSS and its approach, we have attempted to engage with the field in a constructive manner, to root out its problematic epistemic assumptions by re-embedding the field in an alternative meta-theoretical framework, thus allowing its methods and metaphors to be mobilized for different aims. Following the analogy with neoclassical economics, we have suggested an HCSS, founded in an ontology of critical realism. Through this ontological re-embedding, we argue that the concepts and methods of CSS can provide valuable insight, contributing to the aim of not only modeling society, but also to question it. This proposes a CSS in its briefest terms summarized by Wyly (2009: 317): 'Put simply, be careful, be modest, and be critical'.

ORCID iDs

Petter Törnberg https://orcid.org/0000-0001-8722-8646
Justus Uitermark https://orcid.org/0000-0002-5274-1455

References

Andersson C and Törnberg P (2018) Wickedness and the anatomy of complexity. *Futures* 95: 118–138. DOI: 10.1016/j.futures.2017.11.001.

Andersson C, Törnberg A and Törnberg P (2014) Societal systems – complex or worse? *Futures* 63: 145–157. DOI: 10.1016/j.futures.2014.07.003.

Andrejevic M (2014) Big Data, big questions | the Big Data divide. *International Journal of Communication* 8: 17.

Aral S and Nicolaides C (2017) Exercise contagion in a global social network. *Nature Communications* 8(1): 1–8.

Archer M, Bhaskar R, Collier A, et al. (2013) *Critical Realism: Essential Readings*. London: Routledge.

Arendt H (2013) *The Human Condition*. Chicago: University of Chicago Press.

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

Ash J, Kitchin R and Leszczynski A (2018) Digital turn, digital geographies? *Progress in Human Geography* 42(1): 25–43.

Babic M, Garcia-Bernardo J and Heemskerk EM (2020) The rise of transnational state capital: State-led foreign investment in the 21st century. *Review of International Political Economy* 27(3): 433–475.

Bail CA (2014) The cultural environment: Measuring culture with Big Data. *Theory and Society* 43(3-4): 465–482. http://www.jstor.org/stable/43694728

Ball P (2012) *Why Society is a Complex Matter: Meeting Twenty-First Century Challenges with a New Kind of Science*. New York: Springer Science & Business Media.

Barabási A-L and Albert R (1999) Emergence of scaling in random networks. *Science (New York, N Y)* 286(5439): 509–512.

Bauman Z (2013) *Liquid Modernity*. Cambridge: John Wiley & Sons.

Bhaskar R (2005) *The Possibility of Naturalism, A Philosophical Critique of the Contemporary Human Sciences*. London: Routledge. DOI: 10.1111/j.1468-0149.1981.tb02719.x.

Bhaskar R (2013) *A Realist Theory of Science*. London: Routledge.

Bourdieu P (1979) *Distinction: A Social Critique of the Judgement of Taste*. London: Routledge.

Buch-Hansen H (2014) Social network analysis and critical realism. *Journal for the Theory of Social Behaviour* 44(3): 306–325.

Byrne D (2002) *Interpreting Quantitative Data*. New York: Sage.

Byrne D and Callaghan G (2013) *Complexity Theory and the Social Sciences: The State of the Art*. London: Routledge.

Byrne DS (1998) *Complexity Theory and the Social Sciences: An Introduction*. New York: Psychology Press.

Caldarelli G, Wolf S and Moreno Y (2018) Physics of humans, physics for society. *Nature Physics* 14(9): 870.

Carr N (2014) The limits of social engineering. *MIT Technology Review*. Available from: https://www.technologyreview.com/2014/04/16/173156/the-limits-of-social-engineering/

Carter B and New C (2004) *Making Realism Work*. London and New York: Routledge, 5. DOI: 10.4324/9780203624289. (May 24, 2016). Available at SSRN: http://dx.doi.org/10.2139/ssrn.2784278

Centola D (2010) The spread of behavior in an online social network experiment. *Science (New York, N Y)* 329(5996): 1194–1197.

Christakis NA (2012) A new kind of social science for the 21st century. *Edge*. Available at: http://edge.org/conversation/a-21st-century-change-to-social-science (accessed August 2020).

Cilliers P (2002) *Complexity and Postmodernism: Understanding Complex Systems*. London: Routledge.

Cioffi-Revilla C (2016) *Bigger Computational Social Science: Data, Theories, Models, and Simulations–Not Just Big Data*. (May 24, 2016). Available at SSRN: http://dx.doi.org/10.2139/ssrn.2784278

Cioffi-Revilla C (2017) Computation and social science. In: *Introduction to Computational Social Science*. New York: Springer, pp. 35–102.

Collier A (1994) *Critical Realism: An Introduction to Roy Bhaskar's Philosophy*. London: Verso.

Conte R, Gilbert N, Bonelli G, et al. (2012) Manifesto of computational social science. *European Physical Journal-Special Topics* 214(214): 325–346.

Conway J (1970) The game of life. *Scientific American* 223(4): 4.

Couldry N and Hepp A (2018) *The Mediated Construction of Reality*. Cambridge: John Wiley & Sons.

Couldry N and Mejias UA (2020) *The Costs of Connection: How Data are Colonizing Human Life and Appropriating It for Capitalism*. Oxford: Oxford University Press.

Cukier K and Mayer-Schoenberger V (2013) The rise of Big Data: How it's changing the way we think about the world. *Foreign Affairs (Council on Foreign Relations)* 92: 28.

Cvitanovic P, Artuso R, Mainieri R, et al. (2005) Chaos: classical and quantum. *ChaosBook.org (Niels Bohr Institute, Copenhagen, 2005)*.

Dalton CM, Taylor L and Thatcher J (2016) Critical data studies: A dialog on data and space. *Big Data and Society* 3(1): 1–9. DOI: 10.1177/2053951716648346.

Danermark B, Ekstrom M, Jakobsen L, et al. (2001) *Explaining Society: An Introduction to Critical Realism in the Social Sciences*. London: Routledge.

Dastin J (2018) Amazon scraps secret AI recruiting tool that showed bias against women. *San Francisco, CA: Reuters. Retrieved on October* 9: 2018.

Dean J (2005) Communicative capitalism: Circulation and the foreclosure of politics. *Cultural Politics* 1(1): 51–74.

Edelmann A, Wolff T, Montagne D, et al. (2020) Computational social science and sociology. *Annual Review of Sociology* 46: 61–81.

Edwards A, Housley W, Matthew W, et al. (2013) Digital social research, social media and the sociological imagination: Surrogacy, augmentation and re-orientation. *International Journal of Social Research Methodology* 16(3): 245–260.

Epstein JM (2006) *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton: Princeton University Press.

Eubanks V (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St Martin's Press.

Finn E (2017) *What Algorithms Want: Imagination in the Age of Computing*. New York: MIT Press.

Foster JB and McChesney RW (2014) Surveillance capitalism: Monopoly-finance capital, the military-industrial complex, and the digital age. *Monthly Review* 66(3): 1.

Foucault M (2018) *The Order of Things*. London: Routledge.

Fuchs C (2007) *Internet and Society: Social Theory in the Information Age*. London: Routledge.

Fuchs C (2010) Labor in informational capitalism and on the internet. *The Information Society* 26(3): 179–196.

Fuhse JA (2015) Theorizing social networks: The relational sociology of and around Harrison White. *International Review of Sociology* 25(1): 15–44.

Galison P (1997) *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.

Giglietto F, Rossi L and Bennato D (2012) The open laboratory: Limits and possibilities of using Facebook, Twitter, and YouTube as a research data source. *Journal of Technology in Human Services* 30(3–4): 145–159.

Gonçalves B, Perra N and Vespignani A (2011) Modeling users' activity on twitter networks: Validation of Dunbar's number. *PLoS One* 6(8): 1–5.

Halavais A (2014) Structure of Twitter: Social and technical. In Weller K, Bruns A, Burgess J and Mahrt M (eds) *Twitter and Society*. New York, NY: Peter Lang, pp. 29–42.

Hanneman RA and Riddle M (2005) *Introduction to Social Network Methods*. Los Angeles: University of California Riverside.

Harvey DL and Reed M (1996) Social science as the study of complex systems. In Kiel LD and Elliott E (eds) *Chaos Theory in the Social Sciences*. Detroit, Ann Arbor, MI: University of Michigan Press, pp. 295–324.

Hayles NK (2008) *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Riverside, CA: University of Chicago Press.

Hayles NK (2010) *My Mother was a Computer*. Chicago: University of Chicago Press.

Hepp A, Roitsch C and Berg M (2016) Investigating communication networks contextually: Qualitative network analysis as cross-media research. *MedieKultur: Journal of Media and Communication Research* 32(60): 20.

Hirschberg J and Manning CD (2015) Advances in natural language processing. *Science (New York, N Y)* 349(6245): 261–266.

Huang N, Chen P, Hong Y, et al. (2018) Digital nudging for online social sharing: Evidence from a randomized field experiment. In: Proceedings of the 51st Hawaii international conference on system sciences, 2018.

Iggers GG (1959) Further remarks about early uses of the term 'social science'. *Journal of the History of Ideas, JSTOR* 20(3): 433–436. DOI: 10.2307/2708121.

Kitchin R (2014) Big data, new epistemologies and paradigm shifts. *Big Data and Society* 1(1): 1–12.

Kramer ADI, Guillory JE and Hancock JT (2014) Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111(24): 8788–8790.

Lambrecht A and Tucker C (2019) Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of stem career ads. *Management Science* 65(7): 2966–2981.

Langton CG (1997) *Artificial Life: An Overview*. Boston: MIT Press.

Lawson T (2006) The nature of heterodox economics. *Cambridge Journal of Economics* 30(4): 483–505.

Lazer D, Pentland AS, Adamic L, et al. (2009) Life in the network: The coming age of computational social science. *Science (New York, NY)* 323(5915): 721.

Lazer D, Pentland A, Watts DJ, et al. (2020) Computational social science: Obstacles and opportunities. *Science (New York, NY)* 369(6507): 1060–1062.

Legewie J (2016) Racial profiling and use of force in police stops: How local events trigger periods of increased discrimination. *American Journal of Sociology* 122(2): 379–424.

Leszczynski A (2015) Spatial media/tion. *Progress in Human Geography* 39(6): 729–751.

Levenda A, Mahmoudi D and Sussman G (2015) The neoliberal politics of 'smart': Electricity consumption, household monitoring, and the enterprise form. *Canadian Journal of Communication* 40(4): 615–636.

Lohr S (2015) *Data-ism: Inside the Big Data Revolution*. New York: Simon and Schuster.

Manovich L (2011) Trending: The promises and the challenges of big social data. *Debates in the Digital Humanities* 2(1): 460–475.

Manovich L (2016) *Online Post*. Available at: https://www.facebook.com/lev.manovich/posts/10155536491382316 (accessed 2021-09-16).

Marres N (2017) *Digital Sociology: The Reinvention of Social Research*. New York: John Wiley & Sons.

Marres N and Gerlitz C (2016) Interface methods: Renegotiating relations between digital social research, STS and sociology. *The Sociological Review* 64(1): 21–46.

Mas A and Moretti E (2009) Peers at work. *American Economic Review* 99(1): 112–145.

Medina E (2011) *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile*. Boston: MIT Press.

Milan S (2013) *Social Movements and Their Technologies: Wiring Social Change*. New York: Springer.

Miller KD (2015) Agent-based modeling and organization studies: A critical realist perspective. *Organization Studies* 36(2): 175–196.

Mitchell M (2009) *Complexity: A Guided Tour*. Oxford: Oxford University Press.

Molina M and Garip F (2019) Machine learning for sociology. *Annual Review of Sociology* 45: 27–45. DOI: 10.1146/annurev-soc-073117-041106.

Nelson LK (2020) Computational grounded theory: A methodological framework. *Sociological Methods and Research* 49(1): 3–42.

Noble SU (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.

O'Neil C (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Broadway Books.

Pearl J (2019) The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM* 62(3): 54–60.

Pentland A (2015) *Social Physics: How Social Networks Can Make Us Smarter*. New York: Penguin.

Perry WL (2013) *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Santa Monica, CA: Rand Corporation. https://www.rand.org/pubs/research_reports/RR233.html

Pickering A (2010) *The Cybernetic Brain: Sketches of Another Future*. Chicago: University of Chicago Press.

Ricoeur P (1980) Narrative time. *Critical Inquiry* 7(1): 169–190.

Reed M and Harvey DL (1992) The new science and the old: Complexity and realism in the social sciences. *Journal for the Theory of Social Behaviour* 22(4): 353–380.

Rogers R (2013) *Digital Methods*. Cambridge, MA: MIT Press.

Rogers R (2014) Debanalising Twitter. *Twitter and Society, New York, NY*, ix–xxxviii.

Rydin Y (2007) Indicators as a governmental technology? The lessons of community-based sustainability indicator projects. *Environment and Planning D: Society and Space* 25(4): 610–624.

Sadowski J (2019) When data is capital: Datafication, accumulation, and extraction. *Big Data and Society* 6(1): 1–12.

Sadowski J (2020) *Too Smart: How Digital Capitalism is Extracting Data, Controlling Our Lives, and Taking over the World*. Boston: MIT Press.

Sadri AM, Hasan S, Ukkusuri SV, et al. (2018) Analysis of social interaction network properties and growth on Twitter. *Social Network Analysis and Mining* 8(1): 1–13.

Savage M and Burrows R (2007) The coming crisis of empirical sociology. *Sociology* 41(5): 885–899.

Schmidt J-H (2014) Twitter and the rise of personal publics. *Twitter and Society*. New York, NY: Peter Lang. 3–14.

Schuurman N and Pratt G (2002) Care of the subject: Feminism and critiques of GIS. *Gender, Place and Culture: A Journal of Feminist Geography* 9(3): 291–299.

Shaw R (2015) Big Data and reality. *Big Data and Society* 2(2): 1–4. DOI: 10.1177/2053951715608877.

Sober E (1980) Holism, individualism, and the units of selection. In: PSA: Proceedings of the Biennial meeting of the philosophy of science association, pp. 93–121.

Squazzoni F (2012) *Agent-Based Computational Sociology*. New York: John Wiley & Sons.

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

Sweeney L (2013) Discrimination in online ad delivery. *Communications of the ACM* 56(5): 44–54.

Thatcher J, O'Sullivsan D and Mahmoudi D (2016) Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environment and Planning D: Society and Space* 34(6): 990–1006.

Törnberg A (2018) Abstractions on steroids: A critical realist approach to computer simulations. *Journal for the Theory of Social Behaviour* 49(1): 12194. DOI: 10.1111/jtsb.12194.

Törnberg P and Törnberg A (2019) Minding the gap between culture and connectivity: Laying the foundations for a relational mixed methods social network analysis. In Froehlich DE, Rehm M and Rienties BC (eds) *Mixed Methods Approaches to Social Network Analysis*. (pp. 58–71). Routledge.

Törnberg P and Uitermark J (2020) Complex control and the governmentality of digital platforms. *Frontiers of Sustainable Cities* 2: 6.

Turco CJ and Zuckerman EW (2017) Verstehen for sociology: comment on Watts. *American Journal of Sociology* 122(4): 1272–1291.

Uitermark J (2015) Longing for Wikitopia: The study and politics of self-organisation. *Urban Studies* 52(13): 2301–2312.

Uitermark J and van Meeteren M (2021) Geographical network analysis. *Tijdschrift voor economische en sociale geografie* 112(4): 337–350.

Van de Rijt A, Kang SM, Restivo M and Patil A (2014) Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences* 111(19): 6934-6939.

Van den Berg J (2017) Duncan Watts: "sociale wetenschappers, stop met navelstaren." *Sociale Vraagstukken*. On-line: https://www.socialevraagstukken.nl/interview/duncan-watts-sociale-wetenschappers-stop-met-navelstaren/ (accessed 2021-09-17).

Van Dijck J (2014) Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance and Society* 12(2): 197–208.

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

Van Doorn N (2017) Platform labor: On the gendered and racialized exploitation of low-income service work in the 'on-demand' economy. *Information, Communication and Society* 20(6): 898–914.

Wahde M (2008) *Biologically Inspired Optimization Methods: An Introduction*. Southampton, UK: WIT press.

Watts DJ (2004) The 'new' science of networks. *Annual Review of Sociology* 30: 243–270.

Watts DJ (2011) *Everything is Obvious: Why Common Sense is Nonsense*. New York: Atlantic Books.

Watts DJ (2013) Computational social science: Exciting progress and future directions. *The Bridge on Frontiers of Engineering* 43(4): 5–10.

Watts DJ (2014) Common sense and sociological explanations. *American Journal of Sociology* 120(2): 313–351.

Watts DJ (2017) Should social science be more solution-oriented? *Nature Human Behaviour* 1(1): 1–5.

West G (2017) *Scale: The Universal Laws of Life and Death in Organisms*. New York: Weidenfeld & Nicolson.

Wyly E (2009) Strategic positivism. *The Professional Geographer* 61(3): 310–322.

Zachariadis M, Scott S and Barrett M (2013) Methodological implications of critical realism for mixed-methods research. *MIS Quarterly* 37(3): 855–879.

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