

JOURNAL OF COMPARATIVE RESEARCH IN
ANTHROPOLOGY AND SOCIOLOGY

Copyright © The Author, 2019
Volume 10, Number 1, Summer 2019
ISSN 2068 – 0317
<http://compaso.eu>



Sociological perspectives in critical data studies: A literature review

Alexandra Hosszu¹

Abstract

This research note reviews the main sociological directions in Critical Data Studies (CDT) articulated as opposed to “digital positivism”. In the first part of the review, I will highlight the relevance of the critical data studies in the contemporary digitalized world and I will explain the meaning of the interdisciplinary field. In the second part, I plan to analyze two relevant concepts for Critical Data Studies: (1) power and (2) surveillance. I will also consider the implications for the new ways of doing Sociology when using big data. Finally, I will discuss the identified theoretical gaps and the need for empirical studies based on the analyzed concepts.

Keywords

Critical data studies, big data, algorithms, digital sociology, surveillance, power and authority

The days we live in are fully shaped by digital technologies, either if we use them or not. One of the most intriguing aspects of this digitalized era is that even though one does not have access or decides not to use these technologies, the digital still influences that one, as member of the digitalized society.

In 2018, 55% of the world population (Internet World Stats, 2018) has access to internet (more than 4 billion people around the world), 68% are unique mobile users and 42% are active social media users (We are social, 2018). In Europe, 67% of the population believe the technologies have a positive impact on the quality of their life and 64% on the society in general. Still, 18% disagree with the first statement and 25% disagree with the second one (Eurobarometer, 2017).

¹ Faculty of Sociology and Social Work, University of Bucharest, Romania, alexandra.hosszu1@gmail.com

The mass use of internet has led to the development of big data, large volume of data, usually characterized by 3 Vs: volume, variety and velocity. Scholars from different fields have lately analyzed the impact of big data in the society. In the following research note, I will analyze the main critical perspectives on big data and its implications focusing on sociological perspectives, but not exclusively.

Lupton (2013: p.106) analyzes the metaphors used to describe big data; most of them associate data with water – “streams, flows, leaks, rivers, oceans, seas, waves, fire hoses and even floods, deluges and tsunamis” emphasizing the quantitative side of data and also its uncertainty and fluctuation. Another set of metaphors portray data as remaining parts or reflections of human activities: trails, breadcrumbs, exhausts, smoke signals, shadows. Furthermore, big data has been depicted as living creation through metaphors like “rhizome”, “digital nervous system” or “digital exoskeleton” (Lupton, 2013: p.108).

Big data were understood through their quantitative data as huge amount of data; such perspective was labelled by Mosco (2014) as “digital positivism”: “data, suitably circumscribed by quantity, correlation, and algorithm, will, in fact, speak to us” (2014: p. 206).

In response to the “digital positivism”, scholars have developed the so-called Critical Data Studies, an interdisciplinary field focusing on the way Big Data shapes the social structures, interactions and institutions and creates asymmetrical power relationships. This field was initially emerged from geography researchers who concentrated their studies on “geospatial web” (Richterich, 2018: p.11) and continued by scholars from Sociology and other Social Sciences.

In Sociology, Lupton (2013) analyses the development of digital sociology, as sub-discipline. She identifies four categories when defining the sub-discipline: (1) professional digital practice – using technologies as sociological tools (networks, publishing, sharing research); (2) analyses of digital technology use – how to technologies shape the self, the social structures and relations; (3) digital data analysis – using Big Data in sociological inquiries and (4) critical digital sociology – developing reflexive analysis of digital technologies.

Critical Data Studies are extremely relevant for the digital society we live in, as the field questions the accuracy, objectivity, neutrality and unbiased characteristics of big data that are claimed by the positivists.

The research note includes the analysis of (1) **big data, algorithms and power**, (2) **surveillance and sousveillance concepts** and (3) **challenges for doing sociology in a digital context**. All three subthemes are significant for the Critical Data Studies field.

Big data, algorithms and power

There are several questions raised by the big data phenomenon: How does big data impact our lives? What decisions are made on behalf of big data? Who wins and who losses from the use of big data?

Big data were initially presented as objective, neutral and representative for masses even though we live in a subjective, unjust and stratified world. Big data are the product of our world; it would be hard to believe these data are perfect in an imperfect world. “Raw data” do not exist, this phrase is an “oxymoron” according to Helmond (2013). Thus, as Lupton (2013: p.101) believes, big data are “political, social and cultural processes”, they do not have intrinsic value, data are a human product that might be produced, archived and used in corrupt ways.

Boyd and Crawford (2012) claim one of the questions we should ask is who has access to big data? The access to big data is uneven and privileges few individuals who control their use; thus, a new “digital divide” is created between big data rich and big data poor. Manovich (2011) identifies three categories of actors relevant to big data debates: the ones who produce these data by leaving “digital footprints” (regular individuals using digital platforms, apps, devices), the ones who collect and store data (usually corporations owning the “means of production”) and the ones who have access and skill to analyze data and produce knowledge based on it. He labels these three categories as “new data-classes of our big data society” (2011).

Richterick (2018) brings into the debate the problem of *data monopolies* referring to the small group who influence or control the access and use of big data. The relationship between the big group who produces data and the small group who controls it is clearly asymmetrical and it reflects the general relationship in society between the masses and the ones having power. The asymmetry per se wouldn't be intrinsically problematic, it becomes questionable in the moment the asymmetry generates injustice.

Aggregated data are used to develop algorithms that could simplify and improve the human lives. Unfortunately, even though we created algorithms to reduce the human subjectivity and biases, the algorithms are the digital face of the society because they are the product of the same system. One of the existing assumptions is that if the level of prejudices and stereotypes existing in a society is high, the algorithms produced in the same society will be biased. If the level of biases is high, then algorithms will reinforce existing social inequalities. Lupton (2013: p.102) believes algorithms have “their own agential power” because they make predictions and create new data based on the old ones. Cheney-Lippold (2011) explains that algorithms have generated “new algorithmic identities”, they establish human characteristics such as gender, level of education, ethnicity, religion according to mathematic models. Thus, these characteristics are framed based on the previous aggregated data through simplifications and labels. It is quite odd that we fought for reducing people's natural predisposition for judging based on models and we promoted the need for analyzing each case and, now, algorithms are built to create models and to analyze at aggregate level.

Lupton (2013: p.106) concludes we don't have the weapons to fight or resist to algorithm biases and asymmetric power because we don't understand them.

Surveillance and sousveillance

If (1) you have a cell phone, a credit card or an ID card and (2) you use Google, Gmail, Facebook, go to school, have a job or drive a car you are “under surveillance” (Gilliom and Monahan, 2013). We live in a surveillance society. The surveillance idea was developed and analyzed from a long time, but nowadays we face a new social surveillance meaning “scrutiny through the use of technical means to extract or create personal or group data, whether from individuals or contexts” (Gary T. Marx, 2007). The surveillance theory building was structured into 3 clusters by Galič et. al. (2017):

1. Bentham’s initial perspective on Panopticon architecture and Foucault’s further development referring to the control and power the institutions apply to the individual. David Lyon builds up his theory of surveillance using Foucault’s work.
2. Surveillance post-panoptical theories are based on Deleuze’s work
3. Contemporary theories of surveillance use the previous concepts and theories and focus on digital surveillance.

In the third cluster, several contemporary sociologists could be identified. One of them is David Lyon (2006) who developed the surveillance studies, his work has contributed to theorization of the concept and a better understanding. His theoretical approach on digital surveillance maps the concepts used by Marx (capitalism), Weber (rationalization, bureaucratization) or Simmel (urban life and metropolis).

Lyon together with Bauman (2010) also developed the concept of “liquid surveillance”, building on Bauman’s well known concept of “liquid modernity”; the new surveillance is liquid because we don’t know any more who is monitoring us and when, or how they use it: “reduction of the body to data and the creation of data-doubles on which life-chances and choices hang more significantly than on our real lives and the stories we tell about them” (Lyon, 2010).

Haggerty and Ericson (2000) explored the concept of “surveillant assemblage”, derived from Deleuze and Guattari’s work who distinguished between assemblages and flows. Thus, individuals who before didn’t represent the focus of routine surveillance, are now part of a surveillant assemble who uses abstractions of human bodies.

Crawford (2014) captures the “anxieties of surveillance” – “the fear that all the data we are shedding every day is too revealing of our intimate selves but may also misrepresent us”. Such anxieties are not only felt by the ones surveilled, they also belong to the surveillers even though they have an asymmetrical power over the first group. Leszczynski (2015) disagrees with Crawford regarding the individuals’ anxieties towards big data, rather he believes the lived experiences of big data generate “anxieties of control” where individuals have anxieties due to their lack of control on the production, administration and management of their data.

In response to the surveillance, an antonymic concept was developed – sousveillance, meaning “watching from below” (opposed to “watching from above”). Sousveillance promotes the idea of involving regular citizens in surveilling the authority in order to achieve a democratization of surveillance (Lupton, 2013: pp.36-37).

Sociology in a digital society

Many sociologists are analyzing how sociology and social research could answer to the “digital society” by rethinking methods and ways of interpretation. As part of critical data studies, sociologists are critically analyzing the “status of contemporary sociology” (Lupton, 2013: p.46). Back and Puwar were articulating the idea of “live sociology” by developing new ways of doing sociology that are “public and critical” (Lupton, 2013: p.46) through new methods and new research devices. Beer (2014) was re-imagining a “punk sociology”, inspired from the punk ethos, because the contemporary sociology has a “moment of uncertainty”. Punk sociology would be part of the answer to the question: how can sociology respond to a changing world and the particular analytical and structural challenges it brings?” (2014: p.68).

“The coming crisis of empirical sociology” was pointed out by Savage and Burrows (2007) without referring directly to the big data phenomenon, but to the digitalized knowledge in general. At that time the two sociologists, highlighted the need for reframing the ways of doing research in order to respond to the digitalization challenges. Later on, in 2014, Savage and Burrows wrote some reflexive and critical thought on big data and methodological constraints, concluding that “big data does challenge the predominant authority of sociologists and social scientists more generally to define the nature of social knowledge”.

Boyd and Crawford (2012) consider big data could give the impression of improvement in social sciences research by addressing more quantitative and objective methods. In practice, some of the available data might be “unreliable” and no matter the size of the data, its analysis “it is subject to limitation and bias” (Boyd and Crawford, 2012: p.668).

Ruppert et al. (2013) indicate the need for rethinking the theoretical assumptions of research methods, they identify 9 assumptions related to: transactional actors, heterogeneity, visualization, continuous temporality, whole populations, granularity, expertise, mobile and mobilising and non-coherence. They also associate this perspective with the “field analysis” concept from Bourdieu, the digital devices do not have intrinsic capacities, but they are complementary one with another and they become indispensable (Ruppert et al., 2013: p.15).

Sociologists have been divided by the big data phenomenon: at one spectrum, we can find the enthusiastic ones enjoying the challenges and reframes imposed by digitalization and, on the other side, we can identify the skeptical ones who are worried about the side effects that big data will bring not only to the society in general, but to sociology more specific.

The critical analysis from both sides are relevant for reforming the discipline ensuring sociology will remain part of the digital society.

Conclusions

“The era of Big Data has only just begun, but it is already important that we start questioning the assumptions, values, and biases of this new wave of research. As scholars who are invested in the production of knowledge, such interrogations are an essential component of what we do” (Boyd and Crawford, 2012: p.675).

Scholars from different fields have lately analyzed the impact of big data in the society. Sociologists have mainly considered the ways social research should adapt to the digital context in order to better responses and understandings of the contemporary world.

When it comes to understanding the outcomes and changes that big data determine in the digital world, the sociologists work is still limited. Only few concepts have been studied through sociological lens and the Critical Data Studies field needs more sociological perspectives. These concepts have to also be applied in empirical studies that will offer comprehensive and valuable insights regarding big data and digital society.

In this research note I have analyzed the existing sociological perspectives on (1) power, authority and asymmetric relation between the big group who produces data and the small group who controls it; (2) digital surveillance and (3) new way of doing sociology in order to address the challenges brought by the digitalized contemporary world.

Further sociological conceptualization of big data’s layers and components is required in the future for a broader understanding of the phenomenon.

REFERENCES

- Beer, David. 2014. *Punk Sociology*. Palgrave Macmillan.
<https://doi.org/10.1057/9781137371218.0001>.
- Boyd, Danah, and Kate Crawford. 2012. “Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.” *Information Communication and Society* 15 (5): 662–79. <https://doi.org/10.1080/1369118X.2012.678878>.
- Burrows, Roger, and Mike Savage. 2014. “After the Crisis? Big Data and the Methodological Challenges of Empirical Sociology.” *Big Data & Society* 1 (1): 205395171454028. <https://doi.org/10.1177/2053951714540280>.
- Cheney-Lippold, John. 2011. “A New Algorithmic Identity” 28 (6): 164–81. <https://doi.org/10.1177/0263276411424420>.
- Crawford, Kate. 2014. “The Anxieties of Big Data.” *The New Inquiry*. <https://thenewinquiry.com/the-anxieties-of-big-data/>.
- Eurobarometer. 2017. *Attitudes towards the Impact of Digitisation and Automation on Daily Life*. <https://doi.org/10.2759/835661>.

- Galič, Maša, Tjerk Timan, and Bert Jaap Koops. 2017. "Bentham, Deleuze and Beyond: An Overview of Surveillance Theories from the Panopticon to Participation." *Philosophy and Technology* 30 (1): 9–37. <https://doi.org/10.1007/s13347-016-0219-1>.
- Gilliom, John and Torin Monahan. 2013. *SuperVision. An Introduction to Surveillant Society*. The University of Chicago Press, Chicago 60637.
- Haggerty, D., Richard V. Ericson, Kevin. 2000. "The Surveillant Assemblage." *British Journal of Sociology* 51 (4): 605–22. <https://doi.org/10.1080/00071310020015280>.
- Helmond, Anne. 2014. 'Raw Data' Is an Oxymoron. *Information, Communication & Society*. Vol. 17. <https://doi.org/10.1080/1369118X.2014.920042>.
- Internet World Stats. 2018. [Online] Available from: <https://www.internetworldstats.com/stats.htm> [Accessed 27th December 2018].
- Leszczynski, Agnieszka. 2015. "Spatial Big Data and Anxieties of Control." *Environment and Planning D: Society and Space* 33 (6): 965–84. <https://doi.org/10.1177/0263775815595814>.
- Lyon, David. (2006). *Theorizing surveillance: the panopticon and beyond*. Cullompton, Devon, Willan Pub.
- Lyon, David. 2010. "Liquid Surveillance: The Contribution of Zygmunt Bauman to Surveillance Studies." *International Political Sociology* 4 (4): 325–38. <https://doi.org/10.1111/j.1749-5687.2010.00109.x>.
- Marx, Gary T. 2007. "Hey Buddy Can You Spare a DNA? New Surveillance Technologies and the Growth of Mandatory Volunteerism in Collecting Personal Information." *Annali Dell'Istituto Superiore Di Sanita* 43 (1): 12–19. <https://doi.org/17536149>.
- Mosco, Vincent. 2014. *To the cloud: big data in a turbulent world*. Paradigm Publishers.
- Richterich, Annika. 2018. *The Big Data Agenda: Data Ethics and Critical Data Studies*. <https://doi.org/10.16997/book14>.
- Ruppert, Evelyn, John Law, and Mike Savage. 2013. "Reassembling Social Science Methods: The Challenge of Digital Devices." *Theory, Culture & Society* 30 (4): 22–46. <https://doi.org/10.1177/0263276413484941>.
- We are social. 2018. *Global digital report*. [Online] Available from: <https://wearesocial.com/blog/2018/01/global-digital-report-2018> [Accessed 27th December 2018].

Alexandra Hosszu has Bachelor and Master's Degrees in Political Science, both from National School of Administrative and Political Studies from Bucharest. Alexandra is currently a PhD student at Doctoral School of Sociology from University of Bucharest. Alexandra has been working with the Romanian NGO sector for more than 5 years doing research, project writing and management, monitoring and evaluation. At the moment, she is interested in understanding how digital technologies shape the social interactions and how they create knowledge.