

# Philosophical Aspects of Big Data

Nicolae Sfetcu

19.08.2019

Sfetcu, Nicolae, "Philosophical Aspects of Big Data", SetThings (19 august 2019), URL = <https://www.setthings.com/en/philosophical-aspects-of-big-data/>

Email: [nicolae@sfetcu.com](mailto:nicolae@sfetcu.com)



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

A partial translation of:

Sfetcu, Nicolae, "Etica Big Data în cercetare", SetThings (6 iulie 2019), DOI: 10.13140/RG.2.2.27629.33761, MultiMedia Publishing (ed.), ISBN: 978-606-033-228-2, URL = <https://www.setthings.com/ro/e-books/etica-big-data-in-cercetare/>

Big Data can generate, through inferences, new knowledge and perspectives. The paradigm that results from using Big Data creates new opportunities.

One of the major concerns in the Big Data case is that data scientists tend to work with data on topics they do not know and have never been in contact with, being alienated from the final product of their activity (application of analyzes). A recent study (Tanner 2014) states that this may be the reason for a phenomenon known as digital alienation.

Big Data has great influence at the governmental level, positively affecting society. These systems can be made more efficient by applying transparency and open governance policies, such as Open Data.

After developing predictive models for target audience behavior, Big Data can be used to generate early warnings for various situations. There is thus a positive feedback between research and practice, with rapid discoveries taken from practice.

A. Richterich, in "Examining (Big) Data Practices and Ethics", states that the popularization of user activity monitoring was motivated by claims that using (and collecting data with) these devices would improve users' well-being, health and life expectancy, and significantly reduce healthcare costs. (Richterich, 2018) To obtain user consent, many companies offered discounts to those customers who would be willing to provide access to their monitoring data. (Mearian 2015) But there are also concerns about the influence of these technologies on society, especially in issues related to fairness, discrimination, privacy, data abuse and security. (Collins 2016)

Conceptually, Big Data should be understood as an umbrella term for a set of emerging technologies. In their use, we must take into account the cultural, social and technological contexts, networks, infrastructures and interdependencies that may make sense on Big Data. The term "Big Data" refers not only to the data as such, but also to the practices, infrastructures, networks and policies that influence their various manifestations. Understanding big data as a set of emerging technologies seems to be conceptually useful, as it "encompasses digitally enabled developments in data collection, analysis, and utilization." (Richterich, 2018)

In this context, Rip describes the dilemma of technological developments: "For emerging technologies with their indeterminate future, there is the challenge of articulating appropriate values and rules that will carry weight. This happens through the articulation of promises and visions about new technosciences." (Rip 2013, 192) Thus, emerging technologies are places of "pervasive normativity" characterized by articulating promises and fears, conceptualizing it as an approach "in the spirit of pragmatist ethics, where normative positions co-evolve" (Rip 2013, 205)

*Pragmatic* ethics emphasizes that new technologies are developing in societies in which they are discursively associated/dissociated by certain norms and values. At the same time, pragmatism states that increasing the large number of data and research-related practices is not a simple matter of technological superiority. They form a field of normative justification and contestation.

The *neo-pragmatic* approach to ethics addresses epistemological knowledge through the falsification of (scientific) knowledge, with critical evaluations of social power structures. Keulartz et al. have proposed a pragmatic approach to ethics in a technological culture (Keulartz et al. 2004) "as alternative which combines the strengths of applied ethics and science and technology studies, while avoiding the weaknesses of these fields." (Richterich, 2018) Thus, applied ethics is an effective approach in terms of detecting and expressing the norms involved in (inter-) socio-technical actions or resulting from socio-technical actions, but it has no possibilities to capture the inherent normativity and the agent of technologies. (Keulartz et al. 2004, 5)

Keulartz et al. believes that the lack of normative technological evaluations can thus be overcome: "impasse that has arisen from this" (i.e. the respective 'blind spots' of applied ethics and STS) "can be overcome by a re-evaluation of pragmatism." (Keulartz et al. 2004, 14) Ethical pragmatism can be characterized by three common principles: anti-foundationalism, anti-dualism and anti-scepticism.

*Anti-foundationalism* refers to the principle of falsifiability, considering that we cannot reach certainty in terms of knowledge or values ("ultimate truth"), but knowledge, as well as values and norms, changes over time. Moral values are not static but can be renegotiated depending on technological developments.

*Anti-dualism* implies the need to refrain from predetermined dichotomies. Among the dualisms criticized by Keulartz are the essence/appearance, theory/practice, consciousness/reality and

facts/value. Applied ethics tends to assume such dualisms as *a priori*, as opposed to pragmatism, which underlines the blurred interrelations and lines between such categories.

*Anti-scepticism* is closely linked to the need for situated perspectives and explicit normativity, relating to the anti-Cartesian foundation of pragmatism.

In European research, pragmatism was usually dismissed as superficial and opportunistic, being associated with negative stereotypes, (Joas 1993) being accused of "utilitarianism and meliorism." (Keulartz et al. 2004, 15) At the end of the 1990s and 2000s, pragmatism experienced a revival in European research. (Baert and Turner 2004)

European Economic and Social Committee, in "Big Data: Balancing economic benefits and ethical questions of Big Data in the EU policy context", states that Big Data analysis from an ethical point of view involves two main interdependent aspects: a theoretical one (the philosophical description of the elements subject to ethical control) and a pragmatic vision (of the impact on the lives of people and organizations). (European Economic and Social Committee 2017)

There are ethical problems caused by artificial intelligence, and a close link between Big Data and artificial intelligence and its derivatives: machine learning, semantic analysis, data exploitation.

An ethical approach is through the moral agency with at least the three conditions of causation, knowledge and choice. According to Noorman: (Noorman 2012)

- There are causal links between people and the outcome of actions. The person's responsibility derives from the control over the result.
- The subject should be informed, including on possible consequences.
- The subject must give his consent and act in a certain way.

Professor Floridi, in *The Fourth Revolution*, identifies the moral problem of Big Data with the discovery of a simple model: a new frontier of innovation and competition. (Floridi 2014) Another

problem associated with Big Data is the risk of discovering these patterns, thus changing the predictions.

The basic rule of Big Data ethics is the protection of privacy, freedom and discretion to decide autonomously. It is worth noting that there is a continuous tension between the individual needs and those of a community.

It is possible to identify several ethical issues arising from the exploitation of Big Data: (European Economic and Social Committee 2017)

- *Privacy* - The extreme limit of confidentiality is the seclusion, defined by Alan F. Westin as "the voluntary withdrawal of a person from the general society through physical [means] in a state of solitude". Moor and Tavani defined a privacy model called Restricted Access Control (RALC) that differentiates between privacy, justification, and privacy management.
- *Tailored reality and the filter bubble* - The application on a server collects information by learning from it, and then uses that information to build a model of our interests. When a system uses these models to filter information, we may be induced to believe that what we see is a complete view of a specific context, when in fact we are limited by the "understanding" of an algorithm that built the model. The ethical effects can be multiple: some information can be hidden, imposing prejudices which we do not know, our vision of the world can become progressively limited, and in the long term could generate a certain point of view.
- *After death data management* - What happens to the data of a deceased user? Do the heirs become their owners? Can data be removed from the digital world? There are legal and technological problems here.
- *Algorithm bias* - Data interpretation almost always involves certain biases. In addition, there is a possibility that an error in an algorithm may introduce bias forms. An ethical issue is our

implicit trust in algorithms, with high risks when risks are not taken into account due to programming or running errors of the algorithms.

- *Privacy vs. growing analysis power* - It refers to the emergent nature of information as a complex system: the result of data from different contexts is more than the simple sum of the parts.
- *Purpose limitation* - It is very difficult or even impossible to limit the use of data. Privacy is not a single block, with subtle forms of privacy being lost.
- *User digital profile inertia and conformism* - This is about the subject of personalized reality. A model that involves a user's interests is usually based on past behavior and past information. Thus, the algorithms are not based on the actual identity of the person, but on an earlier version. This will influence the real behavior of the user, being pushed to maintain their old interests and therefore not be able to discover other opportunities. If the user is not aware of this problem, the influence of inertia will be much greater.
- *User radicalization and sectarianism* - Big Data can form opinions using filtering/recommendation algorithms, information, personalized articles and posts, and specific recommendations from friends. Thus, users will be more and more in touch with the people, opinions and facts that will support their original position. This tendency is often hidden from the users of Big Data based systems, with the tendency to develop prejudices, ranging from conformity to radicalization. It is possible to postulate the formation of a kind of technological subconscious with impact on the development of the personality of the users, phenomena evident in the case of social networks, where the distance between the real ("physical") world and the Internet is strongly attenuated.
- *Impact on personal capabilities and freedom*

- *Equal rights between data owner and data exploiter* - Usually the person whose data is used is not their legal owner. Therefore, a minimum requirement is for that person to have access to their own data, allowing them to download them and eventually delete them.

## Bibliography

- Baert, Patrick, and Bryan Turner. 2004. "New Pragmatism and Old Europe: Introduction to the Debate between Pragmatist Philosophy and European Social and Political Theory." *European Journal of Social Theory* 7 (3): 267–74. <https://doi.org/10.1177/1368431004044193>.
- Collins, Tom. 2016. "Security Fears Sparked over Wearable Technology." Mail Online. December 19, 2016. <http://www.dailymail.co.uk/~/article-4049154/index.html>.
- European Economic and Social Committee. 2017. "The Ethics of Big Data: Balancing Economic Benefits and Ethical Questions of Big Data in the EU Policy Context." European Economic and Social Committee. February 22, 2017. <https://www.eesc.europa.eu/en/our-work/publications-other-work/publications/ethics-big-data>.
- Floridi, Luciano. 2014. *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. OUP Oxford.
- Joas, Hans. 1993. *Pragmatism and Social Theory*. University of Chicago Press.
- Keulartz, Jozef, Maartje Schermer, Michiel Korthals, and Tsjalling Swierstra. 2004. "Ethics in Technological Culture: A Programmatic Proposal for a Pragmatist Approach." *Science, Technology, & Human Values* 29 (1): 3–29. <https://doi.org/10.1177/0162243903259188>.
- Mearian, Lucas. 2015. "Insurance Company Now Offers Discounts -- If You Let It Track Your Fitbit." Computerworld. April 17, 2015. <https://www.computerworld.com/article/2911594/insurance-company-now-offers-discounts-if-you-let-it-track-your-fitbit.html>.
- Noorman, Merel. 2012. "Computing and Moral Responsibility." *Stanford Encyclopedia of Philosophy*.
- Richterich, A. 2018. "The Big Data Agenda: Data Ethics and Critical Data Studies." <https://doi.org/10.16997/book14.b>.
- Rip, Arie. 2013. "Pervasive Normativity and Emerging Technologies." In *Ethics on the Laboratory Floor*, edited by Simone van der Burg and Tsjalling Swierstra, 191–212. London: Palgrave Macmillan UK. [https://doi.org/10.1057/9781137002938\\_11](https://doi.org/10.1057/9781137002938_11).
- Tanner, Adam. 2014. "Different Customers, Different Prices, Thanks To Big Data." Forbes. 2014. <https://www.forbes.com/sites/adamtanner/2014/03/26/different-customers-different-prices-thanks-to-big-data/>.