Digital Social Science

Innovative

Article

Research Ethics in Digital Social Sciences

Aiman Zahra*

Department of Sociology, Bahauddin Zakariya University, Multan 60000, Pakistan
* Correspondence: aimanzahra9136@gmail.com

Abstract: Since last decade, digital technologies have revolutionized the social science research, due to unlimited access to large-scale behavioral data, computational modelling, and analysis using AI. However, moral issues of privacy, monitoring, algorithmic prejudice, assent, and data breach have emerged simultaneously. Current study explores the relationship between technology ethics, and digital social science, suggesting that methodological innovation in social science is inextricably linked with normative ethical issues. The study explores surveillance capitalism, data ethics, and algorithmic accountability to examine the technological and ethical shifts towards reshaping digital social science. Digital social science can focus pattern recognition over meaningful explanation if it's does not use a strong theoretical foundation and methodological reflexivity. The study highlights that research ethics boards and regulations, needs to be adjusted to computational research environments. An ethically-centered digital social science, needs transparency, collaborative data stewardship and responsible use of AI. The future of digital social science hinges on combining technological innovation with ethical considerations, with the aim of not sacrificing social responsibility or scientific integrity for the sake of computational power.

Keywords: research ethics; digital social science; datafied; algorithmic bias; digital environment

1. Introduction

Digital technologies have changed many aspects of today's social life, including but not limited to communication, political engagement, economic transactions, and identity creation. In datafied societies, research methods in social science have also transformed. Besides traditional survey based and ethnographic methods, computational methods have been added and, in some cases, substituted, that interpret digital traces collected from platforms, sensors, and algorithmic systems (Kitchin 2014). This transformation has created a new discipline called digital social science, which combines methodologies from computer science, big data analysis, machine learning, and network analysis to investigate social phenomena at a large-scale. These tools have more analytical power, but have raised ethical issues that call into question the principles of research.

Accessibility of data and ethical responsibility are among the main dichotomies in digital social science. Large-scale datasets, gathered from social media platforms, mobile devices, and online interactions may seem like they are readily available but, when harnessed, they can pose serious issues of informed consent, contextual integrity, and privacy expectations (boyd and Crawford 2012). Furthermore, the growing dependence on proprietary platforms for accessing data introduces asymmetries of power between researchers, corporations and research subjects. Meanwhile, algorithmic systems and artificial intelligence has increased concerns about issues of bias, transparency and accountability. Structural inequalities, present in training data of machine learning models can be reflected in the resulting model producing misinterpretations or even unethical assessments of social reality (O'Neil 2016). All this, mirrors the need for a strong ethical framework to deal with the technical and normative aspects of digital research.

Shoshana Zuboff (2019), claims that the current digital ecosystems are governed by the 'surveillance capitalism,' which involves platform business extracting, analyzing, and monetizing behavioral data. This context makes the ethical stance of social scientists' dependent on these kinds of data infrastructures difficult, as they have to ask themselves questions regarding complicity, data legitimacy, and epistemic dependence on commercial actors. Luciano Floridi (2014) highlights the need for a philosophy of information that considers the moral dimension of the data processes themselves. The integration perspective regards ethics as a part of the process of information generation, processing and interpretation, and not as an external constraint on research practices.

It is essential not only for technology to keep advancing but for ethical reflexive research infrastructures to be created for digital social science to be the future as per the paper. Otherwise,

Citation: Aiman Zahra.2025. Research Ethics in Digital Social Sciences. *Digital Social Sciences 2(2),* 18-23. https://doi.org/10.69971/dss.2.2.2025.44

computational approaches may reinforce the surveillance systems, exacerbation of inequalities, and loss of public trust in scientific research. Current study explores changes in the methodological practices of social science research due to digitalization, key ethical issues arising due to computational and data-driven methods, existing governance frameworks to these challenges and the principles of developing "ethically responsible digital social science".

2. Theoretical Framework

The principles and theories of digital social science are rooted in emerging fields of information theory, critical data studies, philosophy of technology, and ethical theory. Epithecological foundations in sociology, anthropology, and political science are being reworked as computational approaches are becoming integral part of the generation of social knowledge. The conceptual pillars of digital epistemology, datafication and platform society, and ethical theories of information and computation are discussed below.

Digital epistemology includes the process of knowledge creation, verification and sharing with computational systems. Digital social science is not based on sampled observations, interviews or fieldwork, but rather on massively scaled behavioral data created and recorded due to digital interactions (Kitchin 2014). This indicates an epistemological change as knowledge is no longer obtained through a research design, but from streams of data available online. The social reality is increasingly derived from the use of digital traces like clicks, likes, GPS signals and communication metadata. This suggests some overlap between observation and intervention.

Big data is not objective and neutral (boyd and Crawford 2012). Rather, it is a manifestation of local practices of gathering and reading, as well as algorithmic filtering. The 'data' is a construction of technology design and institutional agendas. This is a critique of the notion, in digital social science, that it affords a more direct, or "raw" entry to reality. Interpretive opacity has new meanings in the epistemic authority of algorithms. There are concerns regarding transparency in machine learning systems, as they can make predictions without providing logical explanations for their actions. It's difficult to understand and explain decisions of machine learning systems, making it hard to ensure scientific accountability. Here, the process of knowledge production is in part shifted to a non-human system whose decision-making mechanism is not necessarily accessible to the researcher.

Datafication is the process of quantifying social action into digital data. Datafication is a socio-political process as well as a technical one, and is inherent in institutional structures and power dynamics (Kitchin 2014). In today's society, there are platforms that are capturing, storing and analyzing user behavior in real time, which is increasingly becoming the norm of everyday life.

Zuboff (2019) proposes the notion of surveillance capitalism to capture behavioral surplus data from digital platforms to predict and manipulate human behavior for commercial gain. This model is not about users as a part of a digital system but about users as raw material for predictive analytics markets. Data sources created in commercial ecosystems like social media, search engines, and mobile apps are frequently used by researchers. Usually, however, these datasets are private and come with epistemic asymmetries. Structure biases affect the design of research as well. Structural biases also arise in the design of research on the data platformization. The algorithms decide what data is being visible, how it is being ranked, and the interactions that are logged. Thus, digital social sciences can be reduced to platform logic, rather than a space for new modes of inquiry.

3. Technology Ethics in Digital social sciences

The ethics of digital systems is understood on a normative basis Floridi (2014). He maintains that the current era is the "infosphere," in which the ontological basis of reality is information entities. In this context, ethics also applies to the integrity and quality of informational contexts, not just human ones. Digital social science, therefore, is not only an analytical method, but is a moral activity that influences the infosphere. Data collection, processing and interpretation demands responsibilities regarding accuracy, transparency and harm prevention.

Algorithmic systems are subject to ethical issues at various phases of their development including collection, design, deployment, and construction (Mittelstadt et al. 2016). Bias is not merely a technical issue, but is an intrinsic property that arises from socio-technical configurations. Cathy O'Neil (2016) highlighted the risks associated with opaque algorithmic systems, which she calls "Weapons of Math Destruction" when they operate on a large scale, lack regulation, and are destructive. They can perpetuate inequities by incorporating historical biases into systems that are automated for making decisions. Ethical computation is not just about technology solutions. It requires reflective consciousness of the production of knowledge and social outcomes as influenced by computational systems. This involves issues of fairness, accountability, transparency and justice in algorithmic design.

Due to digital epistemology and ethical computation, methodological and normative questions are indistinguishable from each other. Digital systems face ethical issues of inclusion, representation, and interpretation in the creation of social scientific knowledge. In the traditional social science, ethical oversight is generally regarded as a formal procedure that is regulated by institutional review boards. But digital settings raise ethical questions as part of the technology of data production. Ethics thus must go beyond the level of research approval and be part of the system design. Digital social science must adopt a different approach and move towards embedded ethics in which ethical reasoning is embedded in all aspects of computational research. This involves both creating the data sets and choosing the algorithms, testing the model, and publishing the results. This would reflect the growing calls for responsible innovation and ethical governance of AI, which center on anticipatory, participatory, and adaptive ethical practices.

Digital environments are distinct from traditional research by its ongoing data generation, passive nature, and massive scale, which are not limited to explicit consent and controlled sampling. Four key ethical domains are explored: privacy and surveillance, informed consent in a digital environment, algorithmic bias of research datasets, and platform dependency and epistemic integrity.

Privacy is one of the key aspects of digital social science. With proliferation of digital traces, created as users browse the internet and engage in social media, mobile phone, wearable devices and online services, the boundary between public and private information has become diffuse. Today's digital ecosystems are characterized by surveillance capitalism which is continuous collection of personal data for predicting behavior and monetizing it (Zuboff 2019). Social scientists are not always directly involved in these commercial activities, but often use data from these data infrastructures. This poses moral dilemma as observation-based

research might be actually dependent on surveillance systems. “Publicly available” digital information doesn't always mean that it's ethically accessible for free. Users seldom expect that their online actions could be collated, analyzed and repurposed for research. Even for accessible data, the traditional notion of privacy is undermined in digital contexts (boyd and Crawford 2012). Thus, privacy in digital social science is not only about confidentiality; it is also about the right information flow in a certain social context, called contextual integrity.

One of the core principles of research ethics is informed consent, which has been traditionally understood to mean that the data subject must be informed and agree to the use of their information. In digital social science, however, it becomes more challenging to get meaningful consent. A lot of the data is gathered on a platform rather than by the people themselves. Behavioral data is collected by social media APIs, through web scraping, and via mobile tracking systems without user consent. The concept of informed consent is then undermined in such cases. Ethical responsibility in infosphere should be not only individual, but systemic as well (Floridi 2014). People understand that they are not able to practically understand or control the entire flow of data in a complicated digital ecology. Consent models, when in place, are often opaque and lengthy, and they are not truly negotiable, such as platform terms of service. This brings into question the concept of procedural consent vs substantive consent, as users might consent to the terms without actually comprehending what it means. Digital social science needs to explore other models of consent, such as dynamic consent frameworks, collective consent models, and ethics-by-design approaches.

Another issue that of algorithmic bias. Algorithms powered by machine learning can perpetuate and exacerbate inequality in society if they are based on social datasets that contain disparities. When algorithmic systems are opaque, harmful and scalable, they can turn into “Weapons of Math Destruction” (O’Neil 2016). Biased algorithms can bias empirical results, which can lead to incorrect social behavior conclusions in social science research. Algorithmic bias is not a technical problem, but a socio-technical phenomenon, which occurs through the process of data selection, model design, and deployment contexts (Mittelstadt et al. 2016). There can be several points of bias in the process, such as the methods used to create the data set (e.g. sampling bias, missing data). Both misreading results and making overgeneralized predictions call into question the epistemic validity in digital social science. If analytical tools are systematically biased, then the knowledge produced by the tools can contribute to the reinforcement of existing social inequalities rather than reflecting social reality.

Access to data is one of the hallmarks of modern digital social science, and it is dependent on private platforms. Social media platforms, search engines, and online markets hold large amounts of information which are difficult to access by other researchers. This introduces an epistemic asymmetry, and corporations are more aware of this social behavior than academic institutions. Researchers need to access APIs, data sharing agreements or third-party data, and they have no control over these restrictions or changes. Platform infrastructures, are primarily built for business optimization, not for science (Zuboff 2019). Therefore, data for researchers has already been processed by proprietary algorithms which define what is measurable and visible. The dependency is problematic in terms of the independence of social science as a subject. Prominent research programs and agendas can foster a danger of science being driven by corporate interests, rather than independent epistemic agendas, due to their platform accessibility.

All above ethical issues are inextricably linked. Platform dependency is not a separate dimension, but rather another facet of the socio-technical system of privacy, consent and bias. Digital social science takes place within infrastructures that facilitate and limit the production of knowledge. Ethical reflection needs to shift away from principles, and towards a more holistic framework which considers the systemic nature of digital research environments. It involves a transition from reactive ethics (in response to harm) to anticipatory ethics (prevent harm in design of research systems).

4. Methodological Transformation in Digital Social Science

The advent of big data has given social scientists the ability to study populations at scales much larger than they had ever been able to do before. Researchers can track the behavior of users in a near real-time manner by capturing digital traces that are produced via social media platforms, mobile devices, sensors, and online transactions. Kitchin (2014) talks about this change as the “data revolution”, wherein massive amounts of data are playing a pivotal role in the study of social phenomena. Unlike traditional survey methods, which sample populations, big data methods may try to sample entire populations or continuously generated data. There are a number of methodological implications in this change. First, it decreases the need to rely on self-reported information which has, in the past, been prone to recall bias and measurement error. Second, it allows for the analysis of social behavior over time, or a longitudinal analysis, and a dynamic analysis. It also brings in new types of bias, though, that stem from the platform design, data availability, and algorithmic filtering. Notably, theory cannot be replaced by big data. If not framed by theory, vast sets of data may yield a correlation with no real significance. However, as boyd and Crawford (2012) warn, having lots of data does not necessarily provide epistemic depth or explanatory power.

Integration of machine learning into digital social science has become a tool for recognizing patterns, categorizing behaviors, and forecasting outcomes from complex data. Techniques like natural language processing, clustering and neural networks in sociological and political analysis are now widely used. However, the use of AI in social research faces various methodological issues. An important one is the interpretability. In many machine learning models, the processes used to reach the output are not explained, and they are considered a “black box.” This makes it difficult to follow the typical scientific approach of explanation and verification. Algorithmic systems can have hidden assumptions that impact outcomes in ways that may be hard to identify (O'Neil, 2016). Computational models can be uncovering social patterns or creating patterns through biased optimization functions. AI systems are very data dependent. Models can actually perpetuate the distortions of past data if the data used to create them also contain such inequalities or sampling errors. Oracle's Mittelstadt et al. (2016) point out that algorithmic bias is not a post-processing problem, but rather a structural one related to data pipelines.

However, digital social science has also led to the development of combinations of quantitative and other methods, like computational ethnography. These methods involve both qualitative (traditional) and computational research techniques, allowing researchers to explore patterns at macro levels and meaning at micro levels. Digital ethnographers, can analyze the social media interactions and complement them with interviews or participant observation to interpret behavior. One of the main drawbacks of big data methods is the lack of interpretation, which this integration helps remedy.

Computational ethnography also poses ethical issues, however. Addressing questions of consent, anonymity and the context in which data is interpreted are important challenges when using socially sensitive data that is accessible to the public. In an online environment, there are blurred boundaries between observation and participation. The recent use of computational approaches has reignited discussions about the issues of validity and reliability in the social sciences. The traditional standards for assessing the quality of research in the era of dynamic and algorithmically mediated data sets needs to be reconsidered.

Data instability is one of the issues. Study replication is often difficult due to many digital platforms changing their APIs, algorithms, and data accessibility policies. Changes in the platforms might make the data set collected at one time difficult to re-create later. Digital traces are frequently indirect measures, rather than direct measurements of social behavior. For instance, "likes" on social media do not necessarily represent endorsement, agreement or interest. If not interpreted carefully, researchers can end up misinterpreting platform-specific interactions with the bigger social message. Moreover, proprietary algorithms and closed datasets make it difficult to replicate in computational social science. If research relies on private data infrastructures, independent confirmation will be hard, which is one of the fundamental aspects of research.

In view of these obstacles, digital social science must be even more reflexive with regard to methodology. They have to critically reflect on the tools they use to analyze, as well as the socio-technical systems that generate their data. This means that data sources are transparent, there is a clear documentation of the preprocessing steps, algorithmic limitations are acknowledged. It also needs to be acknowledged that the computational products are not mere products of reality, but designed artifacts that are influenced by design choices. According to Floridi's (2014) philosophy of information, data processing is considered as an ethical activity. From this perspective, methodological decisions are also decisions of values and morality since they shape the representation and understanding of social reality.

Digital social science is a new and more complex form of scientific learning. While big data and AI offer immense analytical power, they also raise concerns about potential biases, lack of transparency, and the interpretability of the results. Finally, there must be a foundation of ethics and theory behind methodological innovation. If they are not integrated, the digital social science can be technologically cutting edge, but conceptually lacking.

5. Governance and Regulation of Digital Social Science

Digital technologies have introduced new opportunities for researchers not only in the volume of social data that can be collected, but in the underlying principles of social science methodology. In the age of big data, machine learning and computational modeling, the focus has shifted from hypothesis-driven work to pattern-driven work that is heavily data-dependent. This shift opens up new opportunities and challenges with respect to the questions of validity, interpretation, and ethical responsibility.

Current governance and regulatory structures have not been able to keep pace with the rapid growth of Digital Social Science. With the growing importance of platforms, algorithmic systems and transnational digital infrastructures in research, there has been a growing focus on issues of accountability, compliance and ethical governance in research. Four dimensions are explored here, including the global data protection frameworks, institutional ethics review frameworks, platform governance and structures, and jurisdictional issues in digital research. The General Data Protection Regulation (GDPR) by the European Union (EU) is one of the most impactful regulatory changes in the digital age, having entered into force in 2018. Under GDPR, there are clear guidelines for data collection, processing and storage, including principles of lawfulness, transparency, purpose limitation and minimization. GDPR has profoundly transformed research in relation to personal data in digital social science. Researchers have to provide a rationale for data collection mechanisms, ensure anonymization of data if available, and substantiate the ethical and legal basis for the use of data. But applying GDPR to computational research is not easy. There are secondary data sources that often make use of big data methodology, and obtaining the direct consent of the individual is impossible. This puts a strain on the compliance with the regulations and the feasibility of the methodology. In addition to the European context, there are different data protection regimes in other jurisdictions, such as the California Consumer Privacy Act (CCPA) in the United States and new Data Governance initiatives in Asia. Ethical and legal requirements for digital research are not and are not expected to be globally standardized.

The traditional method of ethical oversight in social science research is called an Institutional Review Boards (IRBs) or Research Ethics Committees (RECs). These bodies review research proposals to ensure that they follow ethical guidelines, especially those concerning human subjects. IRBs had been developed primarily for traditional research methods like surveys, interviews and experimental designs. These frameworks are inadequate to address the new challenges of digital social science. IRB's have a difficult time reviewing research projects that include publicly accessible social media data, algorithmic scraping, or massive behavioral inference. Questions emerge as to whether such data is human subjects research, and whether consent is needed when data is public, but contextually sensitive. Ethical regulation of digital environments should become more continuous and adaptive, rather than relying on pre-approval systems (Floridi 2014). This is because digital ecosystems are dynamic by nature and platform architectures evolve quickly.

Digital social science uses privately-owned digital platforms which are large datasets for computational research, and being controlled by social media networks, search engines and cloud service providers. This results in “platform governance”, as the data available on the platform, its structure and whether it can be analyzed are determined by the commercial policies. These platforms are not public research infrastructures, but rather entities that are driven by commercial mandates that are not scientific, such as profitability, user engagement and regulatory compliance. For Zuboff (2019), this concentration of power over data is a new type of imbalance, as the data of human behavior is harvested and sold for profit in the systems of surveillance capitalism. It has a significant impact on researchers, because they will have to rely on empirical material that can be modified at the last minute as a result of corporate interests. This depends upon one another has raised concerns for academic autonomy. The risk of having research agendas driven by the availability of platforms is that some topics may appear more prominently than others for reasons other than scientific interest, for example, because the data to be studied are more easily available.

Digital social science is by nature a transnational entity. The information collected in one jurisdiction is frequently stored, processed or analyzed in another jurisdiction, which presents challenging legal and ethical issues. The concept of privacy, data ownership and research ethics vary across countries. This dis-harmonization sets non-tailored compliance for international research

collaborations. For instance, data that is deemed anonymous in one jurisdiction could be deemed to be personal data in another jurisdiction. Additionally, cloud infrastructures typically store data on a number of servers in various countries, which can create a challenge to identify the applicable legal measures. This creates problems of uncertainty in jurisdiction and overlapping regulations. Kitchin (2014) points out that data infrastructures are not merely technical systems but socio-political arrangements which are negotiated across global power relations. Therefore, geopolitical aspects of data flows need to be taken into consideration when guiding governance of digital social science.

Existing regulatory approaches are limited and an increasing demand is emerging for more adaptive and integrated governance approaches. These mechanisms would be in the form of legal regulation, institutional oversight, and platform accountability mechanisms. Adaptive governance is based on flexibility, meaning that ethical guidelines need to be adaptable to technological changes. It also promotes the cooperation of all research, policy and technology providers, to ensure that ethical standards are practical and enforceable. Research ethics bodies at the institutional level may require additional skills in computational areas to be able to better assess computational studies. Likewise, platforms could be mandated to be more transparent and give researchers more access to data for legitimate research. In conclusion, successful governance of digital social science is about innovation and responsibility, so that technology does not compromise the ethical aspects of this discipline.

6. Future of Digital Social Science

These trends will influence the future of digital social science, potentially shaping its trajectory in the years to come. This is likely to continue to change in methodological direction and in what are the "good" values of the discipline as it enters more deeply into social life. Four aspects are examined: AI-assisted social science, ethical AI ecosystems, decentralized data infrastructures and human–machine co-production of knowledge.

From data collection to data preprocessing, analysis to interpretation, Artificial Intelligence (AI) has been implemented in every step of research in social sciences. Natural language processing models can be used to process massive amounts of text data like news articles, social media posts, and policy documents, and machine learning algorithms can help with predictive modeling of social behavior. The trend is indicative of a future where parts or all of the research process could be semi-automated or fully automated. Efficiency and scale come with such a price as epistemic authority question. While algorithms can hypothesize, identify patterns, and even come up with interpretations, the human's job as a researcher could move from knowledge producer to knowledge curator. Gautier, however, cautions that the automated systems are not objective observers (O'Neil 2016). They represent assumptions that are part of the data and design decisions. Thus, the use of AI in research needs to be strongly supervised to avoid blindly accepting what the AI produces, without critically examining it.

Ethical AI ecosystems, in which AI tools are embedded and built with transparency, accountability, and fairness, will likely be instrumental to realizing the future of digital social science. Information systems are "morally loaded environments (Floridi 2014). Ethical AI ecosystems should be concretely embedded in data infrastructures, so that ethical considerations are not something added on, but as an integral part of the design of those infrastructures. These ecosystems can support explainable AI models, tools for bias detection, algorithmic decision-making audit trails, and participatory design processes that engage stakeholders in shaping the development of an ecosystem. These tools take the approach of computational efficiency and ethical responsibility. Systemic changes to the whole data pipeline are necessary to deal with algorithmic ethics (Mittelstadt et al. 2016). This encompasses the process of creating the models, handling the data, setting up deployment environments, and establishing evaluation procedures after deployment.

Digital social science can move towards open and decentralized data infrastructures. However, because dataset is currently controlled by a central platform, this may yield epistemic asymmetries and limit research autonomy. Decentralized systems, facilitated by blockchain or distributed ledger technologies, can help individuals to control their information, allowing secure and anonymized access to research. Governments can use open data campaigns to contain private platforms. Decentralization, can lead to challenges of data governance complexity, security issues and unequal access to technological infrastructure. Decentralized systems may perpetuate inequalities, if not designed appropriately. Data infrastructures, are dynamic and always socio-technical assemblages that are influenced by institutional power (Kitchin 2014). Thus, open data ecosystems do not only rely on technological possibilities, but also on political commitment and regulations.

Future research may include the use of collaborative epistemic systems whereby humans and machines collectively create insights rather than using AI systems as tools or assistants. This model combines the ability of machine learning algorithms to detect patterns in large data sets with human researchers' contextual interpretation, theoretical framing, and ethical judgment. The separation of analytical and empirical effort can deepen the depth of analysis as well as the range of empirical data. Co-production also challenges the issue of accountability, however. In shared co-creation process, knowledge, responsibility for errors or biases grows more complicated. This highlights the importance of good governance guidelines setting accountability in hybrid research systems. By over-reliance on machine-generated insights, researchers may become epistemic deskilled and rely on automated outputs without being clear on the processes underlying these outputs.

Reflexive science recognizes that social science research is not without agency, rather it co-constructs the social reality that it investigates. Digital data is not "found" but created with intricate socio-technical systems. There are strong relations between digital infrastructures and power and economic interests' systems (Zuboff 2019). Thus, social science education needs to be cautious towards the political economy of data production in the future.

7. Conclusion

Impact of Digital technologies on the transformation of social science is one of the biggest epistemic shifts in the history of social science. The use of big data analytics, the use of artificial intelligence and the platform-based infrastructure have made empirical research more massive and involved, allowing researchers to observe social phenomena in a finer-grained and more rapid manner. However, these methodological advances have raised ethical, political and epistemological issues. Digital tools facilitate

new methods of data collection and analysis, but also carry with them inherent risks of surveillance, invasion of privacy, algorithmic bias and epistemic distortion. Reliance on platform data also poses challenges to the independence of social science as the researchers depend upon infrastructures are run by commercial interests and not academic criteria. Accordingly, there is reconfiguration of the production of knowledge that is not just a methodological shift, but the transformation of digital social science. Based on the digital epistemology, datafication theory and the philosophy of information, it is clear that data is not a neutral resource, but a man-made artefact that is socially and techno-constructed. Thus, ethical issues need to be integrated into the actual processes of data generation, processing and interpretation. An extension of research ethics paradigms to the digital environments is needed. Consent, privacy and fairness issues cannot be handled via the procedural process alone. Rather, they need a systemic and anticipatory ethical attitude which takes the distributed and dynamic nature of digital ecosystems into account. Although the tools of computation are very powerful, they are not epistemically neutral. New types of bias, opacity and interpretive uncertainty are introduced by machine learning systems and big data analytics. Existing systems of governance are unable to catch up with technological change. The fragmentation of the world and the dependency on platforms pose serious challenges for data ethical oversight. The viability of digital social science as a viable academic subject depends on its potential to engage with digital technologies it uses.

References

Boyd, danah, and Kate Crawford. 2012. Critical Questions for Big Data. *Information, Communication & Society* 15: 662–679. https://doi.org/10.1080/1369118X.2012.678878

Floridi, Luciano. 2014. *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford: Oxford University Press.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.

Mittelstadt, Brent D., Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The Ethics of Algorithms: Mapping the Debate. *Big Data & Society* 3:1-21. https://doi.org/10.1177/2053951716679679

O'Neil, Cathy. 2016. *Weapons of Math Destruction*. New York: Crown.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism*. New York: PublicAffairs.