UDC 323.31. https://doi.org/10.33619/2414-2948/109/62

BIG DATA ANALYTICS AND ETHICAL RESPONSIBILITY IN THE DIGITAL AGE

*©**Yoğurtçu G.**, ORCID: 0000-0002-2404-3908, Dr., Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan, gokce.yogurtcu@manas.edu.kg*

БОЛЬШИЕ ДАННЫЕ И ЭТИЧЕСКАЯ ОТВЕТСТВЕННОСТЬ В ЦИФРОВУЮ ЭПОХУ

*©**Югуртчу Г.**, ORCID: 0000-0002-2404-3908, Доц.док., Кыргызско-Турецкий университет «Манас», г. Бишкек, Кыргызстан, gokce.yogurtcu@manas.edu.kg*

Abstract. Big data analytics possesses the capability to process vast volumes of digital content and analyze societal sentiment trends in real time. Such analyses offer comprehensive insights and detailed models on how social awareness forms and evolves, particularly during periods of crisis. However, due to the immense, diverse, and complex nature of these datasets, traditional analytical methods often fall short, necessitating the ongoing development of novel interdisciplinary approaches and methodologies. At the same time, significant debates continue regarding the extent to which these datasets accurately represent the real world, given the existence of the digital divide and the ethical frameworks governing the collection and application of big data. This article explores the theoretical and methodological foundations of big data-driven research, the societal impacts arising from inequalities in data access, and the ethical issues that emerge during the collection and analysis of extensive datasets. The study consists primarily of secondary data obtained from existing literature, case studies, and the application of big data analytics in various research fields.

Аннотация. Аналитика больших данных обладает способностью обрабатывать огромные объемы цифрового контента и анализировать тенденции общественного настроения в режиме реального времени. Такие анализы предоставляют всесторонние взгляды и детализированные модели формирования и трансформации общественного сознания, особенно в условиях кризисных ситуаций. Однако из-за огромного, разнообразного и сложного характера этих наборов данных традиционные методы анализа часто оказываются недостаточными, что требует постоянного развития новых междисциплинарных подходов и методологий. В то же время продолжаются значительные дебаты о степени, в которой эти наборы данных точно отражают реальный мир, учитывая существующий цифровой разрыв, а также о этических рамках, регулирующих сбор и использование больших данных. В данной статье рассматриваются теоретические и методологические основы исследований, основанных на больших данных, социальные последствия, вызванные неравенством в доступе к данным, и этические вопросы, возникающие при сборе и анализе обширных наборов данных. Исследование состоит в основном из вторичных данных, полученных из существующей литературы, тематических исследований и применения аналитики больших данных в различных областях исследований.

Keywords: Big Data, Big Data Analytics, Digital Divide, Data Ethics.

Ключевые слова: большие данные, аналитика больших данных, цифровое неравенство, этика данных.

The continuously expanding network of digital interactions in contemporary society generates vast digital traces that reflect individual perspectives on a wide range of events. These digital traces form complex and interconnected large datasets, deeply entwined with various aspects of social and institutional life, providing profound insights into modern societal dynamics. Big data serves as a crucial domain for accelerating scientific research, analyzing public discourse, and examining societal psychological phenomena.

In the broadest sense, data is defined as “raw elements that can be abstracted from phenomena, measured, and recorded in various ways” [1]. Data differs fundamentally from facts, evidence, information, and knowledge, as it precedes and underpins all these components. Data forms the interconnected elements that constitute information, which in turn leads to the creation of organized knowledge. Consequently, data serves as empirical evidence with the potential to be transformed into both facts and knowledge [1].

Such information and evidence render big data sets invaluable for understanding the dynamics of complex social systems. Big data encompasses both structured and unstructured data from various online and offline sources [2]. The conceptual and operational framework for defining big data is extensive, with two primary approaches emerging. The first emphasizes the need to understand the conceptual and philosophical nature of data, rather than solely focusing on its practical use for generating insights and value [1]. The alternative approach centers on the analysis and practical applications enabled by this data [3].

Both approaches recognize key defining characteristics that establish the scope and functionality of big data. These characteristics are volume, variety, velocity, veracity, and value.

Volume refers to the size of data files used for archiving and dissemination. While definitions of large data volumes are relative, digital environments can process and store millions of texts, audio recordings, photos, and videos per second. Such datasets, varying widely in size, demand diverse data analytics methodologies [4]. Global data volume estimates indicate that the capacity to store data struggles to keep pace with its rapid generation [1].

Another characteristic that necessitates real-time analysis of high-volume data flows is velocity [4]. Velocity, a key distinction between small data and big data, refers to the rapid and dynamic nature of the big data cycle. This characteristic underscores the importance of managing the continuous and fast-changing flow of data [1].

In addition to volume and velocity, data flows originate from a wide array of sources (e.g., medical records, ecological research, consumer behavior) [3]. Variety refers to the structural heterogeneity within a dataset. Traditional methods lack the capacity to process large data streams in real-time in terms of volume, velocity, and variety. Such data requires innovative information processing techniques [4].

Data characterized by high volume, velocity, and variety introduces significant risks, including unpredictable errors and biases [3]. Therefore, veracity refers to the reliability of big data and the extent to which quality standards can be met. For instance, the emotional states of social media users are uncertain due to the volatile nature of emotions. Consequently, specialized tools and methods are needed to process and analyze uncertain data [4]. Without proper measurement and validation processes, this can lead to misleading or entirely false evidence for knowledge claims [3].

Additionally, the variability and complexity of data further heighten this risk. Variability refers to fluctuations in the rate of data flow; big data rates are often inconsistent, exhibiting periodic peaks and troughs. Complexity reflects the fact that big data originates from numerous different sources. This characteristic introduces challenges such as the need to correlate, harmonize, cleanse, and transform data from various sources [4].

Another defining characteristic, virtue and value, refers to the diverse meanings that different segments of society attribute to big data [2]. Emphasizing these characteristics is particularly important for advancing scientific knowledge and improving algorithms [5]. Beyond scientific value, data can also hold economic, political, ethical, and even emotional value. However, these values may not always align with the priorities of researchers. Institutions that own data processing techniques and fund data-related research may interpret the data in ways that reflect their own perspectives [3]. As a result, it remains a subject of debate whether big data will generate insights in the public interest or prioritize private interests.

In addition to these features, it is crucial that big data remains accessible, reinterpretable, and usable despite changes in archiving technologies. Given the rapid obsolescence of tools and techniques used in data generation and analysis, data infrastructures must be regularly updated to ensure long-term access. This is particularly important as data collected from various sources is often converted into digital formats that are suited for algorithmic processing and may be heterogeneous [3]. Creating highly relational data from high-speed, dynamic, and heterogeneous sources allows the data to become more meaningful than the sum of its parts.

The volume of data derived from these digital behaviors cannot be analyzed without machine learning techniques and cannot be fully understood without considering social theories that account for the multifaceted nature of behaviors [6]. Indeed, the strength of big data lies in its capacity to bridge different theoretical frameworks, methodological approaches, and research communities. This ability drives the continuous expansion of the boundaries of data-driven research logic.

Data-Driven Research Logic

In data-driven research, theoretical expectations are not considered the primary driver of the research process. Instead, social information and network theory provide a functional theoretical framework for understanding the complexity of digital interactions.

Social Information and Network Theory views individuals as nodes within social networks, emphasizing that the interactions and emotional diffusion between these nodes play a critical role in shaping social psychology. According to this perspective, digital interactions allow us to study how individual emotional states are transmitted and how collective reactions to social events emerge. Empirical studies based on social network data increasingly demonstrate the dynamics of these interactions and their spillover effects [8, 9].

While social networks enable individuals to share their thoughts and emotions, tracking the spread of ideas and emotions resulting from these interactions is more efficiently understood within the framework of network theory. For instance, a post about a significant event on social media can rapidly propagate through key, centralized individuals within the network and influence large masses. Network theory seeks to explain how this diffusion occurs, which nodes are most influential, and how thoughts and emotions disseminate through the network [9, 10].

In the digital age, one of the most significant aspects of big data analysis is understanding how emotional reactions spread to large populations and how these reactions influence social dynamics. Sentiment analysis and opinion mining play a crucial role in this process, as they analyze data to determine the emotional tendencies of individuals and communities towards particular issues. In sentiment analysis, the emotional tendencies of individuals or communities in digital data, such as text, audio, and images, are classified as positive, negative, or neutral, and the polarity of emotion is identified [4, 11, 12]. Unlike information, emotions and opinions are inherently subjective, making it essential to analyze as many opinions as possible [13].

These analysis methods, which employ techniques such as natural language processing (NLP) and machine learning, offer the ability to capture public feelings, opinions, and attitudes about social events. This facilitates the prediction of potential mass reactions. Sentiment analysis methods use knowledge-driven linguistic patterns and advanced statistical techniques to recognize sentiment and polarity from heterogeneous data types. However, these techniques may struggle to capture implicitly expressed opinions and emotions. Additionally, content may have polarity without an explicit opinion, or it may be challenging to distinguish between related and unrelated views in multi-topic materials. Addressing these issues is crucial to avoid drawing misleading conclusions [14].

However, detecting sentiment and opinion trends alone is insufficient. The real potential of big data is realized when it is used to inform decision-making. In this context, structured phases are required to transform diverse data into meaningful insights. These phases include acquisition and recording; extraction, cleaning, and annotation; integration, aggregation, and representation; modeling and analysis; and interpretation [4]. Structuring these stages effectively is essential to separate data from noise and to integrate data in various formats [13].

The reliability of the data-driven approach hinges on the effectiveness of the methods used to assess whether the patterns extracted from the data are meaningful. There is no guarantee that an algorithm trained to extract patterns from one dataset will perform equally well when applied to another dataset [3]. Algorithms vary greatly in terms of their mathematical structure and the conceptual frameworks on which they are based. These algorithms are designed to learn from new information input into the system. As a result, they have the capacity to adapt and evolve based on new data, enhancing their ability to analyze phenomena and predict future behaviors more accurately [3].

The next generation of deep learning models continues to evolve, with designs that better understand natural language structures, as well as psychological and ethical reasoning. The success of these models is directly proportional to the increase in data. As the volume and variety of data grow, so too does the success of these models in terms of their predictive capabilities [15].

The development of these techniques necessitates integrating social and psychological theories that focus on the nature of digital behavior. Big data should efficiently bridge different theoretical and methodological approaches [6]. A theory-centered perspective predominates in science, acknowledging the importance of methods, data, models, and tools in scientific research, but viewing them as means to reach accurate propositions about social phenomena. However, in data-driven research, theoretical expectations are often not seen as guiding the research process [3].

Big data provides a powerful foundation for seeking correlations rather than theoretical explanations. A large volume of data is sufficient for inductive inference, allowing patterns to emerge without the need for pre-existing hypotheses. This approach contrasts with traditional research logic, where hypotheses are tested to confirm or falsify theoretical models. In data-driven research, theoretical expectations do not guide the process; instead, empirical inputs dictate the direction of inquiry, and correlation replaces causation. A sufficiently large dataset enables statistical algorithms to discover patterns. Anderson (2008) described this shift as the “end of theory” [16].

However, in this context, Elliot and colleagues caution that big data analyses may encourage a casual approach to empirical research, such as “fishing” for correlations or lead to spurious results [3]. In big data, spurious correlations can occur when uncorrelated variables appear falsely related due to the sheer size of the dataset. Scientifically irrelevant variables may be wrongly correlated due to high dimensionality [4]. As a result, the inability of big data analysis to differentiate between spurious and meaningful correlations raises concerns about the validity and reliability of research findings.

Digital Divide

Big data research is also controversial due to its lack of representativeness of the general population. It is argued that big data often reflects the views and needs of certain social groups while marginalizing or omitting the perspectives of others. This exclusion has been criticized for leading to insights that are neither inclusive nor egalitarian [2, 17].

Individuals, groups, and geographic regions without access to digital information are absent from large datasets, which contributes to the digital divide. While two-thirds of the world's population owns mobile devices, one-third lacks digital access because they do not own a cell phone [18]. Educational opportunities, gender, ethnicity, economic status, generational differences, and the level of development in various countries are also factors that influence access to digital information [19]. Research has shown that those without digital access tend to belong to disadvantaged and vulnerable groups, such as the disabled, elderly, incarcerated, and unemployed [18].

The digital divide refers to the techno-determinist divide between those who have access to digital resources and those who do not. However, this perspective is reductionist. Beyond merely having or lacking digital tools, there are also advantages and disadvantages associated with the use of digital tools. Therefore, indicators of the digital divide include technological infrastructure, internet connectivity, user knowledge, and social support [18].

The digital divide is an extension of social inequalities, and it is addressed within the context of its economic, political, and ethical dimensions. There are more complex social, economic, and cultural factors at play than simply access to technology [19]. According to [20], unequal access to digital technologies results in algorithms and artificial intelligence systems that extract patterns of emotions and opinions from datasets, reinforcing socially constructed inequalities. In this way, predictive models built through data mining and decision-making processes exclude the emotions and opinions of those who lack digital access. Algorithmic discrimination may amplify forms of discrimination in everyday life and contribute to existing social polarizations. Therefore, any claims derived from big data analysis are constrained by the social, economic, and cultural representations that shape the data pool [3].

Contrary to techno-optimistic narratives that frame the big data revolution as a harbinger of transparency, democracy, and social equality, the digital divide between those who can access and utilize data technologies and those who cannot continues to grow. As a result of these divides, there is a lack of data on specific subgroups and geographic regions, further limiting the comprehensiveness of available data sources [3, 21].

Thus, the digital divide is intricately linked to the concepts of “data violence” and “data justice,” as it contributes to algorithmic discrimination and social polarization. In this context, the ethics surrounding data and algorithms become increasingly crucial and require careful examination.

Data Ethics

Big data research presents significant opportunities, but it also requires the adherence to ethical values and behaviors when recording, processing, and sharing data. The ethical challenges posed by data science are primarily organized around data ethics, algorithm ethics, and application ethics.

Data Ethics concerns issues related to the recording, processing, and use of data, as well as the ethical implications of algorithms (such as artificial intelligence and machine learning) and their applications (such as programming and coding). Data ethics addresses a broad range of issues arising from big data research, including concerns related to data violence, data justice, data philanthropy, and open data. Due to this expansive scope, data ethics adopts a holistic approach, avoiding narrow or ad-hoc solutions and addressing the impacts and consequences of data applications in a comprehensive framework [22].

Algorithm Ethics focuses on the responsibility and accountability of algorithms, particularly in machine learning applications. This includes concerns about the transparency, fairness, and potential biases inherent in algorithmic decision-making. Ethics of Practice pertains to the obligations and accountability of individuals and institutions responsible for data processes and policies. Together, data ethics, algorithm ethics, and application ethics aim to develop ethical codes that foster both the advancement of data science and the protection of individuals' and society's rights [22]. This framework encompasses a wide array of concerns, ranging from issues related to data collection and analysis to more specific ethical dilemmas, such as algorithmic bias, privacy violations, mass surveillance, anonymity, openness, consent, and user privacy [23, 24]. These concerns necessitate ongoing critical reflection on the ethical implications of data practices to ensure that data science is conducted responsibly and with due consideration for societal impact.

In this context, to create an ethical data ecosystem, several fundamental ethical principles must be established and safeguarded by ethics review boards and legislative frameworks. The first principle is ownership. Data ethics is grounded in the idea that individuals own their data. This means individuals should have the right to freely make decisions regarding their data and take responsibility for its use. The second principle is transparency in data processing. Transparency involves disclosing the methods used to create large datasets and the potential impacts of these processes. This is crucial for ensuring accountability and trust in how data is managed. The principle of consent is also central to the ethical processing of data. It requires individuals to be informed about who will use their data and for what purposes. This principle is deeply connected with the protection of privacy, as individuals have the right to control their personal information. The principle of confidentiality ensures that the privacy of individuals is maintained, and they should also be informed if their data is used for financial gain. The principle of openness of data is another important concept, particularly in promoting social good, progress, and auditability. For big data to be useful for societal benefit, it must be accessible and publicly available. This openness contributes to fostering scientific advancements and ensuring the transparency of algorithms [25,26].

Ballantyne (2018) further emphasizes several additional principles for responsible data use: social value, harm minimization, control, fairness, justice, reliability, transparency, and accountability. Social value refers to the potential of data to generate knowledge and societal benefits, with open data playing a crucial role in advancing scientific knowledge and the development of algorithms. However, the social value of data must be carefully balanced with ethical considerations to minimize harm—whether physical, economic, or psychological—to data subjects. In line with these considerations, control over data is a vital principle, ensuring that data subjects have authority over how their data is used. However, as many data uses occur without explicit consent, ensuring transparency and accountability has become a priority in contemporary data governance [5].

Ballantyne also argues that balancing these competing ethical values requires conscious negotiation and transparency. The demands of different stakeholders often conflict, so it is essential to carefully consider and balance their perspectives. Importantly, the voices and rights of data subjects should be incorporated into this decision-making process, ensuring that their interests are adequately represented and protected [5, 27].

Observing and upholding these fundamental principles is essential for building public trust in the use of data. In this context, developing ethical review committees for sustainable data management is crucial to protect society from the potential risks posed by big data research. Ethics Review Committees (ERCs) are traditional oversight mechanisms designed to ensure that research is conducted in accordance with ethical guidelines. However, the broad scope of big data research, issues related to data anonymization, and the involvement of multiple actors necessitate a re-evaluation of the effectiveness of ERCs in these areas [23]. Specifically, the varying ethical standards across different countries and committees lead to a lack of harmonization. The lack of clarity in the processes and decision-making mechanisms of Ethics Review Committees can erode public trust. Additionally, ERCs may fail to sufficiently consider the societal benefits when assessing the individual risks associated with big data research [23].

The issues specific to big data research pose new challenges for ethics review committees. Since big data projects are often not based on predefined hypotheses, they can involve unforeseen risks. Furthermore, while anonymization is considered the cornerstone of protecting individuals from potential harm, technological advancements have enabled the re-identification of individuals in pseudo-anonymous datasets, rendering anonymization efforts ineffective. This failure can leave individuals vulnerable to risks such as stigmatization or discrimination. Additionally, research using anonymized data is often exempt from ethical review [23].

Given these shortcomings, there is a growing need for interdisciplinary collaboration across fields such as computer science, data analytics, artificial intelligence, communication sciences, and philosophy to develop global ethical standards for big data research. Such collaboration would not only enhance the ethical dimensions of big data research but also increase public trust, promote more inclusive data policies, and support data-driven scientific progress.

Conclusion and Summary

The expansion of digital interactive networks has facilitated the creation of large datasets that reflect individual perceptions of social events, providing valuable insights into social, economic, and political processes. Big data plays a pivotal role in accelerating scientific research, shaping public discourse, and analyzing public psychology. By processing vast amounts of complex digital content, big data analytics enable real-time analysis of social trends and can simulate collective emotions during crises. This rapid development calls for interdisciplinary approaches. However, debates surrounding the extent to which big data reflects the real world, the impact of digital access inequalities, and the ethical regulation of big data applications continue to intensify.

This article explores the theoretical and methodological foundations of big data analysis, the social consequences of access inequalities, and the ethical challenges in data collection processes. While emphasizing the opportunities big data presents, it also highlights the risks, such as deepening social inequalities and data ethics violations. Data-driven research, which prioritizes correlations over theoretical inferences, facilitates the development of new deep learning techniques and inductive analysis methods. However, such analyses can lead to false correlations and misunderstandings. In high-dimensional datasets, unrelated variables may appear falsely correlated. Therefore, the reliability and effectiveness of big data analysis depend on the design of algorithms, the contextual accuracy of analyses, and integration with the social sciences. Multidisciplinary approaches are crucial to preventing false correlations and generating meaningful insights.

Despite the opportunities, big data research has been criticized for its lack of inclusivity and representativeness. Inequalities in access to digital information exacerbate the digital divide between individuals, groups, and regions. Disparities in data access across social groups lead to analyses that reflect the needs and perspectives of certain groups while overlooking others. This increases the risk of generating misleading insights. Consequently, big data analysis is framed within a context shaped by social, economic, and cultural biases. These disparities also result in data gaps for specific groups and geographic regions, limiting the comprehensiveness of analyses. In particular, failures in anonymization and the absence of international ethical standards expose individuals to risks such as discrimination and privacy violations. In this regard, concepts such as “data violence” and “data justice” are crucial for deepening the study of the social impacts of digital divides.

Data ethics and algorithm ethics are becoming increasingly critical in addressing these issues. Over time, addressing these challenges will require interdisciplinary cooperation. Contributions from fields like computer science, artificial intelligence, communication science, and philosophy will strengthen the ethical framework of big data research, enhance public trust, and promote more inclusive data policies. These efforts will protect individual rights while advancing data-driven scientific research.

References:

1. Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

2. Chen, W., & Quan-Haase, A. (2020). Big data ethics and politics: Toward new understandings. *Social Science Computer Review*, *38*(1), 3-9. https://doi.org/10.1177/0894439318810734

3. Leonelli, S. (2020). Scientific research and big data.

4. Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International journal of information management*, *35*(2), 137-144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

5. Ballantyne, A. (2018). Where is the human in the data? A guide to ethical data use. *GigaScience*, *7*(7), giy076. https://doi.org/10.1093/gigascience/giy076

6. Sapienza, A., & Lehmann, S. (2021). A view from data science. *Big Data & Society*, *8*(2), 20539517211040198. https://doi.org/10.1177/20539517211040198

7. Ferrara, E., & Yang, Z. (2015). Measuring emotional contagion in social media. *PloS one*, *10*(11), e0142390. https://doi.org/10.1371/journal.pone.0142390

8. Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788-8790. https://doi.org/10.1073/pnas.1320040111

9. Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., & Hołyst, J. A. (2011). Collective emotions online and their influence on community life. *PloS one*, *6*(7), e22207. https://doi.org/10.1371/journal.pone.0022207

10. Fan, R., Zhao, J., Chen, Y., & Xu, K. (2014). Anger is more influential than joy: Sentiment correlation in Weibo. *PloS one*, *9*(10), e110184. https://doi.org/10.1371/journal.pone.0110184

11. Dickinson, B., & Hu, W. (2015). Sentiment analysis of investor opinions on twitter. *Social Networking*, *4*(3), 62-71. https://dx.doi.org/10.4236/sn.2015.43008

12. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, *2*(1–2), 1-135. https://doi.org/10.1561/1500000011

13. Liu, B. (2017). Many facets of sentiment analysis. *A practical guide to sentiment analysis*, 11-39. https://doi.org/10.1007/978-3-319-55394-8_2

14. Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). Affective computing and sentiment analysis. *A practical guide to sentiment analysis*, 1-10. https://doi.org/10.1007/978-3-319-55394-8_1

15. Çetin, U. (2020). Kompleks Sistemler ve Veri Bilimi. *Yildiz Social Science Review*, *6*(2), 119-130. https://doi.org/10.51803/yssr.833992

16. Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired magazine*, *16*(7), 16-07.

17. Markham, A. N., Tiidenberg, K., & Herman, A. (2018). Ethics as methods: doing ethics in the era of big data research—introduction. *Social Media+ Society*, *4*(3), 2056305118784502. https://doi.org/10.1177/2056305118784502

18. Robinson, T. (2024). Digital Divide Experienced by Formerly Incarcerated Individuals and Its Effect on Their Reintegration into Society. *Open Journal of Social Sciences*, *12*(7), 258-278. https://doi.org/10.4236/jss.2024.127018

19. Livingstone, S., & Helsper, E. (2007). Gradations in digital inclusion: Children, young people and the digital divide. *New media & society*, *9*(4), 671-696. https://doi.org/10.1177/1461444807080335

20. Ragnedda, M., & Ragnedda, M. (2020). *Connecting the digital underclass* (pp. 85-104). Springer International Publishing. https://doi.org/10.1007/978-3-030-49079-9_5

21. Novak, R., & Pavlicek, A. (2021). Data experts as the balancing power of big data ethics. *Information*, *12*(3), 97. https://doi.org/10.3390/info12030097

22. Floridi, L., & Taddeo, M. (2016). What is data ethics?. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), 20160360. https://doi.org/10.1098/rsta.2016.0360

23. Ferretti, A., Ienca, M., Sheehan, M., Blasimme, A., Dove, E. S., Farsides, B., ... & Vayena, E. (2021). Ethics review of big data research: What should stay and what should be reformed?. *BMC medical ethics*, *22*(1), 51. https://doi.org/10.1186/s12910-021-00616-4

24. Häußler, H. (2021). Data Ethics Frameworks: Werte und Machtstrukturen im Spannungsfeld von Technologie und moralischem Handeln. *Information-Wissenschaft & Praxis*, *72*(5-6), 291-298. https://doi.org/10.1515/iwp-2021-2178

25. Davis, K., & Patterson, D. (2012). *Ethics of big data*. O'Reilly Media, Inc.

26. Zwitter, A. (2014). Big data ethics. *Big data & society*, *1*(2), 2053951714559253. https://doi.org/10.1177/2053951714559253

27. Hasselbalch, G. (2019). Making sense of data ethics. The powers behind the data ethics debate in European policymaking. *Internet Policy Review*, *8*(2). https://doi.org/10.14763/2019.2.1401

Список литературы:

1. Kitchin R. The data revolution: Big data, open data, data infrastructures and their consequences. Sage, 2014.

2. Chen W., Quan-Haase A. Big data ethics and politics: Toward new understandings // Social Science Computer Review. 2020. V. 38. №1. P. 3-9.https://doi.org/10.1177/0894439318810734

3. Leonelli S. Scientific research and big data. 2020.

4. Gandomi A., Haider M. Beyond the hype: Big data concepts, methods, and analytics // International journal of information management. 2015. V. 35. №2. P. 137-144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

5. Ballantyne A. Where is the human in the data? A guide to ethical data use // GigaScience. 2018. V. 7. №7. P. giy076. https://doi.org/10.1093/gigascience/giy076

6. Sapienza A., Lehmann S. A view from data science // Big Data & Society. 2021. V. 8. №2. P. 20539517211040198. https://doi.org/10.1177/20539517211040198

7. Ferrara E., Yang Z. Measuring emotional contagion in social media // PloS one. 2015. V. 10. №11. P. e0142390. https://doi.org/10.1371/journal.pone.0142390

8. Kramer A. D. I., Guillory J. E., Hancock J. T. Experimental evidence of massive-scale emotional contagion through social networks // Proceedings of the National Academy of Sciences. 2014. V. 111. №24. P. 8788-8790. https://doi.org/10.1073/pnas.1320040111

9. Chmiel A., Sienkiewicz J., Thelwall M., Paltoglou G., Buckley K., Kappas A., Hołyst J. A. Collective emotions online and their influence on community life // PloS one. 2011. V. 6. №7. P. e22207. https://doi.org/10.1371/journal.pone.0022207

10. Fan R., Zhao J., Chen Y., Xu K. Anger is more influential than joy: Sentiment correlation in Weibo // PloS one. 2014. V. 9. №10. P. e110184. https://doi.org/10.1371/journal.pone.0110184

11. Dickinson B., Hu W. Sentiment analysis of investor opinions on twitter // Social Networking. 2015. V. 4. №3. P. 62-71. https://dx.doi.org/10.4236/sn.2015.43008

12. Pang B., Lee L. Opinion mining and sentiment analysis // Foundations and Trends® in information retrieval. 2008. V. 2. №1–2. P. 1-135. https://doi.org/10.1561/1500000011

13. Liu B. Many facets of sentiment analysis // A practical guide to sentiment analysis. 2017. P. 11-39. https://doi.org/10.1007/978-3-319-55394-8_2

14. Cambria E., Das, D., Bandyopadhyay S., Feraco A. Affective computing and sentiment analysis // A practical guide to sentiment analysis. 2017. P. 1-10. https://doi.org/10.1007/978-3-319-55394-8_1

15. Çetin U. Kompleks Sistemler ve Veri Bilimi // Yildiz Social Science Review. 2020. V. 6. №2. P. 119-130. https://doi.org/10.51803/yssr.833992

16. Anderson C. The end of theory: The data deluge makes the scientific method obsolete // Wired magazine. 2008. V. 16. №7. P. 16-07.

17. Markham A. N., Tiidenberg K., Herman A. Ethics as methods: doing ethics in the era of big data research—introduction // Social Media+ Society. 2018. V. 4. №3. P. 2056305118784502. https://doi.org/10.1177/2056305118784502

18. Robinson T. Digital Divide Experienced by Formerly Incarcerated Individuals and Its Effect on Their Reintegration into Society // Open Journal of Social Sciences. 2024. V. 12. №7. P. 258-278. https://doi.org/10.4236/jss.2024.127018

19. Livingstone S., Helsper E. Gradations in digital inclusion: Children, young people and the digital divide // New media & society. 2007. V. 9. №4. P. 671-696. https://doi.org/10.1177/1461444807080335

20. Ragnedda M., Ragnedda M. Connecting the digital underclass. Springer International Publishing, 2020. P. 85-104. https://doi.org/10.1007/978-3-030-49079-9_5

21. Novak R., Pavlicek A. Data experts as the balancing power of big data ethics // Information. 2021. V. 12. №3. P. 97. https://doi.org/10.3390/info12030097

22. Floridi L., Taddeo M. What is data ethics? // Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2016. V. 374. №2083. P. 20160360. https://doi.org/10.1098/rsta.2016.0360

23. Ferretti A., Ienca M., Sheehan M., Blasimme A., Dove E. S., Farsides B., Vayena E. Ethics review of big data research: What should stay and what should be reformed? // BMC medical ethics. 2021. V. 22. №1. P. 51. https://doi.org/10.1186/s12910-021-00616-4

24. Häußler H. Data Ethics Frameworks: Werte und Machtstrukturen im Spannungsfeld von Technologie und moralischem Handeln // Information-Wissenschaft & Praxis. 2021. V. 72. №5-6. P. 291-298. https://doi.org/10.1515/iwp-2021-2178

25. Davis K., Patterson D. Ethics of big data. O'Reilly Media, Inc., 2012.

26. Zwitter A. Big data ethics // Big data & society. 2014. V. 1. №2. P. 2053951714559253. https://doi.org/10.1177/2053951714559253

27. Hasselbalch G. Making sense of data ethics. The powers behind the data ethics debate in European policymaking // Internet Policy Review. 2019. V. 8. №2. https://doi.org/10.14763/2019.2.1401

Работа поступила в редакцию 08.11.2024 г.

Принята к публикации 12.11.2024 г.

Ссылка для цитирования:

Yoğurtçu G. Big Data Analytics and Ethical Responsibility in the Digital Age // Бюллетень науки и практики. 2024. Т. 10. №12. С. 465-475. https://doi.org/10.33619/2414-2948/109/62

Cite as (APA):

Yoğurtçu, G. (2024). Big Data Analytics and Ethical Responsibility in the Digital Age. *Bulletin of Science and Practice, 10*(12), . 465-475. https://doi.org/10.33619/2414-2948/109/62

