Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 06, June-2023

The moral risks and correction for the big data ecosystem [*]

[1]Chengfei Wei, [2]Zhaodan Xie*

[1] Lecturer, Academy of Marxism, Nanjing University of Finance and Economics, Nanjing, China;

P.h.D, Jindal Institute of Behavioural Sciences, O.P. Jindal Global University, Haryana, India

2 (Corresponding Author) Lecturer, Academy of Marxism, Nanjing University of Finance and Economics, Nanjing, China

Abstract: Big data ecosystem, which promotes knowledge sharing and science & technology innovation, improves the core competitive advantage of participants, and overcomes the barrier of time and space in social interaction, has its negative aspects. Public anonymity freedom has nearly been destroyed because of privacy disclosure. Since the re-identification of trade secrets, the fairness of competition has been broken. Information control differentiation has negative impacts on the right to know and choose. Profit from data acquired by huge companies influences the interests of consumers. Monopoly in data hinders the autonomy of inheritance. Solidification of the decision-making course impacts exploration and innovation. All these above violate basic human rights and risk economic goals. In order to realize a balanced and positive circular for the interests of each side in the Internet of Things, it is necessary to establish a moral correction system based on the principles of human rights, harmless, and responsibility.

Keyword: Big data; Ecosystem; Moral Risks; Moral Correction

I. INTRODUCTION

Big data ecology is the best platform so far, which provides accurate data analysis and reliable real-time prediction, based on multi-type big storage and effective information refining, contributing to the brand new breakthrough in productivity, technology, and social contact structure.[1] On the one hand, integrating physical monitoring system and online positioning system, gradually, wide-area information flow and transformation and related decision-making will be employed in big data marketing systems and customers relationship management systems to improve operating efficiency in the industries like finance, manufacture, retail, and advertising technology, healthcare and so on. On the other hand, with high level, high speed, and diversity, big data mining starts an epoch in which added value information resources that can be identified and utilized in the academy, education, and social interaction. Therefore, the model of thinking and behavior will be improved in all directions.

In the era of information explosion, massive data has already become raw material for production and a new source of huge economic and social value.[2] While smart analysis has changed traditional commercial model, promoted a brand new change of motivation for scientific study from assumption to data technology of distributed storage and computation[3], and perfectly coped with political, cultural, and ecological issues, it also reveals some serious problems, including only benefit-seeking commercial institution, too emotional mining subjects, poor recognition and management of data right, association analysis abused by public authorities and mala fide third party through correlation analysis technology. Moral correction for big data based on the values of fairness, efficiency, mutual benefit, and responsibility is necessary to be put into practice. Consequently, disadvantages that may impede the sound development of the big data industry, including over-monetization of information, little protection for customers' privacy, and stylized choice of behavior, could be overcome.

II. MORAL RISKS FOR THE BIG DATA ECOSYSTEM

[*] The thesis is the achievement of the University Philosophy and Social Science Research Project of Jiangsu Province (2022SJSZ0095) and the Humanities and Social Science Research Project of the Ministry of Education (22YJC860020).

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 06, June-2023

In the context of the data tsunami, the application of big data has extended from simply healthcare statistics to business decisions and even political competition.[4] The big data ecosystem, where the participants are interactive and restricted mutually, has many risks. For example, infringement on privacy and trade secrets by excessive data mining, monetizing unidentified private data, and data monopoly cannot be controlled, which seriously hurts the balance of interests in the big data ecosystem.

2.1 Anonymity freedom hurt by privacy disclosure

The key to the success of the Internet is anonymity in producing and transferring information [5]. The Internet provides the public with an important way to convey their free will, though anonymity may lead to the increase of false judgment and collective violence. There is a negative correlation between data accumulation and privacy protection. With the rising of meta-data collected, it would be easier to uncover privacy and secrets through correlation analysis. Big data technology can recognize true information and thoughts according to massive structured and unstructured data. Reconstructing freedom of expression under the spotlight by exposing 'masks' such as personal sexual orientation, working willingness, health condition, etc, which the obligee used for self-protection. Essential values of the Internet, including the voice of free will, the source of public conscience and creativity, might be damaged.

2.2 Re-identification of trade secrets breaks the competition fairness

In the big data ecosystem, the re-identification which contains continuous surveillance and complex association analysis adds the difficulty in traditional trade secrets protection.[6] In the age of big data, the leaked secret is general information rather than classified one. As advantaged companies in technologies and resources in the area of data mining, they make inductive reasoning on the data collected from newsgroups, electronic bulletin boards, massive scattered data from public document transmission, and lots of related variables. As a result, a considerable number of trade secrets, including ingredients of products, designs, and customers' information could be revealed, contributing to the destruction of competition fairness which is one of the traditional commercial moral risks.

2.3 Information control differentiation against customers' right to know and choose

With data permeating every business and functional area, the management and storage of terminal information are commonly carried out on remote servers[7] and would carry out automatic segmentation before dynamic allocation. Service suppliers who unilaterally control the segmentation technology which can reduce the processing time and transmitting difficulty would not only escape informing the data obligee of the physical location of the data storage device, security condition, and processing details,[8] but also evade their responsibility in the advantages of customers' trust and technology. Still, the vulnerability of "Safe Harbor Rule" contributes to data controllers' infringement. Actually, the wide-area technology of data aggregation which can recognize potential correlations among a variety of data is employed by many data mining companies. Using analysis tools, including Tableau, BigQuery, etc, companies can trace data streams without notification to owners, which seriously hurts the right to know and to choose data obligee.

2.4 Data profit against the earning right

Many global Internet companies, like Amazon, Baidu, and Tencent, are active supporters of big data strategy. Since big data strategy not only promotes a sharp increase in the volume of data resources and commercial value of data but rise market requirements towards prediction and market confidence in services related to big data as well. For instance, data analysis is the evidence for Google to locate targeted advertisements.[9] They also use big data to improve Google Translate with nearly zero cost. Big data means big profit.[10] Companies including Tableau and Marketo continuingly trace access data and analyses user preferences and intentions. Therefore, the model they build is sold to release the profit potential.[11] For example, based on massive results of analysis of data, data mining companies optimize their strategy to increase earnings and lower the cost. Moreover, based on long-term analysis of citizens' data, governments develop the budget and reduce juridical costs.[12] Data mining subjects rejected to inform customers' value increment by big data and share it, which is an infringement on data property.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 06, June-2023

2.5 Data monopoly against autonomous inheritance

Deep, wide-area sharing of data is the very charm of the age of information. Meanwhile, it is also the foundation of the long-term development of big data technology. Mining could not be proceeding without massive metadata. But with the excuse of security need and privacy protection, diversified big data companies usually reject to share the massive data they collected. Few individuals, commercial institutions, and even education research institutions can have free access. Take Facebook as an example, billions of access data and millions of photos from the users are collected on their server every day. On the contrary, because of disadvantaged in technology and abundant capital chain, SMEs lacking support of metadata cannot establish and maintain cross-platform infrastructure construction. Thus the comprehensive treatment mechanism for massive data storage and mining would be an uneconomical choice. Furthermore, making a profit on targeted advertisements and app sales would be impossible. Quitting is the only choice for them. For instance, Google is a giant which cannot be overtaken by any competitors, as it has a long product line, including a search engine, Android, Google Chrome, Gmail, Youtube, Google Maps, Google+, Google Analytics, Google Apps, etc. Similarly, the oligopoly market is developing globally. With the development of giants of IBM, Google, Amazon, Baidu, Tencent, and Alibaba to remain their market positions, they establish rules to restrain data from flowing freely.

2.6 Solidification of decision-making against exploration and Innovation

On the basis of a great number of meta-data and efficient relations mining, big data prediction reduce mistakes and improve the processes of analysis, exploration, and decision-making through millions of hypothesis and validation. While a variety of decision-making behavior, based on big data analysis, can make some financial and social profit, there are still negative impacts on technical innovation, fair competition, and personal free development. The projects would predict the value cycle of products and services. And because of tremendous investment in these low-risk projects, the development opportunities of creative startups would be deprived. Everyone has the right to realize dreams and to pursue a better life. On the way to success, there may also be many chances and difficulties which both cannot be predicted by big data analysis. The wide application of big data prediction in personal decision-making, from the choice of the canteen, and work prospects to health care service, contributes to personal strivings becoming solid calculating functions [13]. As a result, there would be some negative impacts on social diversity and the right to choose.

III. MORAL CORRECTION SYSTEM FOR THE BIG DATA ECOSYSTEM

The industry of big data provides positive promotion to the productivity and competence of enterprises.[14] Besides, the cost of public service is getting lower and the marginal revenue is increasing. Lack of regulation is the main reason which causes market chaos, increases the gap between big data theory and practice, and orderly development of related markets. In order to address the burning problem, the perfect regulation system needs to be built. Despite the US surveillance program, PRISM warned us of the importance of data protection, after which we accelerate the procedure of legislation concerning data globally, like promoting the newly-revised European Union Data Protection, which is the landmark of protecting internet privacy strictly. Defects of legislation, including certainty and complex juridical procedure, make the endeavor on ordinary disputes related to information limited. And the administration, law enforcement, and judicial activities are threatened by the conflict raised between the national data border and the global cyber-world, for which a scientific and reasonable moral correction system is necessary to establish. Exploring the ideal balanced point between the development of business and moral expectation for the whole social big data application by determining the basic value goal of big data and balancing the interests of all parties. Finally, the virtuous cycle of the big data ecosystem, political stability, economic growth, and social harmony can be realized.

3.1 Respecting rules of moral correction

3.1.1 Respecting the right to know

With the increasing appreciation of personal data property recognition, data processing activities are increasingly required to narrow the scope of objects and to be specialized. Traditional big data mining which exacts and gathers data from a narrow resource pool is easy to judge and regulate. Big data mining, which is beyond the traditional way of access, storage, management,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 06, June-2023

and analytical data integration,[15] gets unknown but useful information from massive meta-data. The scattered data separation mode has been changed. All these developments have raised the inequality between the data controller and the owner. Furthermore, owners have no access to their sensitive personal data processing actually. Sunshine is the best fungicide. Owing to advantages in technology and the control of the resource, the data mining subjects are obliged to ensure transparency in data collecting, storing, and processing. Moreover, it is also the controllers' responsibility to notify the owner of the purpose, process, risk, and profit in a popular and easy way and to improve the owner's actual control of personal data in case of illegal use by all parties.

3.1.2 Respecting consent and selection right

The obligee of data ought to have basic data rights, including deciding whether and how to share sensitive data independently without outside interference. Obligees will no longer have access to storage and processing to protect their personal data for the complexity of distributed storage and processing. Big data enterprises should take all kinds of methods to protect data privacy during the whole process of data recombination and polymerization, especially safely, and secretly storing important personal information like birth certificates, marriage certificates, passports & visas, and wage income. So as to truly give the information disclosure freedom back to the obligee. Meanwhile, respecting consent and selection right is an important way to ensure the obligee's freedom, which means that the obligee not only has the right to participate in data life but also that no barriers should be set in transferring personal data between servers. For example, circulation means fortune. No service providers can put up dominant or recessive barriers to prevent users from removing data. Besides, although there are side effects of principles of purpose limitation and data minimizing on data mining, this may reduce business interests. Free choice plays a necessary role in human development and our society. Therefore, the data processed should be minimized without consent.

3.1.3 Respect the right to be forgotten

Sharp development of big data technology extraordinarily lowered the cost of digital storage and global barrier-free access to digital memory. Thus society enters a new era of permanent memory. Interleaving unstructured data in a massive anonymous database allows companies re-identifying massive personal data, including credit, marriage & family situation, and working experience, by which individuals' chance to restart may be deprived. So individuals should have the right to erasure under certain conditions. Thus, in the big data ecosystem, thorough erasure of data should be ensured when asked. And the related analysis result should be prevented from being uncovered disorderly.

3.2 Harmless Activities

Two aspects are involved in harmless activity, which are gaining and maintaining. Gaining indicates that processing made should benefit to realize the profit maximization for the data owner. And maintenance represents that the behavior of big data participants should at least not hurt others' profits. In the ideal big data environment of zero conflict between the free flow and processing of data resources, the almost zero-cost data operation can benefit all parties involved. However, in reality, to maximize business interest, it is popular for data mining companies to ignore basic human rights. All the participants in the big data ecosystem should realize that life safety, freedom needs, and privacy interests, which are basic human rights of data owners should still be respected. Moreover, the collection of data should be based on legal reasons and can't be used for other purposes. And data controller must notify the owner of the processing without threatening the owners' basic rights.

3.3 Identifiable Responsibility and Accountability

Big data enterprises, authorities, and other parties involved in data collection, storage, and processing are responsible to keep the big data ecosystem in good circumstances, in order to escape the prejudice and unfair treatment caused by inequality of speech rights. A series of specific measures can be taken to improve the current situation. To ensure transparency, fairness, and order in data mining, the data controller should establish prevention from re-recognition without the license by ameliorating the identification mechanism, increasing the tracing barrier, and improving key technology. And authority has accountability in regulating data mining to reduce discrimination and unfairness caused by inequality in voice. Besides, authorities should be banned to participate in data mining, despite data mining may reduce the budget of investigation. Moreover, other participants should not infringe on others' rights during the massive data operating process.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 06, June-2023

Self-discipline is also vital for the improvement of the moral correction system in the big data ecosystem. An activating mechanism should be added to the comprehensive standard and review system. To convey that the principle of data regulation improves self-discipline, training on rules of moral correction should be provided. Meanwhile, the ombudsman and anonymous impeachment should be set to minimize the negative effects of the infringement.

CONCLUSION

The big data ecosystem provides an unprecedented epoch to employ information and data. Volume, stream, and analysis are important factors in reviewing the strength of a state. Scientific and efficient big data strategy focuses not only on brand-new economic growth points of massive data analysis but also involves many aspects such as development, transmission, management, and use in developing an ideal big data strategy to balance the interests of all parties. Still, there are a fair distribution of resource through continuous improvements and publicity of rules of moral correction. Our cyber-world should be built with a stable meta-date stream, a safe resource pool; efficient algorithms and harmless predictions on innovation consist to prevent artificial intelligence from being out of control and loss of freedom. The purpose of sustained development of big data technology is to predict potentials, instead of telling answers.

REFERENCES

[1] Jordan Robertson, The Health-Care Industry Turns to Big Data, available online at: http://www.businessweek.com/printer/articles/26016-the-health-care-industry-turns-to-big-data, accessed on 21 June 2023.

[2] Omer Tene, Jules Polonetsky, Privacy in the Age of Big Data: A Time for Big Decisions, available online at: http://www.stanfordlawreview.org/online/privacy-paradox/big-data, accessed on 21 June 2023.

[3] Chris Anderson, The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, available online at: ttp://www.wired.com/science/discoveries/magazine/16-07/pb_theory, accessed on 21 June 2023.

[4] Thomas Edsall, Let the Nanotargeting Begin, available online at: http://campaignstops.blogs.nytimes.com/2012/04/15/let-the-nanotargeting-begin/, accessed on 21 June 2023.

[5] Karthick Ramachandran, Thomas Margon, Mark Perry, Clarifying Privacy in the Clouds, available online at: ttp://www.ivir.nl/publications/margoni/PSICTLAES_2011.pdf, accessed on 21 June 2023.

[6] Paul Ohm, Broken Promises of Privacy: Responding to the SurprisingFailure of Anonymization, UCLA Law Review, 2010(57), p.1701.

[7] Arvind Narayanan, A CriticalLook at Decentralized Personal Data Architectures, available online at: http://arxiv.org/abs/1202.4503, accessed on 21 June 2023.

[8] Aleksandar Hudic, Data Confidentiality using Fragmentation in Cloud Computing, International Journal for Communication Networks and Distributed Systems, 2012(1), p.2.

[9] John Markoff, How Many Computers to Identify a Cat?, New York Times,Vol.B1 June 2012.

[10] Ben Rooney, Big Data' s Big Problem: Little Talent, available online at: http://blogs.wsj.com/tech-europe/2012/04/26/big-datas-big-problem-little-talent/?mod=google_news_blog, accessed on 21 June 2023.

[11] Rebecca Greenfield, Facebook Now Knows What You're Buying at Drug Stores, available online at: http://www.theatlanticwire.com/technology/2012/09/facebook-tracking-you-drug-store-now-too/57183, accessed on 21 June 2023.

[12] Amanda Conley, Anupam Datta, Sustaining Privacy and Open Justice in the Transition to Online Court Records: A Multidisciplinary Inquiry, Maryland Law Review, 2012(71), p.772.

[13] Steven Overly. Mobile Health Apps Prompt Questions about Privacy, available online at: http://www.washingtonpost.com/business/capitalbusiness/mobile-health-apps-prompt-questions-about-privacy/2012/04/27/gIQAk17FqT_story.html, accessed on 21 June 2023.

[14] Erik Brynjolfsson, Strength in Numbers: How Does Data-DrivenDecisionmaking Affect Firm Performance?, available online at: ttp://ssrn.com/abstract=1819486, accessed on 21 June 2023.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 06, June-2023

[15] James Manyika, Michael Chui, Brad Brown, Big Data: The Next Frontier for Innovation, Competition, and Productivity, available online at: http://www.mckinsey.com/insights/business_technology/big_data_the_next_

[16] frontier_for_innovation, accessed on 21 June 2023.