Journal of Humanities and Social Sciences Studies
ISSN: 2663-7197
DOI: 10.32996/jhsss
Journal Homepage: www.al-kindipublisher.com/index.php/jhsss

JHSSS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Research on Personal Data Privacy Security in the Era of Big Data

Xuanting Wu[1] and Yi Chen[2] ✉
[1]*School of Information, Guangdong University of Finance & Economics, Guangzhou, China*
[2]*College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen, China*
Corresponding Author: Yi Chen, **E-mail**: 2695868600@qq.com

| ABSTRACT
Big data privacy security has become a hot research topic in contemporary society. Based on the data relevance and life-cycle in the era of big data, this paper analyzes the causes of security problems in China's data privacy. It puts forward suggestions from three aspects to provide references for subsequent research. Based on the current research progress, this paper first sorts out the definitions of data privacy and data privacy protection, then summarizes the causes of privacy security from the perspectives of technology and management and reveals the consequences of data privacy security issues. The demonstrated results trigger an insight into the solution strategy of data privacy problems and offer suggestions for solving problems from the perspective of management based on the data life-cycle model. Finally, starting from other stages of the data life-cycle and the application scenarios of big data, this paper looks forward to the future research direction. This study found that the present study needs to focus on the combination of system and technology, the improvement of laws and regulations, and the data life-cycle model in both technical and institutional management.

| KEYWORDS
Data privacy, privacy security, data protection technology, protection system, data Life-cycle

| ARTICLE INFORMATION
ACCEPTED: 30 August 2022 **PUBLISHED:** 06 September 2022 **DOI:** 10.32996/jhsss.2022.4.3.24

1. Introduction

Since the development of the computer industrial revolution (also known as the third industrial revolution), the continuously accumulated information is retained in the virtual world in an exponentially explosive growth trend, which promotes the rapid development of information technology and then drives all fields of society into the era of big data (Jin, 2020). As the core technology of big data, the data management system has the advantages of high compatibility, strong distribution, and excellent memory. It is used to process big data with large capacity, multiple types, rapid change, and low quality (Du et al., 2019). Under the processing of a data management system, big data has favorable properties such as relevance, non-expendable, sharing, and fidelity. These unique properties of big data make data gradually become the "contemporary oil" in the mind of more and more people (Wen, 2010; Yan, 2020). While the value of data is constantly explored, more and more data infringement cases are loaded into the database, and data privacy security has gradually become the focus of people's attention.

In recent years, research on big data privacy has been carried out mainly from protection countermeasures, contradictory relations, technical support, etc. First is the Protection Countermeasures analysis of personal privacy infringement under the digital background. This research refers to the research direction of providing reasonable suggestions for data privacy protection countermeasures by being familiar with the complexity of data privacy infringement and re-parsing personal sensitive data (Zhou & Xu, 2015; Jiang, 2019) under the influence of the diversification and sophistication of infringement caused by big data. The second is research on personal data privacy protection and data sharing behavior. It refers to the research direction of tracing the source of data privacy security issues by analyzing the objective contradiction between data sharing and privacy protection (Tian

& Huang, 2014; Zhang & Zhu, 2014) and provides reference materials for solving data privacy issues. Third, the expansionary theory of data privacy and the research on the expected positioning of rights. It refers to the research direction of systematically improving the data privacy protection and countermeasure system based on existing privacy protection laws and regulations (Wang & Zhao, 2015; Zhang, 2015) and then supplementing the content of data privacy and data privacy protection from a theoretical perspective. The development of the three research directions of data privacy protection has promoted the improvement of data security systems in many industries, announced the progress of digitization of relevant laws and regulations, and is of great significance to the development of digital civilization.

Data privacy research is still in its infancy. From the dimension of time, the research on data privacy is 2-4 years later than the research on the development and application of big data on average. Hence, the research on big data privacy is insignificant compared to the research on the development and application of big data. The amount of research on the disadvantages of big data is far less than the amount of research on the advantages, leading to the current data privacy research is not enough to deal with the complex problems in the application. In addition to the small amount of research, immature research on data privacy is also reflected in the following three aspects: First, the academic community has not yet formed a unified interpretation of data privacy (Peng, 2021). Second, most studies on new data protection technologies have not been transferred to the stage of practical application (Wang et al., 2020). Third, there have been no systematic suggestions on data privacy protection strategies (Zhu & Li, 2021). The above phenomenon shows that data privacy research only stays in specific areas and lacks the macroscopic overhead of its complete system. Therefore, this paper will summarize recent research on privacy data. First, the definition of data privacy and data privacy protection are sorted out in section one. Then the causes and impacts of new data privacy problems are summarized in section two. Finally, in the third section, this paper puts forward suggestions to solve the big data privacy security problem from different perspectives, hoping to provide some references for the follow-up research.

2. privacy and privacy protection in the context of the era of big data new interpretation

Data privacy and protection are indispensable to the research on data privacy security. A basic understanding of data privacy and protection is the premise of studying data privacy security. The following will discuss the two aspects of data privacy and data privacy protection to study data privacy security issues and focus on the differences and connections between them and traditional privacy.

2.1 Data Privacy and Data Privacy Rights

Data privacy has not only the complex nature of traditional privacy but also the other unique nature of digitized content. With the rapid development of big data, the connotation of privacy is constantly transmuting. Traditional privacy interpretations have developed in various fields such as psychology, sociology, economics, and law, and the era of big data has enriched its explanations from the information systems perspective. Allen (2004) proposes that "control of all one's information is a privacy," which links privacy and information as an early explanation for data privacy. There are many interpretations of data privacy in the academic community and different views. Most of these scholars believe that data privacy is part of the sensitive data that an individual or organization is not allowed to be known to outsiders, including the connection that can be expressed after the data has been processed, such as the browsing records of individuals and the financial status of the company (e.g., Huang et al., 2015; Meng & Zhang, 2015; Xu & Hu, 2019). Compared with traditional privacy, the unique feature of data privacy is that it includes information and data that can be inferred, which supplemented the new privacy connotation and gradually became a new paradigm that provides for data privacy (Lu, 2018; Ren et al., 2022). This expansive change in privacy connotation will lead to changes in related rights, and privacy rights closely related to privacy connotations will also develop new theories and rights positioning.

Privacy and privacy rights are closely related, and the change in privacy data has led to expansionary variations in the connotation of privacy. From the jurisprudential perspective, privacy is essentially a right. This claim can be traced back to the 1890s when Louis Brandeis and Samuel Warren (1890) proposed in the United States that the right to privacy is a right to be alone. In general, although privacy and privacy rights cannot be confused, in the study of privacy security issues, the two are inextricably linked, and the content of privacy increases will even directly affect the scope of privacy rights protection (Gu & Fan, 2018). The range of both, privacy is secret information that individuals do not want to disclose, while data privacy refers to the digitized form of this part of confidential information (Gu, 2021). In essence, privacy and data privacy are informational, and data privacy is the product of expressing private details through binary code. And this closely linked relationship has prompted the expansion of the scope of the object of privacy rights to all data that can be directed to infer or identify personal information (Wang & Yang, 2017; Shang, 2020). All in all, privacy is a right at some level, and its data-driven reform indirectly affects the positioning of its rights, which inaugurates new connotations about data privacy. To research security issues of data privacy rights, it is necessary to have a preliminary understanding of data privacy rights.

The era of big data has given new attributes to data privacy and expanded the theoretical boundaries of traditional privacy, which makes it more difficult to protect data. The conventional right of confidentiality refers to citizens' right to control information inherent in individuals, does not harm the public interest, and is unwilling to be disclosed (Wang, 1994). Although similar to traditional privacy rights, the subject of data privacy rights is generally a natural person whose goal is also to protect vital personal interests (Meng & Zhang, 2015). Still, with the continuous expansion of the definition of privacy rights in the era of big data, data privacy rights have gradually derived new attributes. First, the value weight of privacy rights from the economic perspective has increased after entering the age of big data. This change in property attributes stems from the ability of privacy to be given value creation in the era of big data. Due to the inference and practical nature of private data, the proportion of personal data as a property function has increased compared with the traditional privacy focus on personality rights (Mao, 2019). Second, compared with conventional privacy rights, the object of data privacy rights has the characteristics of the data life-cycle, data type diversification, and inferred transmission relationship between data, thereby expanding the scope of protection of data privacy rights. Third, due to the characteristics of difficulty in identifying the power boundary, difficulty in controlling the consequences of infringement, and difficulty in applying current laws and regulations (Meng & Zhang, 2015), compared with traditional privacy, the protection of data privacy is much more difficult. The above three new changes come from the technological development of the information age and the big data era. The derived data privacy and data privacy rights put forward higher requirements for privacy protection.

2.2 Data Privacy Protection and Protection Technologies

Data privacy protection needs to focus on the dynamic balance of privacy confidentiality and data availability, which is different from traditional privacy protection. Legal privacy protection aims to protect private information and prevent personal privacy from being maliciously stolen (Zhu & Zeng, 2021). Compared with the original privacy protection, data privacy protection aims to minimize the risk of private data leakage while reasonably grasping data availability. Compared with the two protection purposes, data privacy protection seems more likely to be used as a balancing mechanism. The main reason for this discrepancy is that the privacy protection object includes two dimensions. One of the dimensions is the protection of the individual, which refers to the safety of the individual's private data from leakage and the provision of technical operation security support for the individual when the user accesses or uses the data information. Another dimension is the protection of data, which is the use of relevant technologies to protect personal privacy data and maintain the regular availability of data when transmitting or encrypting data (Qian, 2013). Based on the above dimension of protection requirements analysis, data privacy protection is based on traditional privacy protection to put forward more refined requirements. The upgrade of data privacy protection requirements has increased the difficulty of data privacy protection, and the protection technology closely related to privacy protection also needs to be innovated.

The difficulty of big data privacy protection catalyzes the secondary growth of data privacy protection technology, and data protection technology can be divided into different types according to different angles. First, the methods of protecting data privacy by technology are divided into active and passive types (Zhang, 2019). Active type technology refers to adding identification attributes to personal data so that the data can be proved to be the source of the data in the use stage; passive method technology is currently widely used, mainly refers to the encryption, reconstruction, and control of personal privacy data, which to interfere with the identification and inference of privacy data during the data use phase. Secondly, based on the life-cycle model of data privacy, data privacy protection technology can be divided into protection technology in the data collection stage, protection technology in the data release stage, and protection technology in the data access stage. Among them, the protection technology for the data release stage is currently a hot topic in academic research (Wang et al., 2020), which can be roughly divided into three categories: grouping technology (Samarati & Sweeney, 1998), encryption protection technology (Qiu & Li, 2018) and release distortion technology (Li et al., 2013). At the same time, these three types of technologies have been extended to privacy protection technologies in other life-cycle-based stages, supplementing the technical gaps in the protection methods at the acquisition and access stages. These complementary and constantly derived new technical theories have become the development trend of technology research of data privacy protection.

3. Data privacy security issues under the contradiction between "disclosure" and "protection."

The contradiction between developing big data technology and ensuring data privacy security is challenging to avoid. The development of big data technology depends on data disclosure. However, the practice of disclosing data will objectively infringe the privacy of organizations or individuals, which makes the relationship between the development of big data technology and the protection of privacy implicit deviation and contradiction, and producing the paradox of "data disclosure" and "privacy protection" (Yan, 2020). This paradox is widely present in the privacy security issues in the era of big data, which dramatically increases the difficulty of privacy protection. Furthermore, this contradictory relationship may become one of the motivations for privacy theft. On the one hand, enhancing the property attributes of data privacy has increased the proceeds of stealing private data.

On the other hand, data disclosure makes private data relatively easy to obtain, coupled with the objective paradox that creates activity space for the theft of privacy. The combined effect of the two reduces the cost of illegal data theft. Faced with the temptation of such economic benefits, more and more people have chosen the unlawful act of stealing private data. Therefore, to protect personal privacy and develop science and technology simultaneously, many scholars have researched the causes of privacy security issues around the contradiction between the two (e.g., Tian & Huang, 2014; Loukis, 2016; Zhu & Li, 2021). Studies have shown that taking data and data subjects as reference objects can mainly attribute privacy security issues to two aspects: technology and management.

3.1.1 Privacy data definition technology and data protection technology

The traditional privacy data identification technology has lost its original role due to the popularity of big data, failing sensitive data screening technology. At present, privacy data identification technology divides highly recognizable data such as the name, residential address, and education of data subjects into sensitive data and data with low recognition, such as birthdays and hobbies, into ordinary data. With the development of network communication technology and big data technology, the above two will become possible to be interconnected and transformed: attackers can collect ordinary data of objects, take advantage of the transmutation of privacy content, the ambiguity of privacy boundary, and obtain sensitive data through link attack models, which makes traditional privacy definition technology gradually lose the ability to isolate and effectively distinguish standard privacy (Li et al., 2010; Yan, 2020; Zhu & Li, 2021). Not only that, but the rapidly evolving technologies of data collection at scale, modern data access and extraction technology, and late data model analytical technologies, as well as the barriers to data sharing that these technologies have spawned (Meng, 2015; Yan, 2020), which not only makes it more difficult to define and protect privacy but also widens the gap between individuals and enterprises and governments with massive databases (Zhang, 2019; Zhu & Li, 2021). Traditional privacy identification technology fails because the conventional privacy positioning and theory fail to effectively cover the scope of privacy authority attribution under big data (Long, 2014). Failed privacy-identification technologies and the widening data divide have upset the balance between individuals and data resource owners. Therefore, the present study needs solid and effective privacy data identification technology, data protection technology, and relevant laws and regulations to adjust the relationship between the two.

Traditional privacy data protection technologies focus on more minor data and are less effective at dealing with extensive and multi-type big data. On the one hand, the conventional Anonymous method commonly used at present has become easy to crack in big data, the Fuzzing technology is cumbersome and lowly applicable in the face of large amounts of data, and the traceability query method is too limited in function construction, and it isn't easy to adapt to the complex situation of reality (Meng, 2015; Feng, 2019). On the other hand, the new data protection technologies generated under big data also have the limitations of difficulty in the popularization and immature technology. Although the much-concerned differential privacy can effectively prevent data traceability, it is difficult to independently and appropriately determine the input privacy parameters (Meng, 2015). Although Homomorphic encryption technology can encrypt private data, it has the limitation of higher requirements for computing performance and the disadvantage of higher cost (Liu et al., 2022). Although secure multi-party computing can improve the security of encryption and the efficiency of the guarantee protocol, it still has the defect of being unable to effectively defend against attackers who break the protocol and are too costly (Li, 2007). Other new protection technologies, such as Blockchain, have also been technically Bottlenecked by insufficient data throughput capabilities and limited application scope (Liu et al., 2022), resulting in the gradual elimination of old protection technologies and the fact that new technologies have not yet been rolled out. This phenomenon reflects that traditional protection technologies can no longer meet the needs of privacy protection in the era of big data, which is one of the leading causes of data privacy security issues.

3.1.2 Data subject awareness management and data privacy policy management

Internal factors at the management level mainly come from the lack of awareness of the behavior of data subjects. This paper takes the data subject as the reference object. It divides the factors at the management level into the internal consciousness management of the issue and the external policy management. With the continuous enrichment of the content of the big data industry and the continuous clarification of the industrial division of labor, data subjects can gradually be divided into users, data owners, and data producers. In China, the privacy protection awareness of the three types of data subjects has different degrees of defects. First of all, data owners hold a large number of information resources. They should shoulder the responsibility of data disclosure and privacy protection. Still, in the face of economic temptations, some data owners, such as enterprises and governments, have shown a lack of sense of responsibility and insufficient self-control capabilities, which in turn has spawned a variety of privacy cases that are complex and difficult to deal with by existing laws (Zhang, 2019; Zhu & Li, 2021). Second, most users and data producers, as objects of data privacy protection, are ignored by their increasingly open values and weak awareness of data protection, resulting in their personal privacy information being snooped, stolen, and exploited (Di, 2016; Feng, 2019). Finally, the privacy awareness of most data subjects has limitations and lags (Mao, 2019). First, the infringer believes that as long as it does not involve the loss of economic and life interests, it is unnecessary to pay attention to the infringement. In addition, the infringe often becomes aware of data privacy protection after being violated by data privacy. This lag and limitation create a large amount of activity space for

various privacy violations. Together with the lack of data privacy management awareness of data subjects, it constitutes the main internal management factor of data privacy leakage problems, which objectively increases the difficulty of solving data privacy security issues.

External factors at the management level mainly come from the shortcomings in managing data privacy protection policies. Shang (2020) stressed that the fundamental element of personal privacy leakage in the era of big data lies in the lack of a suitable data privacy legal system. In other words, due to the late start of research on data privacy rights, compared with developed countries, countries still have problems such as unclear division of rights, imperfect laws and regulations, and poor ways to protect rights (Wen, 2010). The flaws of this system have led to continuous contradictions and conflicts between traditional privacy protection strategies and big data-guided development trends, which have made the original network normative order questionable (Zhang, 2019). The failure of the normative order is embodied in the declining moral cultivation of the infringer and the increasing demand for privacy protection of the infringed party (Xie, 2019). This also reflects the lack of attention paid by the government to data privacy issues in another way. In most countries' current judicial processing, there are missed examples of judicial authorities making expansive interpretations of privacy rights based on data characteristics (Shang, 2020). In terms of international information management legislation, countries have long focused on ensuring the regular operation of the Internet and passively withstanding hacker attacks, thus ignoring the formulation of systematic laws and regulations, which finally leads to the current situation of weak handling of complex data privacy cases (Yin & Wang, 2016; Ning & Li, 2020; Zhu & Li, 2021). Whether in the implementation of the system or legal attention, the degree of attention to data privacy and security needs to be improved.

3.2 Information security issues are becoming increasingly prominent in the era of big data

The security and harm caused by data privacy have been further amplified in the Internet era. The value of personal data has been more deeply excavated in the period of big data, and it has become a commodity that occupies a vast supply, demand, and market. Organizations or individuals have been able to collect and analyze data on a larger scale, resulting in various "overlord clauses" for collecting personal data information and even making profits by reselling personal data, constantly breeding eclectic illegal collection of personal data (Zhu & Li, 2021). Once personal data is collected and uploaded to various databases in the online world, it cannot escape the fate of automated surveillance and secondary exploitation (Yuan, 2015; Zhu & Li, 2021). On the one hand, this result will lead to the misuse and dissemination of personal data, resulting in the invasion of personal privacy (Xu & Dong, 2014).

On the other hand, due to the fast, large volume, and significant scope of information dissemination, it is more likely that individuals with weak awareness of specific economic, cultural, and political situations and privacy awareness will be illegally collected and forced to leak personal data (Shang, 2020) repeatedly. Once personal privacy data is leaked, citizens' right to know and right to information self-determination will be seriously infringed (Zhu & Li, 2021), resulting in serious interference in the expected life of the victim, which in turn violates personal real-life, amplifies social contradictions, and endangers the public interest (Mao, 2019; Peng, 2021), infrastructure construction that ultimately causes losses to the real economy and threatens the safety of individual lives and national defense security (Wang et al., 2016). The above phenomena show that data privacy security issues affect different individuals to varying degrees, and the harm caused by them has touched the fundamental interests of the majority of people.

In the era of big data, the privacy disclosure of personal data leads to multiple interests, affecting all walks of life. Personal data includes primary personal data, personal privacy data, and non-primary essential information between two, and the consequences of privacy data and information theft are becoming increasingly severe. Information security has become an urgent problem to be solved by all walks of life (Gu, 2021). To obtain the convenience brought by big data, people need to continuously disclose personal data, which leads to data privacy being under monitoring at any time, and the risk of leakage will be significantly improved; however, if people are cautious in their words and deeds to avoid the leakage of privacy, they cannot enjoy the "freedom" brought by the era of big data (Chen & Huang, 2016; Feng, 2019). Moreover, more and more people find that their behavior of disclosing personal data begins to change from active to passive. The fear of losing control of information affects the development of all walks of life, so people urgently need a safe and efficient information protection system to ensure the quality of data release and protect data that may leak privacy (Wang et al., 2020). It can be seen that it is imperative to solve the problem of personal privacy and security in the era of big data, and it is urgent to need a relatively complete data privacy protection management system. To this end, the following privacy management suggestions based on the life-cycle model will be proposed for the causes of data privacy security issues.

4. Data privacy protection recommendations based on the data life-cycle model

Cracking the dilemma of data privacy security requires seeking a balance between privacy protection and data openness. The realization of this balance can take the data life-cycle model as a reference and implement every step of theoretical and technical research on the actual situation. This section will put forward guiding suggestions for data privacy protection in three aspects: technical management, system management, and awareness management based on the management perspective aiming to provide specific references for solving data privacy security issues.

4.1 Technical management level

In terms of technical management, maintaining data privacy security needs to start from three aspects: privacy protection technology, big data algorithms, and industry standards. First, privacy protection technology needs to be carried out around the data life-cycle (Meng, 2015), with professional technology research and development for each data stage. On this basis, the advanced technology protection theory of all parties is integrated, then promotes the implementation of protection technology from theory to practical application, and finally establishes a set of safe, comprehensive, efficient, and professional technology management systems for computer security. Secondly, big data network algorithms include data operations of the whole data life-cycle, which need to pay more attention to the identification and protection of personal privacy at the beginning of the design of requirements (Zhang, 2022), thereby enhancing people's control over high-order algorithms of big data, and finally establishing a series of more humanitarian computer algorithms and algorithm management standards from data collection to data push. Finally, computer data technology and data protection technology also need to follow the data life-cycle model (Zhu & Li, 2021). Different technical means and guidelines should be adopted for different data stages to establish technical standards that can maintain the harmonious development of people and technology. If the data life-cycle model can be introduced into the structural design and functional implementation of privacy protection technology, and the people-oriented standards can be integrated into technology research and development and standard-setting, data privacy security issues may be improved.

4.2 Institutional management level

In terms of system management, improving data privacy security requires paying attention to the management system based on the data life-cycle model. Zhu & Li (2021) emphasized the importance of combining the data life-cycle with the management system and derived a scientific analysis framework from the model that combines data sensitivity and data stage. According to the framework of the above analysis, this paper proposes the following five steps of data privacy protection methods: First, in the data deletion stage, it is necessary to establish a data operating system that the data subject can operate, reduce the uncontrollability of the data removal stage, and thus protect the individual data forgetting right (Wang & Zhao, 2020). Second, in the data release stage, the government must take the lead in formulating fair data transaction guidelines and a complete data management system. At the same time, enterprises must pay attention to personal data security protection, standardize the "overlord clause" of the application, and cooperate in purifying the network environment of data release. Third, in the data application stage, the government needs to set up a particular sensitive information protection agency, build a multi-party reporting system for the government, enterprises, and individuals, and form a triangular relationship of mutual restraint and mutual supervision between the three parties, thereby ensuring the user's data privacy. Fourth, in the data transportation stage, all parties can establish a data transportation confidentiality mechanism through the supervision of an independent third-party agency. The government should introduce relevant policies to crack down on the illegal theft of private data severely. Finally, in the data production and collection stage, all parties should first clarify the purpose of data collection, divide the data's sensitivity, and ultimately refer to different sensitivity levels and combine the protection technology to protect the data hierarchically. The management method based on the data model will be the multi-faceted and multi-level refinement of data privacy protection. To a certain extent, it can improve the overall framework of data protection management.

4.3 Awareness management level

In terms of awareness management, improving awareness of data privacy protection for individuals and enterprises is an essential part of awareness management. The ultimate goal of addressing data privacy security issues is to protect privacy while preserving the positive impact of technology on productivity. Although the sound development of protection technology and management norms can effectively help realize the goal, the article emphasizes the importance of raising citizens' awareness of data privacy protection. First, individuals should cultivate healthy online habits, pay more attention to and reject illegal websites, reduce the public release of sensitive personal information on social software, and develop self-disciplined privacy protection habits (Gu, 2021). Secondly, enterprises and organizations should cultivate a sense of privacy and ethical responsibility (Li, 2022), fully understand the harm of data privacy leakage, and refuse the illegal transaction of personal privacy data, to establish a sense of industry ethics, and ultimately establish an industry benchmark and corporate culture with moral responsibility. Finally, the government and enterprises should implement data privacy education and create a privacy protection atmosphere, and carry out education related to a dialectical view of data privacy and data privacy protection, to reduce the limitations and lag of people's awareness of protection, thereby improving media literacy and personal quality in the era of big data, and ultimately establishing a correct view of privacy and values. Any awareness cultivation requires the joint efforts of individuals to organizations from all walks of life, and the awareness cultivation related to data protection also requires mutual help from all walks of life to break the boundaries of multi-party relationships under big data and ultimately create opportunities for exploring a balance to solve privacy problems.

5. Conclusion

This paper mainly starts from the current research situation of data privacy protection, researches the data-sharing stage in the data life-cycle, and introduces the external and internal factors of data privacy security issues. Starting from the internal and

external factors introduced, the present study summarizes the shortcomings of existing data protection technology and the omissions of current data privacy management, sums up the existing and potential harms of privacy security issues, and then puts forward references from the perspective of the data life-cycle. There is a widespread problem in data privacy protection that the protection mechanism can not keep up with the development of technology, which still needs to be sorted out and discussed in depth to prospect the future research direction.

At present, the research on data privacy and security can expand the scope of the study. First, future research needs to focus on other data life-cycle stages. Privacy objects at all data life-cycle stages deserve to be analyzed and studied. The direction of follow-up research can be extended to the data production and destruction stages of the life-cycle. Second, future research should be more integrated with big data application scenarios. Big data has spawned many new formats, and different industries and technologies have additional requirements for data privacy protection. The research on data privacy protection needs to spread to specific issues of other objects, and many contents are worthy of further study.

Acknowledgments: We would like to thank Jialiang Chen for his insightful suggestions and constructive feedback on content relevance, content sufficiency, organization, and language quality in an earlier version of this paper.
Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Allen, A. L., & Torkington, R. C. (2004). *Privacy law in the United States: Doctrinal jurisprudence and legislation.* Beijing: China Democracy and Legal Publishing House.
[2] Chen, S., & Huang, X. (2016). Ethical governance of privacy protection in the era of big data. *Academia,* (01), 85-95.
[3] Di, L. (2016). Impact assessment of privacy disclosure in the big data environment. *Journal of Intelligence, 35*(4), 6. https://doi.org/10.3969/j.issn.1002-1965.2016.04.026
[4] Du, X., Lu, W., & Zhang, F. (2019). History, current situation, and future of big data management system. *Journal of Software,* (01), 127-141. https://doi.org/10.13328/j.cnki.jos.005644
[5] Gu, L., & Fan, H. (2018). The academic field of ten years' research on network privacy: An analysis of visual scientific knowledge map based on citespace (2008-2017). *Journalism and Communication Studies, 25*(12), 18. https://doi.org/CNKI:SUN:YANJ.0.2018-12-005
[6] Gu, Z. (2021). Research on personal information security in big data environment. *Information Science, 39*(12), 5. https://doi.org/10.13833/j.issn.1007-7634.2021.12.014
[7] Huang, L., Tian, M., & Huang, H. (2015). A review of cryptography for big data privacy protection. *Journal of Software, 26*(4), 15. https://doi.org/10.13328/j.cnki.jos.004794
[8] Jiang, P. (2019). Review of research on personal information protection in the era of big data. *Library and Information Service, 63*(15), 9. https://doi.org/10.13266/j.issn.0252-3116.2019.15.016
[9] Jin, Y. (2020). On disclosure and protection of personal privacy data in the era of big data. *Journal of Tongji University: Social Science, 31*(3), 18-29.
[10] Li, A. (2022). On the legitimacy of enterprise data property rights -- from the perspective of locke's property rights theory. *Science and Technology and Law (Chinese and English),* (01), 91-100. https://doi.org/10.19525/10.19685/j.cnki.cn11-2922/n.2022.01.011
[11] Li, N. (2007). T-closeness: privacy beyond k-anonymity and l-diversity. *Proc.ieee Int.conf.on Data Engineering, 22*(7), 106-115. https://doi.org/10.1109/TKDE.2009.139
[12] Li, X., Sun, Z., Deng, J., & Song, G. (2013). A review of privacy protection technology. *Journal of Computer Science, 40*(11A), 4. https://doi.org/10.1109/TKDE.2009.139
[13] Long, X. (2014). Definition and protection of digital privacy in the era of "All-seeing Eye". *Journalist,* (8), 7. https://doi.org/CNKI:SUN:XWJZ.0.2014-08-019
[14] Loukis, E. (2016) A taxonomy of open government data research areas and topics. *Journal of Organizational Computing & Electronic Commerce, 26*(1), 41-63. https://doi.org/10.1080/10919392.2015.1124720
[15] Lu, H. (2018). The changing trend of sci-tech industrial revolution in the next 30 years and suggestions on China's innovation development. *Beijing: Globalization,* (3), 89-97,135. https://doi.org/CNKI:SUN:QUQH.0.2018-03-010
[16] Mao, D. (2019). *Research on Big Data Privacy Protection Technology and Governance Mechanism.* Beijing: Tsinghua University Press.
[17] Meng, X., & Zhang, X. (2015). Big data privacy management. *Journal of Computer Research and Development, 52*(2), 17. https://doi.org/10.7544/issn.1000-1239.2015.20140073
[18] Peng, N. (2021). Review of domestic data privacy protection research. *Changsha, Hunan: Library,* (11), 7.
[19] Qian, R. (2013). *Research on Mmproving Anonymous Algorithm Based on Data Privacy Protection.* (Dissertation, Hangzhou Dianzi University). https://doi.org/10.7666/d.D318807
[20] Qiu, C., & Li, C. (2018) Application of data encryption technology in computer networks. *China Management Informatization, 21*(4), 131-132.
[21] Ren, Z., Li, D., & Tan, L. (2022). Research on global science and technology development trend in 2035 and China's coping strategy. *Science and Technology Management Research,* (10), 34-40.
[22] Samarati, P., & Sweeney, L. (1998). Generalizing data to provide anonymity when disclosing information (abstract). *Seventeenth Acm Sigact-sigmod-sigart Symposium on Principles of Database Systems,* (p.188). ACM. https://doi.org/10.1145/275487.275508
[23] Shang, X. (2020). The realization path of personal information privacy interest and self-determination interest. *Law Science (Journal of Northwest University of Political Science and Law),* (03), 71-85. https://doi.org/10.16290/j.cnki.1674-5205.2020.03.009

[24] Tian, X., & Huang, Z. (2014). The paradox of "public data openness" and "personal privacy protection". *University of Journalism,* (6), 7. https://doi.org/CNKI:SUN:XWDX.0.2014-06-009
[25] Wang, J., Liu, C., & Fang, B. (2016). A review of data privacy protection for internet of things search. *Journal of Communications, 37*(9), 142-153. https://doi.org/10.11959/j.issn.1000-436x.2016186
[26] Wang, L. (1994). *New Theory of Personality Right Law.* Jilin People's Publishing House.
[27] Wang, M., Li, W., Zhang, Q., & Li, X. (2020). A review of data publishing oriented privacy protection technology. *Microcomputer Systems, 41*(12), 11.
[28] Liu, Y., Chen, H., Liu, Y., & Li, C. (2022). Privacy Protection Technology in Federated Learning. *Journal of Software,* (03), 1057-1092. https://doi.org/10.13328/j.cnki.jos.006446.
[29] Wang, X., & Zhao, X. (2015). Exploration on the protection of privacy by integrating public and private laws -- from the perspective of personal information privacy in "Big Data Era". *Hebei Law School, 33*(5), 63-71. https://doi.org/10.16494/j.cnki.1002-3933.2015.05.008
[30] Wang, Y., & Zhao, J. (2020). Realistic logic and local construction of the right to be forgotten in the Era of big data. Journal of Nanchang University (Humanities and Social Sciences), 51(06), 103-111. https://doi.org/10.13764/j.cnki.ncds.2020.06.012
[31] Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review, 4*(5), 193-220. https://doi.org/10.2307/1321160
[32] Wen, W. (2010). Legal discussion on the protection of internet privacy in China. *People's forum: Mid-day Journal,* (9), 2. https://doi.org/CNKI:SUN:RMLT.0.2010-26-035
[33] Xie, L. (2019). Exemption of legitimate interests in the use of personal information in the era of big data. *Forum of Politics and Law,* (1), 11. https://doi.org/10.3969/j.issn.1000-0208.2019.01.007
[34] Xu, L., & Dong, L. (2014). Discussion on personal data privacy protection of cross-border business in China in the era of big data. *Contemporary Economic Management, 36*(11), 4. https://doi.org/10.13253/j.cnki.ddjjgl.2014.11.008
[35] Xu, Y., & Hu, X. (2019). Research on data privacy protection mechanism in SaaS environment. *Modern Electronics Technology,* (17), 68-74. doi:10.16652/j.issn.1004-373x.2019.17.015.
[36] Yan, K. (2020). The Sharing of big data-Privacy Paradox. *Journal of Dalian University of Technology: Social Sciences, 41*(5), 6. https://doi.org/10.19525/j.issn1008-407x.2020.05.003
[37] Yin, J., & Wang, Z. (2016). Research on personal data traceability management systems in big data environments. *Information Science, 34(2*), 5. https://doi.org/CNKI:SUN:QBKX.0.2016-02-028
[38] Yuan, M. (2015). "The right to be forgotten" debate: Digital memory and privacy boundaries in the Era of big data. *Xuehai, 000*(004), 55-61. https://doi.org/10.3969/j.issn.1001-9790.2015.04.009
[39] Zhang, F. (2019). Ethical dilemma and countermeasures of privacy protection in the era of big data. *People's BBS Academic Frontier,* (15), 76-87. https://doi.org/10.16619/j.cnki.rmltxsqy.2019.15.008
[40] Zhang, H. (2022). Re-differentiation of information privacy and personal information function positioning. *Beijing Social Science,* (01), 98-108. https://doi.org/10.13262/j.bjsshkxy.bjshkx.220110
[41] Zhang, X. (2015). From privacy to personal information: Theory and institutional arrangement of interest remeasurement. *Internet Finance Law Review,* (2), 6. https://doi.org/CNKI:SUN:IFLR.0.2015-02-005
[42] Zhang, Y., & Zhu, Q. (2014). Review of foreign information privacy Research. *Library and Information Service, 58*(13), 9. https://doi.org/10.13266/j.issn.0252-3116.2014.13.023
[43] Zhou, S., & Xu, K. (2015). User privacy protection in information services under the background of big data. *Modern Information, 35*(11), 6. https://doi.org/10.3969/j.issn.1008-0821.2015.11.007
[44] Zhu, L., & Zeng, R. (2021). Privacy Information protection in the era of big data opening: Core issues and frontier hot spots. *Journal of Intelligence,* (9), 115-123.
[45] Zhu, Y., & Li, X. (2021). An analysis framework for personal data privacy security protection in the era of big data. *Journal of Information Science,* (01), 165-170. https://doi.org/10.3969/j.issn.1002-1965.2021.01.024