

Reflections on the Legal Boundaries of Scientific Research Processing of Personal Data from a Comparative Law Perspective

Zifan Huang*

School of Law, Queen Mary University of London, London, United Kingdom

**Corresponding author: Huangzifan1029@outlook.com*

Keywords: Personal data; public interest; scientific research

Abstract: This article aims to explore the legal boundaries of handling personal data in scientific research from a comparative law perspective. Currently, there are some deficiencies in the legal framework in China regarding the handling of personal data in scientific research. Therefore, it is worth considering the addition of provisions related to "scientific research" in the Personal Information Protection Law of the People's Republic of China. Additionally, it is possible to mitigate the personal rights of data subjects by appropriately exempting researchers from data processing obligations and imposing additional conditions, as well as adjusting the consent mechanism, in order to establish incentive-compatible requirements for data processing behavior. Through these improvement measures, it is hoped that clearer and more flexible legal boundaries can be provided for scientific research while respecting the rights of personal data subjects.

1. Introduction

In the ever-changing global landscape, the rational utilization of data and harnessing its value has become a central issue for national development. Scientific research, as a crucial domain for data utilization, has witnessed the prominence of large-scale data processing during the Fourth Industrial Revolution. However, the extensive handling of data in scientific research raises concerns regarding the infringement of personal data subject rights, which cannot be ignored. To earn the trust of data subjects and ensure the sustainability of scientific research while harnessing the benefits of technology, it is imperative to effectively safeguard the legal rights of data subjects. Therefore, when conflicts arise between the public interest of scientific research and the rights of personal data, the law must strike a balance to achieve a win-win outcome. Answering the question of what legal boundaries scientific research should adhere to in handling personal data holds urgent practical significance and requires thorough research and careful consideration.

2. Conflicting Interests Arising from the Handling of Personal Data in Scientific Research

Scientific research is a public endeavor dedicated to understanding, exploring, and shaping the world^[1]. In the era of big data, the public benefits of scientific research have been further amplified. By aggregating, analyzing, and utilizing massive amounts of data, scientific research enables accurate

descriptions and precise predictions, facilitating rapid and efficient decision-making. For instance, the integration of digital technology and medical research has given rise to the emerging field of "digital health," which effectively monitors individuals' health, predicts health issues in advance, assists patients in dealing with emergencies, provides personalized treatment plans, reduces healthcare costs, improves efficiency, and promotes medical discoveries and drug development^[2]. Among these advancements, the development of digital pills can activate microcircuits upon contact with gastric fluid, informing patients through external sensors if and when they have taken their medication. The fusion of big data and public administration also facilitates novel governance methods such as social credit ratings and predictive law enforcement, shifting the focus of governance from reactive measures to proactive prevention, thereby reducing the cost and enhancing the effectiveness of governance. The significant achievements in epidemic management in China can be attributed, to some extent, to the application of digital technology in emergency management.

However, the extensive handling of personal data in scientific research can potentially infringe upon individuals' data rights. Even seemingly ordinary data, such as the number of steps taken in a day, can be used to infer and predict personal behavior, even revealing whether an employee is genuinely unable to work due to illness^[3]. Research has shown that the personal rights involved in personal information protection form a concentric pattern, with human dignity at the core, followed by general personality rights, privacy rights within individual personality rights, information privacy rights within privacy rights, and finally, the right to information self-determination^[4]. In scientific research, apart from individual rights, the handling of personal data also involves the research freedom of researchers and research institutions, as well as the public interest. Existing studies have pointed out that research freedom faces two limitations: the public interest and the rights and freedoms of others. Therefore, scientific research encounters conflicts between multiple interests when dealing with personal data, with the fundamental conflict lying in the balance between individual rights and the public interest.

Utilizing personal data for scientific research carries both value and risks. The core issue at present is not whether data should be used for research but rather how to obtain the benefits of scientific research while respecting ethical and privacy principles. Scientific research is not exempt from the law, and when it has the potential to harm research subjects or other relevant individuals, the law needs to intervene promptly and establish boundaries. It is essential to comply with relevant legal provisions to ensure reasonable restrictions and protections while handling personal data in scientific research.

3. Analysis of Legal Regulations on the Handling of Personal Data in Scientific Research in China

3.1 Analysis of the Personal Information Protection Law of the People's Republic of China

On August 20, 2021, the 30th session of the Standing Committee of the 13th National People's Congress officially passed the Personal Information Protection Law of the People's Republic of China (hereinafter referred to as the "Personal Information Protection Law"), which laid the legal foundation for protecting personal information. However, the law does not directly specify provisions regarding the purpose of "scientific research." Before the release of the draft of the Personal Information Protection Law, some scholars called for exceptions such as "personal information comparison necessary for criminal investigation by investigative agencies, personal information comparison for supporting academic research or statistical projects, obtaining statistical data, etc." ^[5], to be included as exceptions for information comparison conducted by national public authorities, attempting to define exceptions to the "purpose limitation" in another way. However, the final draft submitted for deliberation did not include exceptions that would allow the use of personal information beyond the

collection purpose. According to the existing regulatory content, collecting personal information for purposes other than the collection purpose is not permitted. Whether it is a public authority or a private entity, if they want to exceed the purpose limitation, they must obtain the consent of the information subject again. Therefore, activities that involve the handling of personal information for purposes such as academic research, artistic expression, and literary creation face the issue of exempting the principle of purpose limitation and require specific guidelines from relevant government departments for regulation.

From the potential space for interpretive definitions, the articles that may be relevant to scientific research include Articles 5 to 9, which establish general principles for the processing of personal information, emphasizing that the processing must be "lawful and proper," and they set out the principles of good faith, clear and reasonable purpose, limited use, openness and transparency, accuracy of information, and information security. Article 13, item 6 stipulates, "Processing of personal information for implementing news reporting and public opinion supervision activities for public interests shall be done within a reasonable scope." Article 28, paragraph 2 states, "Sensitive personal information may only be processed under specific purposes and with sufficient necessity and strict protection measures." Additionally, Article 32 further stipulates, "If laws and administrative regulations provide that the processing of sensitive personal information shall obtain relevant administrative permits or impose other restrictions, such provisions shall be followed." From a textual perspective, if scientific research follows the general principles for personal information processing in Articles 5 to 9, it will be treated no differently from other data processing activities, and scientific research will be subject to strict limitations. If scientific research is included in the "public interest" category in Article 13, it can process information within a "reasonable scope." If scientific research falls under the "specific purpose" in Article 28, it must also have "sufficient necessity" to process sensitive personal information.

Based on the above analysis, the provisions in the Personal Information Protection Law regarding scientific research are unclear and may have two consequences: firstly, scientific research is placed on equal footing with general data use purposes, and the relatively relaxed development conditions required for scientific research are not guaranteed; secondly, even if scientific research is interpreted in terms of text, and included in the "public interest" or "specific purpose" categories, the legal conditions for the handling of personal information in scientific research remain ambiguous. Therefore, it is necessary for relevant government departments to issue specific guidelines for regulation.

3.2 Analysis of Other Relevant Legal Regulations

Article 20 and Article 47 of the Constitution of the People's Republic of China mention scientific research, demonstrating the state's positive attitude towards scientific research. In particular, Article 47 stipulates the freedom of citizens to conduct scientific research and explicitly encourages and supports creative work that benefits the people, which is widely regarded by the academic community as the constitutional basis for academic freedom. At the same time, Article 51 of the Constitution is recognized as a provision that restricts the fundamental rights of citizens. According to the provisions of Article 51, the external limits of academic freedom are public interests, while the internal limits are the rights and freedoms of others. Article 1009 of the Civil Code and Article 29 of the Law of the People's Republic of China on Progress in Science and Technology both explicitly state that the bottom line of basic values such as public interests and ethical morals must not be violated. In addition, Article 1020, item 1 of the Civil Code stipulates that publicly available portraits can be used in classroom teaching or scientific research without the consent of the rights holder within a necessary scope.

3.3 Deficiencies in Existing Legal Regulations

Observing the current legal regulations in China, the regulations regarding scientific research activities are scattered and not systematic. Although there are regulations concerning the handling of personal data in scientific research, there is a lack of clear and balanced legal mechanisms that consider the purposes of scientific research and personal data protection. Firstly, the Personal Information Protection Law, as the fundamental law in the field of personal data protection, fails to provide clear and appropriate guidance for the handling of personal data in scientific research, both in terms of its content and regulatory interpretation. Other regulations also do not give special treatment to the purposes of scientific research, and scientific research cannot enjoy a privileged position different from general data utilization scenarios. The strict provisions are not conducive to the development of scientific research. Secondly, there are inconsistencies in the legal provisions for similar research activities in the existing regulations. Some regulations require the signing of informed consent, the obligation to provide detailed explanations, and the reacquisition of informed consent when there is a change in the purpose of use, while other regulations state that the use of de-identified data for research does not require the authorization or consent of the data subject. This inconsistency needs to be systematically sorted out and regulated to ensure effective guidance through unified guidelines. Furthermore, the frequent use of terms such as "public interest," "reasonable scope," "sufficient necessity," and the use of uncertain legal concepts in the regulations make the requirements for handling data in scientific research ambiguous and confusing for researchers and data subjects. It is necessary to classify the specific types of personal data handling in scientific research and define the connotation and scope of uncertain concepts by setting clear requirements. Finally, there is a lack of specialized data processing regulatory agencies and sound supervision mechanisms in the regulations. Although some regulations establish "ethics committees" and other institutions, the functions of these institutions tend to focus on safeguarding scientific research activities or personal data protection, with few organizations that can oversee and balance both aspects. In addition, the independence of regulatory agencies is insufficient, and the effectiveness of supervision is difficult to guarantee.

4. Establishing Regulatory Framework for the Legal Regulation of Personal Data in Scientific Research in China

4.1 Positioning and System Construction of Basic Laws and Special Laws

Incorporating the public interest purposes of scientific research and other activities involving the handling of personal data into the formal provisions of the Personal Information Protection Law can clarify the legal status of research activities. This measure holds two important implications: firstly, it demonstrates the importance attached to scientific research activities and declares the value of scientific research as a special data application scenario; secondly, it establishes exemption conditions for the handling of personal data in scientific research, providing benchmark guidelines for various specific research activities and achieving this purpose through the establishment of rules for limitations and exceptions. After the Personal Information Protection Law provides the "basic legal positioning," specialized laws for various research activities should provide more detailed provisions based on their specific circumstances. These provisions can be moderately adjusted within the scope defined by the Personal Information Protection Law but should not exceed the legal boundaries set by the law. The specialized laws need to be interconnected and coordinated with each other, promptly cleaning up and integrating any duplicate or conflicting provisions to maintain the systematic effect of legal norms for scientific research. Currently, China has specific regulations for biomedical research, such as the "Ethical Review Measures for Biomedical Research Involving Humans" and the

"Information Security Technology - Guidelines for Health and Medical Data Security." However, further strengthening the legal regulation of data handling in other types of research is necessary. The Personal Information Protection Law provides a foundation, which is why it directly regulates scientific research in personal data protection legislation in many other countries. China can draw on the experience of this legal framework to provide value guidelines for personal data protection in scientific research.

4.2 Reshaping Requirements for Data Processing Behavior under an Incentive-Compatible Approach

4.2.1 Exemption of Obligations for Data Processors

Considering the expansion of public interest, researchers in scientific research enjoy special treatment when handling personal data. Firstly, the types of personal data that researchers can process should not be restricted, as the research outcomes for each type of data may bring uncertain benefits to the public interest. Secondly, the exemption principle of purpose limitation applies to scientific research. The goal of scientific research is to explore patterns within uncertainty, often transitioning from complete uncertainty to relative uncertainty. Therefore, exempting scientific research from purpose limitation helps unleash the potential of research. Finally, the exemption of obligations for researchers in scientific research should be kept to a minimum, and researchers should adhere to basic research obligations.

The European Union has adopted a two-stage structure of "narrow in, wide out" in general data application scenarios to implement the purpose limitation principle. During the data collection stage, the agreed purposes should follow the principles of "lawfulness, specificity, and clarity." During the data utilization stage, both the agreed purposes and additional purposes that do not violate the agreed purposes are permitted within the authorized scope, achieving a balance between information protection and data flow. Although the allowance for additional purposes exists, it is still subject to standards such as "reasonable expectations" or "balancing tests." Therefore, when researchers overlook common sense or violate regulations for appropriate data usage, they will be subject to scrutiny and penalties from regulatory authorities.

4.2.2 Updating the Conditions for Data Subject Consent and Proportionate Control over Rights Reduction

The application conditions for consent outside the domain are determined based on data identifiability, without delving deep into the forms of consent themselves. With advancements in data technology, the traditional all-or-nothing rigid consent form has faced criticism. Broad consent models have been proposed, where data subjects provide one-time consent for the subsequent use of their personal data in research, although the approval of research review bodies is still required for such subsequent research^[6]. In fact, international health research governance bodies such as the World Medical Association, the Council for International Organizations of Medical Sciences, and the World Health Organization have approved broad consent as an acceptable alternative^[7]. Broad consent models have been widely adopted in Germany and the Netherlands as a means to legitimize secondary research with medical data, and reasons for participants choosing broad consent include altruism, reciprocity, solidarity, and gratitude^[8]. In the face of the risk that broad consent may undermine the substantive function of consent, dynamic consent models have been proposed, treating consent as an ongoing process of negotiation and adjustment in response to changing circumstances. Participants can freely choose to consent to join or withdraw from the research and select the degree to which they prefer to receive information processing^[9]. This places high demands on the level of participant

involvement and transparency in the research process. Consent is a concept that evolves, and as the types of personal data processed in scientific research become progressively defined, the accurate application of categorized consent and layered consent models can highlight the effectiveness of consent.

5. Conclusion

In the era of big data, the rapid development of the digital economy is particularly prominent. However, the application scenarios of scientific research involving the processing of personal data have not received the attention they deserve. Faced with the tension between personal data protection and data utilization, there is a need to shift from the current ambiguous regulatory stance under the uncertain legal conceptual framework to a meticulous analysis and construction of the legitimacy requirements for data processing in scientific research. It is crucial to establish clear legal boundaries for the handling of personal data in scientific research. Only by doing so can we provide guidance to ensure the lawful, transparent, and ethical use of data in the face of the ongoing advancement of digital technologies. This will also promote a balance between scientific research and individual privacy rights.

References

- [1] Chai, Z. (2017). *Scientific research is a public welfare undertaking*. *Science and Technology Daily*, 35(01), 9.
- [2] Vayena, E., Haeusermann, T., Adjei-Kwayisi, A., et al. (2018). *Digital health: Meeting the ethical and policy challenges*. *Swiss Medical Weekly*, 148, w14571.
- [3] Bietz, M., Patrick, K., & Bloss, C. (2019). *Data donation as a model for citizen science health research*. *Citizen Science: Theory and Practice*, 4(1), 1-11.
- [4] Zhu, S., & Wang, J. (2020). *Criminal protection and implications of special personal data in Taiwan*. *Taiwan Research Journal*, (04), 84-91.
- [5] Zhou, H. (2006). *Personal Information Protection Law (Expert Draft) and Legislative Research Report*. Beijing: Law Press.
- [6] Rothstein, M. A., Knoppers, B. M., & Harrell, H. L. (2016). *Comparative approaches to biobanks and privacy*. *Journal of Law, Medicine & Ethics*, 44(1), 161-172.
- [7] Council for International Organizations of Medical Sciences. (2021). *International ethical guidelines for epidemiological studies*. Retrieved from <https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-Ethical-Guidelines.pdf>
- [8] Richter, G., Borzikowsky, C., Lesch, W., et al. (2021). *Secondary research use of personal medical data: Attitudes from patient and population surveys in the Netherlands and Germany*. *European Journal of Human Genetics*, 29(3), 495-502.
- [9] Kaye, J., Whitley, E. A., Lund, D., et al. (2015). *Dynamic consent: A patient interface for twenty-first-century research networks*. *European Journal of Human Genetics*, 23(2), 141-145.