U.S. Government Open Data Privacy Protection Policies and Their Implications for China: A Content Analysis of 52 Policy Documents (Postprint)

Authors: Chu Jiewang, Ding Hui

Date: 2023-04-01T16:02:50+00:00

Abstract

[目的/意义] Personal privacy protection constitutes a critical component in the government open data process. By systematically reviewing personal privacy protection policies in U.S. government open data practices, this study provides valuable references for personal privacy protection in China's open government data initiatives.

[方法/过程] Through the collection of relevant policy texts on U.S. government open data, this study utilizes NVivo text analysis software to conduct content coding analysis, thereby mapping out the privacy protection framework for U.S. government open data and deriving beneficial insights and references.

[结果/结论] The study constructs a personal privacy protection system for China's government open data from five dimensions: leveraging the functional roles of multiple stakeholders, constructing a privacy protection assessment system, governing the utilization of privacy data, building privacy-preserving computation models, and establishing legal safeguards, thereby maximizing the contemporary value of government data openness.

Full Text

Personal Privacy Protection Policies in U.S. Government Open Data and Their Implications for China: A Content Analysis Based on 52 Policy Texts

Chu Jiewang, Ding Hui
School of Management, Anhui University, Hefei 230601

Abstract:
[Purpose/Significance] Personal privacy protection is a critical component in the process of government open data. By examining privacy protection policies in U.S. government open data practices, this study provides reference for personal privacy protection in China's government data opening initiatives. [Method/Process] Through collecting relevant U.S. government open data policy texts and employing NVivo text analysis tools, this research conducts content coding analysis to systematically 梳理 the U.S. government open data privacy protection framework and derive beneficial insights. [Result/Conclusion] To maximize the era value of government data opening, China should construct a personal privacy protection system across five dimensions: leveraging multi-stakeholder functions, building a privacy protection evaluation system, governing privacy data utilization, constructing privacy computation models, and establishing legal safeguards.

Keywords: Government open data; Policy texts; Personal privacy protection; U.S. government open platform
Classification: G203
DOI: 10.13266/j.issn.0252-3116.2021.08.015

With the convergence of internet technology and economic-social development, the open data movement has been widely implemented. In 2009, the Obama administration issued the Open Government Directive, making government information data resources open and accessible to the public and thereby launching the global government data opening movement [1]. China's Fifth Plenary Session of the 18th Central Committee proposed "implementing the national big data strategy," marking the formal elevation of big data strategy to a national-level priority. China has the world's largest internet user base and abundant data resources, giving government open data strong market advantages and development potential. However, inevitable personal privacy leakage issues arise during data opening. Therefore, constructing a personal privacy protection system for government open data holds significant practical and contemporary value for enhancing national big data governance capabilities, promoting economic transformation, building service-oriented governments, and fostering new industrial development models.

Research on personal privacy protection in government open data has yielded substantial results, primarily from legal, technical, and institutional perspectives. First, regarding the legal basis for privacy rights protection, T. Jaatinen points out that protecting personal privacy data is a fundamental right stipulated in the Charter of Fundamental Rights of the European Union, requiring fair and lawful processing of personal privacy data [2]. Columbia professor A.F. Westin defines privacy as informational self-determination—the right of individuals, groups, or institutions to determine when, how, and to what extent information about them is communicated to others [3]. Second, institutional breakthroughs are sought. D.L. Baumer argues that lagging and incomplete personal privacy protection systems manifest in the lack of policy regulations

for cross-border data flows [4]. D.O. Stephens proposes principles for protecting personal privacy information in government data opening [5]. Third, technological innovation strengthens privacy protection applications. C.A. Ardagna suggests combining XACML with PRIME to achieve a privacy-aware access control solution based on certificate management and privacy support functions [6]. Hu Qiping proposes using xBook, a privacy control platform for third-party applications, to enhance privacy protection [7].

Existing research on government open data privacy protection provides important insights, but most studies focus on researchers' analytical induction and theoretical deduction, with relatively few employing content analysis methods for quantitative analysis of relevant privacy policy texts. Therefore, this study uses content analysis to analyze 52 U.S. data security-related policy texts from 1966-2020, comprehensively examining the U.S. government open data personal privacy protection system and proposing countermeasures for China's government open data personal privacy protection challenges.

2 Research Design

2.1 Data Sources

Using "Privacy protection" as the search term on U.S. government websites including "whitehouse.gov," "nascio," "digital.gov," "congress.gov," "nitrd.gov," and "strategy.data.gov," we searched policy text categories with keywords including "Information opening," "Privacy data," "Privacy protection," "Sensitive data," "Intimate information," and "High intrusion data." After sorting by publication date and removing duplicates and low-relevance texts, we obtained 52 privacy data protection documents since 1966 (see Table 1). These U.S. privacy protection policy documents include acts, implementation guidelines, memoranda, presidential orders, and other types, covering privacy rights, data management, privacy technology, data opening, and other content. To enhance research focus, we extracted relevant content on privacy protection in government open data from these policy documents for coding analysis.

2.2 Research Methods

Based on research needs, we employed content analysis to analyze policy texts through three steps: (1) Collected 52 policy documents were analyzed individually and preliminarily categorized using keywords such as "information disclosure," "privacy data," and "privacy protection"; (2) NVivo 12 Plus software was used for word frequency analysis, core keyword coding, and visualization chart generation; (3) Based on NVivo statistical analysis, we systematically 梳理 the U.S. government open data privacy protection framework to provide reference for China's government open data personal privacy protection.

3 Content Analysis of Privacy Protection Policy Texts in U.S. Government Open Data

3.1 Overview of U.S. Privacy Protection Policy Texts

Analysis of the 52 policy texts reveals that U.S. personal privacy protection practices originated early, marked by the 1960s Freedom of Information Act, which opened the prelude to citizens' information access rights and information privacy protection. Subsequently, U.S. personal privacy protection policy documents were continuously introduced, with expanding content scope, increasingly standardized privacy protection procedures, and a maturing privacy protection framework. We 梳理 the basic overview of U.S. government open data privacy protection from policy issuance time and themes.

3.1.1 Policy Issuance Timeline The policy text compilation shows that the U.S. gradually began formulating privacy policy documents before 2000. Since the Obama administration's promotion of government open data practices in the 21st century, U.S. privacy policy documents have increased sharply, with documents issued after 2010 accounting for 69% of the total. This indicates that: (1) With deepening government open data practices, privacy data protection has become a critical challenge; (2) The U.S. government open data policy system framework is maturing, providing important reference for China's privacy policy formulation.

3.1.2 Policy Themes Through thematic extraction of the policy texts, existing U.S. government open data personal privacy policy themes show diversified distribution across eight categories: privacy rights, children's privacy protection, electronic information privacy, government information security, data utilization, privacy protection directives, action plans, and privacy acts (see Figure 1 [Figure 1: see original paper]). Privacy acts and privacy protection directive policies constitute the majority of samples, indicating that the existing U.S. privacy protection framework employs both macro-level legislation to grant citizens privacy protection legal effect and privacy protection directives to detail departmental privacy protection approaches and procedures, ensuring actionable privacy protection and enhancing feasibility.

High-frequency keyword cloud statistics (see Figure 2 [Figure 2: see original paper]) show that U.S. government open data privacy protection policy texts focus on four thrusts: (1) Advancing privacy protection legislation; (2) Strengthening utilization governance across the privacy information data lifecycle (collection, utilization, control, evaluation); (3) Implementing privacy protection 主体责任 (OMB, third-party websites, NASCIO, etc.); and (4) Improving privacy protection procedures. These four dimensions demonstrate that the U.S. government data opening privacy protection framework possesses both macro-level guiding significance and micro-level operational feasibility.

3.2 Content Coding Analysis of U.S. Government Data Opening Privacy Security Policies

In content analysis, "analysis units must be determined, i.e., various factors to be examined in the research, which should have a necessary connection with the analysis purpose and be convenient for extraction and operation" [54]. We treat relevant content on data privacy protection in the 52 sample policy documents as basic analysis units. For policy node type classification, since the selected documents are U.S. government policies, we adopt common public management domain classification methods, dividing U.S. privacy protection policies into procedural policies and substantive policies. To ensure reference node reliability and consistency, two researchers independently coded policy texts and calculated consistency rates, which met validity and reliability test standards. Accordingly, we established nine analysis node coding rules from procedural and substantive policy supply perspectives (see Figure 3 [Figure 3: see original paper]). Using NVivo 12 Plus keyword search functions, we extracted 79 reference points at the procedural policy level and 114 reference points at the substantive policy level from the 52 policy texts.

3.2.1 Coding Analysis of Privacy Protection Procedural Policies Procedural policies refer to policies that, while not producing substantive consequences, regulate and constrain government activities by clearly defining government activity procedures. Based on existing literature and obtained text materials, we categorize privacy protection procedures, evaluation review, privacy agencies, and privacy frameworks as research dimensions for procedural policies. Table 2 presents the unit coding for U.S. government open data privacy protection procedural policies.

Table 2: Procedural Policy Level Node Coding

Dimension	Key Phrases
Protection Procedures	Implementation procedures for privacy protection; methods for determining management and supervision of DHS records, management policies and procedures; processing PII through procedures and information systems; establishing new procedures and providing updated guidance and agency requirements for using network measurement and customization technologies

Dimension	Key Phrases
Evaluation Review	Establishing privacy data evaluation review mechanisms; NASCIO developed version 1.0 of the Federal Privacy Law Compendium to help states identify and evaluate federal laws that may impact their information systems and policies; monitoring changes to third-party privacy policies and regularly re-evaluating risks; all departments must establish corresponding data review mechanisms
Privacy Agencies	PRA typically requires federal agencies to publish Federal Register notices to solicit public comment on proposed collections; OPCL drafts Privacy Act nondisclosure provisions; OMB Director must formulate detailed rules for federal departmental data publication work
Privacy Framework	Basic framework for privacy data protection policies; creating favorable policy frameworks for open government; specifying basic frameworks and procedures for federal government information resource disclosure and access

Table 2 coding analysis reveals that U.S. policy guidance emphasizes both legal policies, privacy protection procedures, and basic frameworks to provide fundamental guidance for privacy protection, and leveraging multi-stakeholder agency roles to implement 主体责任. Through content and coding analysis of relevant texts, the U.S. procedural privacy protection framework can be summarized as:

(1) Legal System: Consolidating the Foundation of Privacy Protection. Beginning with the 1967 U.S. Freedom of Information Act, U.S. privacy rights legislation has continuously improved. With the 2009 establishment of the first U.S. government open data platform www.data.gov, privacy protection acts for government open data were successively introduced. Guided by the Obama administration's Freedom of Information Act Memorandum [55] and following acts such as the Privacy Policy [56], Open Data Policy, and Open Government Data Act, two main aspects emerge: (1) Data collection review: Federal agencies must conduct routine reviews of collected data, considering privacy leakage, security risks, legal liabilities, and intellectual property restrictions to determine data disclosure; (2) Enhanced data transparency: The personal data collected, its necessity, usage purposes, estimated deletion dates, whether shared with

third parties and sharing purposes must be explained to citizens to enhance privacy protection awareness. Perfecting privacy acts is a crucial guarantee for the successful outcomes of current U.S. government open data practices.

(2) Management System: Strengthening Privacy Protection Responsibilities. Organizational establishment is essential for fulfilling privacy protection responsibilities. As shown in Table 3 , the U.S. privacy protection management system sets up multiple stakeholders with different functions, responsibilities, and data lifecycle management stages to strengthen privacy protection responsibilities and maximize government open data benefits.

Table 3: U.S. Privacy Protection Agencies and Their Functions

Agency	Function
Office of Management and Budget (OMB)	Guidance and supervision of privacy protection
Office of Information Policy (OIP)	Providing privacy policy guidance and legal training
Chief Information Officers Council (CIO)	Formulating privacy policy implementation guidelines
Federal Privacy Commission (FPC)	Privacy program management and personnel collaboration
National Technical Information Service (NTIS)	Privacy protection during information preservation
General Services Administration (GSA)	Conducting privacy impact assessments

OMB is an agency within the President's Office dedicated to ensuring proposed legislation aligns with administrative policies. The 2002 E-Government Act made OMB the supreme authority for protecting U.S. citizen privacy, providing necessary guidance and oversight [57]. OIP's mission is to encourage and monitor agency compliance with the Freedom of Information Act, formulate government policy guidance on all aspects of FOIA administration, and provide legal consultation and training to agency personnel [58]. CIO Council is the primary interagency forum for improving federal information resources design, acquisition, development, modernization, use, sharing, and performance, formulating implementation guidelines for standardized digital privacy controls and agency privacy education [59]. FPC improves agency privacy program management by identifying and sharing lessons learned and best practices, promoting collaboration among agency privacy professionals to reduce unnecessary duplication and ensure effective, efficient, and consistent privacy policy implementation government-wide [60]. NTIS provides innovative data services to federal agencies through private sector partnerships to advance federal data priorities and promote economic growth, with necessary preservation measures taken at NTIS to protect data privacy [61]. GSA conducts regular reviews of Personally Identifiable Information (PII) in accordance with privacy acts to ensure personal

data integrity, accuracy, and necessity, while using Privacy Impact Assessment (PIA) as a key tool to ensure privacy issues and protections are addressed in any IT system containing PII [62].

3.2.2 Coding Analysis of Privacy Protection Substantive Policies Substantive policies refer to policies with adequate material resource investment, clear execution authorization, provision of clear benefits to policy adjustment targets or clear behavioral norms, and rigorous organizational procedures to support implementation. Based on substantive policy definitions and obtained text content, we subdivide substantive policies into five research dimensions: citizen privacy rights, data protection types, data utilization, privacy protection standards, and privacy protection technologies, as shown in Table 4 .

Table 4: Substantive Policy Level Node Coding

Dimension	Key Phrases
Citizen Privacy Rights	Protection of citizen privacy rights; DOJ's Office of Privacy and Civil Liberties drafts the Privacy Act; government prohibited from intercepting or monitoring private electronic communications without permission; electronic communication privacy rights protected; company obligations for privacy protection included in "Privacy Principles"
Data Protection Types	Refined data types with different privacy data protection policies; effective protection of personal privacy rights and personal information use on the internet; focus on improving privacy protection in electronic information communications; act requires FTC to formulate rules on personal information collection to enhance consumer privacy; personal information collected from or about children necessary for responding to judicial processes or providing information to law enforcement

Dimension	Key Phrases
Data Utilization	Privacy protection during data acquisition and utilization; core is integrating public power, strengthening acquisition, improving management, and giving the public more voice; any technology use must respect privacy, openness, and transparency; guidelines instruct agencies on processing personal information when using IT to collect new information
Privacy Protection Standards	Standards and regulations for data opening and utilization; establishing standards for sensitive information to reduce excessive concealment from the public; requiring government agencies to build an online resource center for the public and adopt new data standards to improve readability; standardizing data formats and establishing unified open data standards to enhance data usability
Privacy Protection Technologies	Application of technologies in privacy data protection domains; improving government data use and protection methods through privacy protection technologies (e.g., secure multi-party computing); protecting data integrity by incorporating state-of-the-art data security as part of IT security practices for every updated, built, or replaced system to address current and emerging threats; promoting innovation and utilizing new technologies to maintain protection

Table 4 shows that while procedural policies provide top-level design for U.S. privacy data protection, substantive policies transform procedural policy requirements from top-level design into systems that consult the public. Through node coding of privacy data protection substantive policies and content analysis of policy texts, the U.S. government data opening privacy policy system can be summarized into three categories:

(1) Strategy System: Enhancing the Core of Privacy Protection. Data use is changing the world. The federal government's unique position in provid-

ing, maintaining, and using data in society makes maintaining trust in federal data critical to democratic processes. To meet evolving data roles and democratic needs, the federal government has formulated a coordinated and integrated data strategy to better deliver data in missions, serve the public, and manage resources while respecting privacy and confidentiality [63].

Establishing PII inventories: To properly assess and mitigate privacy risks of services or programs, agencies must first know what PII can be collected, maintained, used, or disclosed (see Table 5). No PII inventory or catalog can be sufficiently complete to understand what personal information should be considered "identifiable." In common understanding, PII only includes data that can directly identify or contact individuals (e.g., names) or particularly sensitive personal data (e.g., bank account numbers). OMB and NIST define PII more broadly and dynamically. Through integration analysis of multiple fragmented data elements (e.g., age, height), PII can still be constituted. In other words, if data is linked or can be linked to specific individuals ("linkable"), it may be PII.

Table 5: PII Inventory [64]

Category	Low, Medium, High
Personal Identifiers	Name, Social Security Number, Driver's License Number, Credit Card Number, Other Financial Account Numbers (banks, etc.), Other Government IDs or Unique Identifiers
Contact Information	Email Address
Other Personal Data	Username, Avatar, Other Physical Descriptions (eye/hair color, height, etc.), Marital Status/Children/Relatives
Health, Insurance, Treatment or Medical Information	Other PII (e.g., unstructured data fields filled by users)
Biometric Data	Signature, Fingerprint, Handprint, Photos, Scans (retina, facial), Physical Movements (e.g., finger swipes, keystrokes), DNA Markers
Device-Related Data	Username, Unique Device Identifier, Location/GPS Data, Camera Controls (photos, video, videoconferencing), Microphone Controls, Audio/Voice Data, On/Off Status and Controls, Contact Lists and Directories, Other Hardware/Software Controls, Other Device Sensor Controls or Data, Data Collected by Applications (itemized), Cell Tower Records (e.g., logs, user location, time, date)

Category	Low, Medium, High
Website-Related Data	Log Data (e.g., IP address, time, date, browser type), Tracking Data (e.g., single-session or multi-session cookies, beacons), Device Settings or Preferences (e.g., security, sharing, status), Biometric Data or Related Data (see above), SD Card or Other Storage Data, Network Communication Data

Individual notification and remediation: Under Privacy Act requirements, organizations must formulate and publish remediation policies and procedures in the Federal Register for correcting or updating inaccurate, irrelevant, untimely, or incomplete data [65]. This involves: (1) **Notification**: Informing individuals about collected information, collection purposes, information usage, disclosure and sharing recipients, individual rights, available remediation plan types, information retention periods, and consequences of failing to provide requested information; (2) **Managing Privacy Complaints and Remediation**: Organizations should establish policies and procedures for managing privacy complaints or inquiries to ensure all complaints are documented, tracked, and addressed; (3) **Documentation**: Policy and procedure documents provide detailed privacy complaint resolution procedures, including remediator rights, personnel training on remediation policies and procedures, and transparency of remediation and complaint handling processes; (4) **Reporting**: Organizations should track internal and external privacy complaints to identify areas requiring further organizational attention. Notification and remediation are crucial for achieving information transparency and individual participation—two fundamental fair information practice principles.

(2) Technology System: Weaving a Protection Net for Privacy Data. The Networking and Information Technology Research and Development (NITRD) Program is the primary federal funding source for advanced IT R&D in computing, networking, and software. NITRD is one of the oldest and largest formal federal programs, coordinating multi-agency activities to address multidisciplinary, multi-technology, and multi-sector R&D needs [66]. In government open data processes, NITRD program technical support serves as strong backing for maintaining privacy data.

Privacy-preserving data matching: Record matching is typically performed between different data sources to identify common information shared across sources. However, matched records from different sources often conflict with privacy requirements for relevant data. To alleviate the contradiction between security and privacy, the U.S. federal government employs secure multi-party computing (SMC) and data sanitization methods (e.g., differential privacy and k-anonymity) during data collection to encrypt data in large datasets and reduce privacy leakage risks in data matching.

Privacy feature desensitization: In distributed environments where users must interact with many different service providers, protecting sensitive information in biometric templates becomes more complex. Federal agencies therefore adopt perceptual hashing technology, classification technology, and zero-knowledge proof of knowledge (ZKPK) protocols [67]. Under this approach, user biometric templates are processed to extract a bit string, which undergoes further processing through classification and other transformations. The resulting bit string, combined with random numbers, generates cryptographic commitments. These commitments represent identification tokens without revealing any original biometric input information, used in ZKPK protocols for user authentication.

(3) Evaluation System: Enhancing Privacy Protection Effectiveness. The U.S. has a complete privacy assessment leadership structure (see Figure 4 [Figure 4: see original paper]). The Department of Homeland Security (DHS) Privacy Office conducts digital privacy impact assessments (PIA) on technologies, rulemaking, programs, and activities regardless of data classification type to ensure privacy considerations and protections are integrated into all departmental activities [68]. PIA analyzes how personally identifiable information is collected, used, disseminated, and maintained, demonstrating that program managers and system owners consciously incorporate privacy protections throughout system or program development lifecycles [69]. PIA analysis ensures information processing complies with applicable privacy laws, regulations, and policy requirements, identifies risks and impacts of collecting, maintaining, and disseminating such information, and examines and evaluates alternative information processing protections to mitigate potential privacy risks. PIAs must be documented.

Through content coding analysis of the 52 policy texts, personal privacy protection policy formulation and implementation have been key focal points in U.S. government open data processes. The protection of personal privacy data involves five dimensions—legal system, management system, strategy system, technology system, and evaluation system—spanning the entire government open data cycle, providing policy-level guarantees for creating value from U.S. government open data. Simultaneously, this offers dual reference value for building open governments and safeguarding citizen privacy rights in China.

4 Implications of the U.S. Government Open Data Personal Privacy Protection System for China

Drawing on the U.S. government open data personal privacy policy system and measures, China can improve its government open data personal privacy protection through the following approaches:

4.1 Leveraging Multi-Stakeholder Functions

To address privacy leakage issues in data opening, the U.S. established the CIO system and DHS agencies, creating chief information officer positions dedicated

to government data opening work. Both OMB and GSA undertake review and partial regulatory functions for privacy processes. China has established the Cyberspace Administration of China and the State Internet Information Office. Therefore, in government data opening processes, these agencies should perform functions to assume responsibilities for privacy protection and cybersecurity: (1) **Issue standardized recommendations for personal privacy data**: Before obtaining personal privacy data, identify data types and develop personal privacy data project lists; when citizens submit personal data on government websites, corresponding privacy statements should inform citizens of collection purposes and authorities; (2) **Establish principles for personal privacy data utilization**: Collect and utilize personal data according to principles of transparency, individual participation, integrity, consistency, minimization, usage notification, usage limitation, and modification permission; (3) **Review and supervise data opening processes**: Privacy agencies' third responsibility is strict review and supervision throughout the entire government open data lifecycle; (4) **Strengthen staff training**: Building a professional information security and privacy protection team as active agents makes personnel training crucial, requiring enhanced training on employee privacy awareness and professional ethics; (5) **Clarify and implement 主体责任**: Privacy agencies must strictly follow relevant laws and regulations, clearly assign responsibilities to relevant entities in personal privacy leakage incidents to prevent buck-passing, and penalize violations to build transparent privacy work teams.

4.2 Constructing a Privacy Assessment System

Personal privacy data impact assessment examines the legal compliance of personal privacy data collection and utilization, judges various risks to personal privacy data subjects' legitimate rights and interests, and evaluates the effectiveness of measures to protect personal information subjects. In 2003, the U.S. issued the Privacy Impact Assessment (PIA) Guide, providing a framework for conducting privacy impact assessments—analyzing information collected, stored, protected, shared, and managed in identifiable ways to assess how PII is managed in information systems [70].

This offers important reference for constructing China's government open data privacy impact assessment system: (1) **Establish assessment entities**: Government-led privacy assessment bodies composed of experts from different fields should assess personal privacy data compliance and security in opening processes; (2) **Improve assessment measures**: Drawing on the U.S. PII inventory's classification of personal privacy data characteristics, implement different assessment standards according to personal privacy data sensitivity levels; (3) **Perfect assessment processes**: As shown in Figure 5 [Figure 5: see original paper], conduct pre-assessment of government data platform datasets to determine the importance of privacy risk impact assessment, then prepare assessments by documenting data throughout collection, processing, use, sharing, and deletion flows, analyzing impacts on personal rights and

information security at different stages to determine risks, and finally produce assessment reports.

4.3 Strengthening Data Utilization Governance

4.3.1 Building Data Inventories Include "identifiable" data elements (e.g., phone numbers, ID numbers, photos) in data inventories and rank them by personal privacy leakage risk levels. When government collects personal data, use the data inventory as reference, especially for data elements with significant privacy leakage impact. Follow data collection minimization principles by only collecting personal data necessarily connected to government functions while retaining only essential personal information to achieve data lifecycle minimization.

4.3.2 Enhancing Privacy Protection Technologies Technological advancement is crucial for improving government data opening levels and personal privacy data protection. Technical measures should be emphasized: (1) **Data encryption**: Government databases contain vast amounts of personal information, making encryption of sensitive data a critical privacy protection pathway. Using encryption functions or keys for cryptographic operations when government database management systems interface with external hardware, employing data encryption calculations to determine whether database system platforms allow data access and transmission, and timely feeding back database security risks to network data operation ports can effectively protect important government information [71]; (2) **Establishing data transmission channel protection systems**: If data transmission communication between government open data platforms and users is not secure, data can be easily intercepted, tampered with, or polluted by third parties, compromising data transmission authenticity, confidentiality, and integrity [72]. Controlling IP addresses, ports, and some transmission content between communication parties (server and client) optimizes data transmission environments, enhances transmission capabilities, and ensures secure, efficient data transmission channels; (3) **Audit logs**: Auditing stores all user database operation records in audit logs, which is important for future investigation and analysis when problems arise. When system issues occur, illegal data access times, content, and relevant persons can be quickly identified. From a software engineering perspective, current data protection through access control and data encryption alone is insufficient. Therefore, as an important supplementary measure, auditing is an indispensable part of secure database systems and the final important security line of defense for database systems.

4.4 Constructing a Privacy Computation Model Based on the Government Open Data Lifecycle

Widespread IT application greatly facilitates government data collection, storage, protection, publication, and destruction, while also making personal pri-

vacy data leakage risks pervasive throughout the entire government open data lifecycle. Drawing on U.S. practices of differential privacy, k-anonymity, data desensitization, and other technologies, China can build a privacy computation model based on the data opening cycle. Privacy computation is a computable model and axiomatic system for protecting privacy information throughout its lifecycle [73]: (1) **Privacy information collection**: Specially process (mark or encode) sensitive data containing personal privacy in government-collected information to establish privacy metadata sets; (2) **Privacy data storage**: Use technologies like deduplication on privacy variable sets established during collection to reduce duplicate homogeneous data and minimize privacy data stock; (3) **Privacy data protection**: Employ SMC, cryptography, and differential privacy to meet protection needs for retained privacy data sources; (4) **Privacy data publication**: During government data opening, use appropriate technologies like perceptual hashing, data desensitization, and ZKPK protocols based on publication information attributes, contexts, and data sensitivity levels; (5) **Privacy data destruction**: After the data opening cycle ends, use deterministic deletion technology to thoroughly destroy data containing privacy information that no longer needs to be opened. In the information age, constructing a privacy computation model based on the government open data cycle to provide a systematic, theoretical computation model for personal privacy protection is an important component of China's current open government construction.

4.5 Improving the Privacy Protection Legal System

Although China has formulated legal documents such as the Decision on Strengthening Network Information Protection and Regulations on Protecting Personal Information of Telecommunications and Internet Users in recent years, it lacks a dedicated basic law for personal information protection. Current Chinese laws on government open data focus more on information disclosure rather than data acquisition benefits. Therefore, China should formulate a Government Data Opening Law. First, legally stipulate procedures and content for government data opening to standardize processes and reduce arbitrariness while balancing "open government" and "personal privacy" conflicts. Second, legally define "personal privacy" concepts to determine violations in government open data and reduce privacy leakage risks. Finally, improve citizen "personal privacy" review and remediation systems to enable traceback analysis of violations, accurately identify privacy leakage links, conduct tracking processing, and implement remediation to protect citizens' legitimate rights. Through enacting the Government Data Protection Law, reduce personal privacy leakage risks in government open data, enhance the era value of government open data, and safeguard China's open government and service-oriented government construction.

References

[1] Office of management and budget (OMB). Open government directive, M-10-06 [EB/OL]. [2020-07-20]. http://www.whitehouse.gov/omb/assets/memoranda-2010/m1006.pdf. [2] JAATINEN T. The relationship between open data initiatives, privacy, and government transparency: a love triangle [J]. International data privacy law, 2016, 6(1): 28-38. [3] WESTIN F. Science, privacy, and freedom: issues and proposals for the 1970s. Part 1 the current impact of surveillance on privacy [J]. Columbia law review, 1966, 66(6): 1003-1050. [4] BAUMER D L, EARP J B, POINDEXTER J C. Internet privacy law: a comparison between the United States and the European Union [J]. Computers & security, 2011, 23(5): 400-412. [5] STEPHENS D. Protecting personal privacy in the global business environment: in the electronic world, protecting personally identifiable information is a critical challenge for all companies and governments [J]. Information management journal, 2007, 41(3): 56-60. [6] ARDAGNA C A, DC DV MERCATIS, PARABOSCHI S, et al. An xacml-based privacy-centered access control system [C]//Proceedings of the first ACM workshop on Information security governance. New York: ACM, 2009: 49-58. [7] 胡启平, 陈震. 试析社交网络环境中个人隐私保护 [J]. 信息网络安全, 2010(8): 43-44. [8] Freedom of information act [EB/OL]. [2020-08-01]. http://www.whitehouse.gov/the_{{{press}}{{office}}}/FreedomofInformationAct. *[9] Privacy protection act [EB/OL]. [2020-08-02]. https://www.justice.gov/opcl/overview-Privacy-act-1974-2015-edition. [10]* 谢恩平. 美国的《信息自由法》与媒体 *[D]. 北京: 中央民族大学, 2006. [11] Digital government [EB/OL]. [2020-08-03]. https://digital.gov/resources/paperwork-reduction-act-fast-track-process/. [12] Data quality act of 2000 [EB/OL]. [2020-08-13]. http://www.whitehouse.gov/sites/default/files/omb/fedreg/reproducible2.pdf. [13] E-government bill [EB/OL]. [2020-08-04]. https://obamawhitehouse.archives.gov/omb/memoranda*{m03}-22/. [14] Federal privacy act [EB/OL]. [2020-08-05]. https://www.nascio.org/resource-center/resources/federal-privacy-law-compendium-version-1-0/. [15] Open government act of 2007 [EB/OL]. [2020-08-06]. https://www.govtrack.us/congress/bills/110/s2488/text. [16] Freedom of information act (2009) [EB/OL]. [2020-08-07]. http://www.whitehouse.gov/the_{{{press}}{{o [17] Transparency and open government [EB/OL]. [2020-08-09]. http://www.whitehouse.gov/the_{{{press}}{{ *[18] Open government directive [EB/OL]. [2020-08-10]. http://www.whitehouse.gov/sites/default/files/omb/ass*06.pdf. [19] Executive order 13526 - classified national security information [EB/OL]. [2020-08-11]. https://www.whitehouse.gov/the-press-office/executive-order-classified-national-security-information. [20] Executive order 13556 - classified national security information [EB/OL]. [2020-08-11]. https://www.whitehouse.gov/the-press-office/executive-order-classified-national-security-information. [21] Office of management and budget [EB/OL]. [2020-08-14]. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/memoranda_{2010}/m10-22.pdf. [22] Office of management and budget [EB/OL]. [2020-08-14]. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/memoranda_{2010}/m10-23.pdf. [23] Consumer data privacy in a networked world [EB/OL]. [2020-08-16]. http://www.whitehouse.gov/sites/default/files/privacy-final.pdf. [24] Executive order 13642–making open and machine readable the new default for govern-

ment information [EB/OL]. [2020-08-17]. https://www.Whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government. [25] Second national action plan [EB/OL]. [2020-08-18]. http://www.whitehouse.gov/sites/default/files/docs/us_{{{national}}{{action}}}{{{plan}}{{6p}}}.pdf. [26] Executive order-making open and machine readable the new default for government information [EB/OL]. [2020-08-20]. http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-. [27] Common core metadata schema v1.0 [EB/OL]. [2020-08-22]. https://project-open-data.cio.gov/schema/. [28] RASHM K, YUKIKA A. Liberating data for public value: The case of Data.gov [J]. International journal of information management, 2016, 36(4): 668-672. [29] US-open-data-action-plan [EB/OL]. [2020-09-03]. https://www.whitehouse.gov/sites/default/files/microsites/ostp/us{{{open}}{data}}{{{action}}{plan}}.p [30] Big data, open data & the federal agencies-digital.gov [EB/OL]. [2020-09-05]. https://digital.gov/2014/07/01/big-data-open-data-the-federal-agencies/. [31] The electronic communications privacy act amendments act of 2015 [EB/OL]. [2020-09-11]. https://www.govtrack.us/congress/bills/114/hr283/text. [32] Office of management and budget [EB/OL]. [2020-09-14]. https://obamawhitehouse.archives.gov/sites/defa 15-13.pdf. [33] Open, public, electronic, and necessary government data act or the open government data act [EB/OL]. [2020-09-17]. https://www.congress.gov/bill/114th-congress/senate-bill/2852. [34] Office of management and budget [EB/OL]. [2020-09-18]. https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/ [35] National privacy strategy [EB/OL]. [2020-09-21]. https://www.nitrd.gov/PUBS/NationalPrivacyResearchS [36] Privacy guide [EB/OL]. [2020-09-24]. https://ec.europa.eu/info/sites/info/files/2016-08-01-ps-citizens-guide{en}.pdf_. [37] Cyber.dhs.gov-binding operational directive 18-01 [EB/OL]. [2020-09-30]. https://cyber.dhsgov/bod/18-01/. [38] Office of management and budget [EB/OL]. [2020-10-03]. https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2017/m-17-06.pdf. [39] Privacy protection technology [EB/OL]. [2020-10-05]. https://digital.gov/resources/privacy-preserving-collaboration-using-cryptography/. [40] Digital.government [EB/OL]. [2020-10-11]. https://digital.gov/2017/08/22/nist-crafts-next-generation-safeguards-for-information-systems-and-the-internet-of-things/. [41] Privacy guidelines [EB/OL]. [2020-10-13]. https://www.nascio.org/privacy-policy/. [42] Children's internet bill [EB/OL]. [2020-10-17]. https://www.congress.gov/bill/116th-congress/senate-bill/783?%7B%22search%22%3A%5B%22Privacy+protection%22%5D%7D&s=2&r=3. [43] Social media privacy protection and consumer rights law [EB/OL]. [2020-10-20]. https://www.congress.gov/bill/116th-congress/senate-bill/189?%7B%22search%22%3A%5B%22Privacy+ [44] Data privacy law [EB/OL]. [2020-10-24]. https://www.congress.gov/bill/116th-congress/senate-bill/583?%7B%22search%22%3A%5B%22Privacy+protection%22%5D%7D&s=2&r=17. [45] Personal health data protection act [EB/OL]. [2020-10-28]. https://www.congress.gov/bill/116th-congress/senate-bill/1842?%7B%22search%22%3A%5B%22Privacy+protection%22%5D%7D&s=2&r=50. [46] Privacy act [EB/OL]. [2020-11-02]. https://www.congress.gov/bill/116th-congress/senate-bill/1214?%7B%22search%22%3A%5B%22Privacy+protection%22%5D%7D&s=2&r=6. [47] Data protection act [EB/OL]. [2020-11-07]. https://www.congress.gov/bill/116th-congress/senate-bill/2961?%7B%22search%22%3A%5B%22Privacy+protection%22%5D%7D&s=2&r=57. [48] Cyber.dhs.gov-binding operational directive 20-01 [EB/OL]. [2020-11-10].

https://cyber.dhsgov/bod/20-01/. [49] Data security law [EB/OL]. [2020-11-13]. https://www.congress.gov/bill/116th-congress/senate-bill/4626?%7B%22search%22%3A%5B%22Privacy-
[50] Data protection act 2020 [EB/OL]. [2020-11-13]. https://www.congress.gov/bill/116th-congress/senate-bill/3300?%7B%22search%22%3A%5B%22Privacy+protection%22%5D%7D&s=2&r=32.
[51] 2020 action plan [EB/OL]. [2020-11-14]. https://strategy.data.gov/practices/.
[52] Privacy office enhancement act [EB/OL]. [2020-11-15]. https://www.congress.gov/bill/116th-congress/house-bill/5678?%7B%22search%22%3A%5B%22Privacy+protection%22%5D%7D&s=2&r=6.
[53] Application privacy and security protection law [EB/OL]. [2020-11-16]. https://www.congress.gov/bill/116th-congress/house-bill/6677?%7B%22search%22%3A%5B%22Privacy+prot
[54] 冉连, 张曦. 地方政府数据开放中的数据安全政策研究——基于全国 33 个地级市政策文本的内容分析 [J/OL]. [2020-11-23]. http://kns.cnki.net/kcms/detail/61.1167.G3.20200811.1014.002.html.
[55] President memo: Freedom of information act [EB/OL]. [2020-11-16]. https://www.justice.gov/sites/default/files/oip/legacy/2014/07/23/presidential-foia.pdf. [56] Open data policy-Managing Information as an Asset [EB/OL]. [2020-06-26]. https://project-open-data.cio.gov/policy-memorandum/. [57] Office of management and budget [EB/OL]. [2020-11-16]. https://www.whitehouse.gov/omb/memoranda_{m03}-22. [58] Organization, mission and functions manual: Office of information policy [EB/OL]. [2020-11-17]. https://www.justice.gov/jmd/organization-mission-and-functions-manual-office-information-policy. [59] Chief information officers council [EB/OL]. [2020-06-28]. https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/1151/2016/10/Standardized_{{{Digital}}*{{Privacy}}}*{Controls}.pdf.
[60] Federal privacy council [EB/OL]. [2020-11-17]. https://www.fpc.gov. [61] National technical information service [EB/OL]. [2020-11-18]. https://www.ntis.gov/privacy/index.html.
[62] General services administration [EB/OL]. [2020-11-18]. https://www.gsa.gov/reference/gsa-privacy-program. [63] Federal data strategy [EB/OL]. [2020-11-18]. https://strategy.data.gov/background/. [64] Personally identifiable information [EB/OL]. [2020-11-18]. https://www.investopedia.com/terms/p/personally-identifiable-information.asp. [65] Chief information officers council [EB/OL]. [2020-11-18]. https://www.cio.gov/policies-and-priorities/#subject=*&role=.privacy-filter&status=*. [66] Networking and information technology research and development [EB/OL]. [2020-11-19]. https://www.nitrd.gov/about/index.aspx. [67] Networking and information technology research and development [EB/OL]. [2020-11-19]. https://www.nitrd.gov/cybersecurity/nprsrfi102014/BigData-SP.pdf. [68] Chief information officers council [EB/OL]. [2020-11-19]. https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/1151/2016/10/DHS-Privacy-Office-Guide_{June}-2010.pdf. [69] Department of homeland security [EB/OL]. [2020-11-20]. https://www.dhs.gov/xlibrary/assets/privacy/privacy_{{{pia}}*{{guidance}}}*{may20
[70] Privacy impact assessment (PIA) guide [EB/OL]. [2020-11-21]. https://www.sec.gov/about/privacy/piaguide.pdf. [71] 罗静怡. 政府部门计算机网络安全中数据加密技术的运用研究 [J]. 通讯世界, 2018(04): 46-47. [72] 丁红发, 孟秋晴, 王祥等. 面向数据生命周期的政府数据开放的数据安全与隐私保护对策分析 [J]. 情报杂志, 2019, 38(7): 151-159. [73] 李凤华, 李晖, 贾焰等. 隐私计算研究范畴及发展趋势 [J]. 通信学报, 2016, 37(4): 1-11.

Author Contributions:
Chu Jiewang: Designed the research topic and guided paper revisions;

Ding Hui: Wrote and revised the paper.

Note: Figure translations are in progress. See original paper for figures.