Government data governance framework based on a data middle platform

Zijun Mao
Huazhong University of Science and Technology, Wuhan, China
Jingyi Wu
College of Public Administration, Huazhong University of Science and Technology, Wuhan, China, and
Yali Qiao and Hong Yao
Huazhong University of Science and Technology, Wuhan, China

Received 22 March 2021
Revised 25 July 2021
23 September 2021
Accepted 31 October 2021

Abstract

Purpose – The present paper constructed a new framework for government data governance based on the concept of a data middle platform to elicit the detailed requirements and functionalities of a government data governance framework.

Design/methodology/approach – Following a three-cycle activity, the design science research (DSR) paradigm was used to develop design propositions. The design propositions are obtained based on a systematic literature review of government data governance and data governance frameworks. Cases and experts further assessed the effectiveness of the implementation of the artifacts.

Findings – The study developed an effective framework for government data governance that supported the digital service needs of the government. The results demonstrated the advantages of the framework in adapting to organizational operations and data, realized the value of data assets, improved data auditing and oversight and facilitated communication. From the collection of data to the output of government services, the framework adapted to the new characteristics of digital government.

Originality/value – Knowledge of the "data middle platforms" generated in this study provides new knowledge to the design of government data governance frameworks and helps translate design propositions into concrete capabilities. By reviewing earlier literature, the article identified the core needs and challenges of government data governance to help practitioners approach government data governance in a structured manner.

Keywords Digital government, Middle platform concept, Data governance, Data governance framework, E-Government

Paper type Research paper

1. Introduction

Governments are in the process of datafication, which requires an analysis and mining of all valuable data generated in the administrative process. These data must be used in government management and public services to generate more public benefits (Dawes, 1996). Data governance is an emerging trend in government data management. Researchers and practitioners assumed that data governance referred to full decision-making authority and responsibility for data asset management (Khatri and Brown, 2010; Weber *et al.*, 2009). Successful data governance increases the speed of data processing, manipulates large data sets and adapts to the dynamic changes required by government administrative processes

This work was supported by the National Social Science Fund of China (grant number: 20ZDA038), Huazhong University of Science and Technology Double First-Class Funds for Humanities and Social Sciences; the Fundamental Research Funds for the Central Universities, HUST (grant number 2020JYCXJJ037), and the Graduates' Inovation Found, Huazhong University of Science and Technology (grant number 2021ygsCXCY).

Aslib Journal of Information Management
Vol. 74 No. 2, 2022
pp. 289-310

2050-3806
DOI 10.1108/AJIM-03-2021-0068

(Coleman *et al.*, 2009). However, governments have not been entirely successful in data governance (Thompson *et al.*, 2015). Many audits showed that data asset management in many public sectors lacks effective controls over data entry, processing and reporting. The government data governance framework is defined as enabling various levels to collaboratively manage government-wide data, and it provides the ability to align various data-related programs and information technology (IT) infrastructures with public service objectives. A robust government data governance framework (DGF) coordinates disparate data integration and addresses data quality issues, which have significant value in accelerating proactive business and addressing technical complexity (Panian, 2010). Therefore, government DGF must address technical issues and organizational challenges, such as defining organizational roles, assigning responsibilities, constructing data guidelines and conforming to organizational strategies. The present paper constructed an effective government DGF to control and expand the core organizational data assets associated with administrative processes and systems.

The Shanghai Pudong New Area proposed the "front, middle and back platform" e-government construction model for the first time in 2003. It constructed an e-government framework with the middle platform as the core in 2007. Large Internet companies currently provide commercial data services to the world by governing massive amounts of data. With the development of information systems (ISs), large Internet companies face the same information-building challenges as governments. With the increasing amount number of user data, the scope of business continues to expand. However, the traditional front- and back-end of the enterprise also are imbalanced in the matching of operating speeds (Zhiji, 2020). Chinese Internet companies are starting to plan business system frameworks and design their own information technology solutions. Alibaba explored the "big middle platform, small foreground" model for upgrading organizational structure based on the middle platform concept in 2015 (Cities, 2020). Tencent also proposed opening its capabilities in the middle platform in 2019 to help with industrial upgrading. The middle platform concept of data governance subsequently emerged in Internet enterprises (Wei *et al.*, 2020). An increasing number of enterprises are building a data middle platform-based DGF, and they have achieved remarkable results. The data middle platform reconstructs the original technical organizational structure of an enterprise and provides basic, standardized and guiding technical services for the upper-level business. The DGF of Chinese Internet enterprises and the government early governance experience based on the middle platform concept provide reference experience for the construction of government DGF (Chen *et al.*, 2020). Compared to enterprises, the development of information technology for government administrative processes requires long-term strategic planning. The government must develop a sound DGF and find a technical approach to the data middle platform.

Our research used a design science research (DSR) approach to create new artifacts of a government DGF. The design proposition was derived from a systematic literature review of the existing literature in the context of the scaling up of e-government and the iterative construction of administrative systems and platforms. The main problem addressed was facilitation of the datafication of administrative processes to provide an effective DGF for government organizational operations and public service decision-making. The structure of this paper followed the research steps of DSR. We introduced the DSR methodology in Section 2. The problem identification and motivation for the study were presented in Section 3, and the existing challenges and needs of DGF were summarized based on a systematic literature review. Section 4 described the development and application of the research topic "Data Middle Platform". Section 5 presented a description of the artifacts. Section 6 evaluated the artifacts using case studies, and Section 7 showed the evaluation results. Section 8 provided the main conclusions and contributions of this paper and some future research topics.

2. Methodology

DSR is a well-developed research paradigm in the ISs community and in management research for the creation of innovative artifacts, including frameworks (Costa *et al.*, 2020). A complete DSR study consists of three steps: discovering the motivation and contributions of the artifact, designing the artifact and evaluating the artifact (Peffers *et al.*, 2012). Hevner *et al.* (2010) viewed the process of DSR as three-cycle activities: (1) the relevance cycle initiates DSR in the context of requirements and acceptance criteria; (2) the rigor cycle connects the research project to past theories and methods; and (3) the design cycle evolves and evaluates design artifacts. The research process in this paper is based on the design science framework proposed by Hevner *et al.* (2010), which is shown in Figure 1.

First, we systematically collected and analyzed past knowledge to ensure that the artifact generated in this paper was innovative in its research and application contributions. Therefore, we applied the systematic literature review (SLR) method to define and analyze the existing DGF in relevant studies. The SLR method is a systematic, explicit and repeatable approach. The researchers reviewed the relevant literature and summarized, analyzed and synthesized a set of relevant studies to understand the gaps in existing work and topic that require further exploration (Xiao and Watson, 2017). Our SLR process followed the main guidelines in ISs research provided by Okoli and Schabram (2010): (1) search for relevant literature; (2) practical screening; (3) quality appraisal; (4) data extraction; and (5) synthesis of the identified studies. Because there were fewer government-specific DGFs, we first summarized and generalized government DGFs from the existing research. We then selected representative and valuable frameworks from related work in DGFs for inclusion in the comparisons and analyses (see Figure 2). The Web of Science (WoS) and Google Scholar search engine were used as literature search tools. A literature database was formed by searching for high-quality journal articles, conference papers and review articles using the topic "government data governance framework." We reviewed the identified studies in terms of form and content censoring. We checked the accessibility of the identified studies and excluded inaccessible and non-English articles. We analyzed whether the reviewed articles contained statements about data governance frameworks and the supporting evidence for statements. Valuable articles were added using a snowballing method based on references.

The value of the DGFs constructed from the different studies was filtered using the analysis, categorized and summarized. We analyzed 59 articles containing DGFs, and 12 articles explicitly constructed data governance frameworks, three articles used empirical evidence as support, primarily from surveys or case studies. Two of these articles were directly related to government data governance framework. We categorized, aggregated and

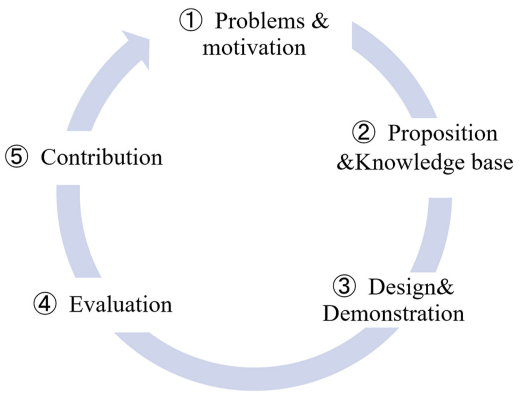

Figure 1.
The research methodology

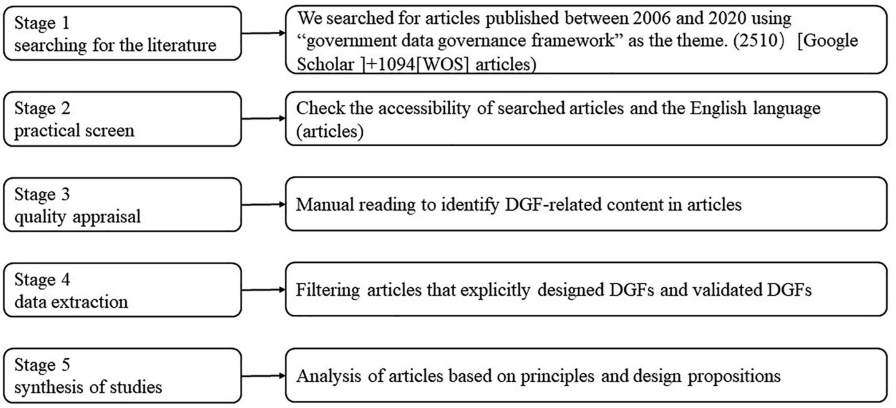

Figure 2.
The systematic literature review process

combined the data governance frameworks proposed in the articles based on the principles of data governance proposed by Brous (2016). The development of the design proposition was based on existing environmental knowledge, practical experience and the theories involved. We combined the strengths of the previous data governance framework and data middle platform and designed a new artifact of the government data governance framework to improve the functional and inherent quality deficiencies of the previous government data governance. Based on the data life cycle and data flow path of government administrative processes, we integrated the data middle platform-based DGF with the current digital government public service expansion environment.

Rigorous evaluation is an essential activity in design science. The present study evaluated the design of the artifacts based on the design proposition and the actual implementation of the artifacts. The case study is the most popular form of qualitative research in ISs research, which is demonstrated in the context of use to affect the real-world situation in organizations. In contrast to other evaluation methods, this method provides more empirical facts of the phenomena and performance (Peffers *et al.*, 2012). Our study used the construction of the government data middle platform in Hainan Province as a case to verify the feasibility and advantages of the artifact.

We further evaluated the designed framework with practical application cases using the expert scoring method. We asked the experts to rate the implementation of data governance principles by scoring the applications based on artifacts and classifying the artifacts from 1 to 5. Based on the expert feedback, we summarized the contribution of the middle platform-based government DGF.

3. Problem and motivation

3.1 Challenges faced by data governance frameworks

A DGF is a process for organizing data asset management (Sarsfield, 2009). DGFs use data aggregation and standardization to reduce the time, human and financial investments required in repeated data collection (Tomusange *et al.*, 2017), aid government decision-making via data analysis and processing, secure data and enable data evaluation. Organizational structures are often characterized by distributed development. It is necessary to ensure the ready availability, accuracy and integration of key business data (Janssen *et al.*, 2020). An imperfect organizational DGF will cause problems such as inconsistent data standards, poor data quality and low management efficiency. Therefore, DGFs are increasingly valued in government and enterprise informatics (Mcguirk *et al.*, 2015).

A total of 59 articles were reviewed for this study, 12 of which explicitly referred to a DGF (see Table 1). MS Excel spreadsheets were developed for open coding in this study. Open coding is the decomposition, examination, comparison and conceptualization of concepts and thought processes related to the research object hidden in the textual data of the literature (Corbin and Strauss, 1990). The two most prominent methods of open coding are making comparisons and asking questions. Making comparisons is the comparison of concepts with other concepts to identify similarities and differences. Asking questions categorizes concepts and create category attributes (Glaser, 1992). After coding the textual data, frequency counting is the main method of presenting the text in accordance with the relevant DGF. We used the open coding and analysis steps proposed by Finney and Corbett in the present study (Finney and Corbett, 2007). Two scorers initially analyzed a sample of literature, and a third scorer was assigned to oversee the analysis. Three personnel repeatedly discussed scoring differences and further refined and modified the scores until the scoring was consistent. The two scoring schemes were guaranteed to exceed the suggested consistency threshold (>0.8) during each iteration.

Principles are very effective in assessing the complexity and irrationality of structures and frameworks. Brous and Jassen (2016) identified four principles of data governance based on a review of previous research and devised a long list of key conceptual affiliations for data governance (see Table 2). They synthesized the literature by reviewing concepts of the papers related to "data governance" and "principles" and grouping them according to Webster and Watson's concepts (Webster and Watson, 2002). Their review grouped the material around concepts, compiled a concept matrix and developed a logical approach to grouping and presenting the key concepts that were identified and synthesized the literature by discussing each identified concept. They identified four principles around achieving the formal goals of organizational data governance, including organization, alignment, compliance and common understanding.

This list of principles provides general guidelines for our review of the existing literature related to DGFs. We compared similarities and differences of concepts in the literature based on concept lists in an iterative manner and categorized these concepts into higher abstract principles, which were considered components of the DGF (FR: fully expressing the principle; PR: partially reflecting the principle; NA: not reflecting the principle). Table 3 shows the ratio of the fit between the different DGFs studied and the four principles of data governance. Considering the length of the paper, the complete categorization is shown in Appendix 2.

For example, Khatri and Brown (2010) constructed DGF for the decision domain of organizational asset management performed well in the evaluation. They defined an organization's decision-making authority using data principles. The roles of data owners such as business managers, data architects and data security officers, data consumers and data modeling engineers were clarified and balanced in the areas of data quality, data access, data life cycle and metadata. Their resulting framework fully expressed the organizational principles. At the compliance level, accountability and data security were explicitly mentioned in the text, and data quality measurement was defined in the data quality module. However, due diligence and data privacy were not covered in the article, and the openness and policy enforcement levels were also described from the side. Therefore, the compliance principle was partially expressed. Based on the literature, we have sorted out the challenges facing DGFs.

3.1.1 Differences in objectives and context. The purpose of the new framework blueprint was to create value from big data using new analytics, based on the varying nature and mission of the organization (Kim and Cho, 2018). However, most of the currently constructed DGFs are designed to capture business value via data asset management (Panian, 2010). There is a gap between this purpose and government data governance that performs government functions and pursues public value (Rajagopalan and Vellaipandiyan, 2013). This functional difference also leads to different strategies, governance approaches and

Frameworks	Author	Year	Components
*F*1: Conceptual framework for data governance	Abraham *et al.*	2019	Governance mechanisms (structural, procedural, and relational mechanisms); organizational scope (intra-organizational and inter-organizational); data scope (data decision domains); antecedents; consequences
*F*2: DG framework for Industry 4.0	Yebenes and Zorrilla	2019	Planning; organization; operation; implementation (security and risk management, data quality, metadata; data life cycle management); monitoring
*F*3: Big data governance	Kim and Cho	2018	Objective; strategy (personal information protection, data quality level, data disclosure); components (organization, standardization and guidelines, policies and processes); IT infrastructure (audit and control, big data infrastructure, data collection/process/analysis/visualization)
*F*4: The proposed big data governance framework	Al-Badi *et al.*	2018	Communication and data management; identify organization structure; stakeholders selection; data storage; measure and monitor quality; optimize compute; policies and standards setting; big data scope determination
*F*5: Citizen-centered big data analysis-driven governance intelligence framework	Ju *et al.*	2018	Big data (data types, data sources); data-merging (populating a citizen-centered panoramic data set for each citizen); knowledge- discovery (designing citizen profile and persona via the citizen-centered data set); decision making (building ontology models supporting governance decision-making)
*F*6: Model of big data governance	Zhang *et al.*	2017	Governance objectives; the top-level design; governance objects; governance methods; the internal and external environments; contributing factors
*F*7: Data governance in smart city	Paskaleva *et al.*	2017	Data collection; data sharing and management; data identification; data generation; data use and legacy; project context
*F*8: Conceptual framework for design data governance for cloud computing services	Al-Ruithe	2016	Data governance structure; data governance assessment; data governance function; negotiation; data governance level agreement
*F*9: Big data framework for government	Rajagopalan, M. R	2013	Resource management; data organization and management; data analytics and discovery of unexplored opportunities; report–decision support and visualization
*F*10: The big data governance framework	Soares	2013	Big data types; information governance disciplines; industries and functions
*F*11: Framework for data decision domains	Khatri and Brown	2010	Data principles; data quality; metadata; data access; data lifecycle
*F*12: Building blocks of the data governance framework	Panian	2010	Data governance; standards; policies and processes; organization; data integration infrastructure

Table 1.
List of the data governance frameworks

Data governance principles	Key concepts	
Organization (*O*)	Decision rights (*O*1)	Separation of duties and concern (*O*4)
	Balanced roles (*O*2)	Improved coordination of decision making (*O*5)
	Ownership (*O*3)	Stewardship (*O*6)
Alignment (*A*)	Meeting business needs (A1)	Aligning business and IT (*A*2)
	Developing data strategy (*A*3)	Defining data quality requirements (*A*4)
	Reducing data error (*A*5)	Effective policies and procedures (*A*6)
Compliance (*C*)	Accountability (*C*1)	Openness (*C*5)
	Policy enforcement (*C*2)	Security (*C*6)
	Due diligence (*C*3)	Data quality measurement (*C*7)
	Privacy (*C*4)	
Common understanding (*U*)	Shared data commons (*U*1)	Standardized data models (*U*4)
	Use of standards (*U*2)	Standardized operations (*U*5)
	Metadata management (*U*3)	Facilitates communication (*U*6)

Table 2.
Data governance principles and the list of key concepts

Frameworks	FR:PR:NA	Frameworks	FR:PR:NA
*F*1	17:3:5	*F*7	6:0:19
*F*2	15:0:10	*F*8	9:7:9
*F*3	15:5:5	*F*9	3:2:18
*F*4	10:4:11	*F*10	13:5:7
*F*5	6:3:16	*F*11	16:5:4
*F*6	13:5:7	*F*12	17:2:6

Table 3.
The ratios of 12 typical data governance frameworks

designs in the implementation of data governance decisions and changes in the focus and formulators of the basic decision focus (Khatri and Brown, 2010). Therefore, government DGFs must specifically pursue public service value as a research objective, such as enhancing citizenship and promoting social equity.

3.1.2 Lack of cross-sectoral organizational linkage. Some large DGFs are case or domain-specific governance frameworks that do not address multisectoral data governance subjects. Although there are specific frameworks and recommendations to address data governance in cloud computing (Al-Ruithe *et al.*, 2016), smart cities (Ju *et al.*, 2018), Industry 4.0 (Yebenes and Zorrilla, 2019) and AI (Janssen *et al.*, 2020). There is a lack of frameworks that enable cross-sectoral data governance. The lack of cross-organizational sector linkage leads to barriers in long-term data collection and sharing (Paskaleva *et al.*, 2017). Inadequate sectoral collaboration leads to a lack of common understanding between stakeholders to act collectively and follow common goals and rules. The layout of government departments across regions and levels is more prone to this problem. Although organizational roles and decision rights have received much attention as an important part of the solution to this problem, data governance decisions on the separation of duties and responsibilities and data ownership-related decisions should receive further attention (Alhassan *et al.*, 2016).

3.1.3 Conceptual framework and implementation. Most of the frameworks in the current literature are conceptual in nature and were not implemented and validated (Al-Badi *et al.*, 2018). One of the typical conceptual frameworks is Abraham's framework (Abraham *et al.*, 2019). They summarized different levels of DGFs and sorted the holistic process of DGF using a literature review. Soares described the core elements in big data governance and constructed a top-level design framework for data governance (Soares, 2013). However, we do not have access to the specific implementation strategy knowledge related to the framework

(Zhang *et al.*, 2017). We should learn from the validation of the data analysis framework constructed by Ju and Kim for blood donation governance in China and the Korean National Pension Service (NPS), which highlight the potential value of the framework for public services decision-making (Ju *et al.*, 2018; Kim and Cho, 2018). The framework constructed by Paskaleva (2017), which drew on three cases and stakeholder surveys in European countries, also provides solutions that may be used for the construction of government DGF. Therefore, the construction of a DGF requires a combination of theory and practice, and a DGF that may be validated is more relevant in practice.

3.2 Current requirements for government data governance framework

There are some common requirements for data governance frameworks in government and large companies. First, a clear and logical organizational structure always determines the allocation of data governance roles and decision-making authority. A standardized process for data governance promotes consistency of action for data-related activities (Abraham *et al.*, 2019). Second, the design of the data life cycle management process and the application of technology are also inevitable considerations for data governance in government and large companies. This process must ensure data quality and data security. Finally, cross-organizational stakeholder collaboration on data is also critical to the success of data governance. The main difference in data governance needs between government and large companies is the difference in organizational scope and governance purpose (Thompson *et al.*, 2015). Large companies focus on their own data governance for economic benefits, within an organizational scope that covers head and branch offices. In contrast, government data governance covers all local governments and all types of organizations across the country. Therefore, this paper outlines the main requirements of the current government DGFs.

3.2.1 Generalizability and replicability. Governments build different administrative process systems and platforms as the scale of e-governments expand (Rajagopalan and Vellaipandiyan, 2013). Government service subsystems are constructed repeatedly, and resources are seriously wasted. Various government departments accumulate a large amount of business data, and the degree of reuse of data resources and administrative service capabilities is low. Although governments formed an information ecosystem with a complex network structure (Basmi *et al.*, 2020), the overall rationality and welfare are impaired. Over time, ISs become increasingly larger and difficult to maintain and expand and become unable to be uniformly scheduled and used. There are various hardware and software systems in various cities, and different platforms and systems must be integrated, but the cost of reconstruction is huge (Cordella and Paletti, 2019). Therefore, a common and replicable government DGF is needed.

3.2.2 Collaboration and interaction. From the perspective of basic databases, these data are not owned by certain information equipment or systems. Rather, these data are controlled by different departments and institutions. Part of the data infrastructure lacks effective organization and management, which results in data inconsistency and difficulty in data exchange (Fan *et al.*, 2014). ISs are independent, and standards are inconsistent. It is difficult to communicate and collaborate between government service subsystems. Interaction between stakeholders is also rare (Magalhaes *et al.*, 2013). There are long-term interconnection problems between various subjects of administrative processes, which lead to different degrees of process breakpoints. Structural barriers may restrict data collection and exchange in the government data framework. The competition in performance assessment in each department complicates information integration and resource circulation (Ruijer *et al.*, 2017). Data providers, demanders and managers cannot clarify their rights and responsibilities in the process of data circulation.

3.2.3 Agility and servitization. The need for agility in government services is growing as the ecosystem of government services continues to evolve. The purpose of government data

governance is to be able to quickly and flexibly use information systems to provide appropriate public services and improve the efficiency of government administrative processes (Janssen and Estevez, 2013). The demand subjects of data services change across different business sectors and service areas, including services for the public and services for special individuals or groups, and each business subject has its own data governance approach (Cao *et al.*, 2020). Because data and IT are scattered across multiple platform facilities, it cannot currently meet the multiple data needs of government practice. Government data governance cannot provide personalized services in the face of different service subjects, and the means of data governance cannot meet the objectives of business, which lead to data governance failure.

3.2.4 Quality and security. As the volume of government data increases and the range of data sources becomes broader, inconsistent data formats and poor data quality become more likely to occur. Many current government data platforms have a limited capacity to carry out data asset storage and fusion management, and they cannot fully integrate multiple data sources, which may result in duplicate data collection and inconsistent data formats and standards, leading to a lack of data standardization, completeness and timeliness (Panian, 2010). The special nature of citizen data leads to government data governance focusing on data security and privacy protection (Kim *et al.*, 2017). Multiple data sources must be accessed during the performing of government services, and there is a large amount of heterogeneous data from different manipulation systems and different types of computers in the basic government database (Thompson *et al.*, 2015). The data that are not processed will lead to a serious waste of data resources and potential data leakage (Lv *et al.*, 2018).

4. Design proposition

4.1 Why the middle platform concept?

The IBM mainframe system in the 1960s had an intermediate service transaction processing layer, the customer information control system (CICS), to manage user information (Chen *et al.*, 2020). With the expansion of information system architecture, a more complex and larger DGF is needed, and the search for a "common middleware" that can achieve integration has become an important direction for data governance. The concept of "middleware" was highlighted in a large number of data governance practices. Etymologically, the term "middle platform" comes from the efficient operation strategy of the Finnish mobile game company Supercell's "big middle platform structure" (Guangqian, 2019). Alibaba performed in-depth research on the concept and was the first to propose the middle platform concept. It reformed its corporate structure based on the "big middle platform, small foreground" organizational principle. The foreground in the middle platform concept is composed of various applications that directly reach users, and enterprises directly interact with end-users via the foreground. The background is composed of management systems, and each background business system manages a piece of the enterprise's business. The background solves the problem of enterprise management by focusing on the stability of the structure and the cautiousness of changes. Therefore, the background cannot respond to user needs with rapid innovation or support the constantly updated user needs of the foreground. Therefore, the construction of a middle platform becomes the optimal choice to connect the foreground and background.

After 2018, the middle platform derived many new concepts in practice, such as the data middle platform, organization middle platform and business middle platform, which were derived from practical experience. The core of the middle platform concept is the data middle platform. This platform is a DGF design based on data technology and management methods to achieve full life cycle data governance for organizations. The data middle platform strategy established by China's e-commerce (Internet) enterprises is highly significant. This strategy shows that China's Internet companies have moved on from the copy and imitate dilemma and started to think independently about their business system planning.

4.2 Government data middle platform

Compared to enterprises, governments are organizations that continue to exist and grow. Long-term strategic planning is required for the development of information technology business systems. Therefore, it is easy to develop a sound DGF and find a technical approach for the middle platform. It is logical that the e-government's DGF developed a middle platform concept very early. After the development of e-government to a certain stage, it was naturally faced with a strengthening of system integration, to overcome the problem of duplication of construction. Therefore, many countries, including the USA and China began their own e-government construction of a top-level architecture design in 2000 using a process naturally derived from the data middle platform (Guangqian, 2019).

At the beginning of the 21st century, the term "middle platform" already appeared in the field of government services. The Pudong New Area of Shanghai proposed the "front, middle and back platform" e-government construction model in 2003. The "Implementation Outline for Promoting E-Government Construction in Pudong New Area" further defined, the construction tasks of front offices (external services), middle offices (data exchange) and back offices (business support) (Feng *et al.*, 2009). However, at that time, the e-government construction model was limited to the construction of e-mail, official document exchange and general office systems. Since the establishment of the big data center in Shanghai in 2018, the data middle platform has been officially applied to the construction of the Shanghai One Network Office. The One Network Office government affairs middle platform continues to build the capability of the middle platform from three aspects: business, data and application that gradually form a unified standard; decoupled construction; and a mutually supportive government affairs middle platform system (Chen *et al.*, 2020). Therefore, e-government developed the middle platform thinking earlier than e-commerce at a practical level.

The government data middle platform is a DGF under multiplatforms and systems. It includes global data warehouse planning, data specification definition, data modeling research and development, data connection extraction, data operation and maintenance monitoring and data asset management tools. The government data platform was developed based on the service demand. The government middle platform is based on building and improving the government's data governance capability. The data middle platform realizes the highly centralized storage of data from all government departments, and it plays a supporting role for all business data of the government.

5. Artifact definition and design: building a government data governance framework based on a data middle platform

Many of the ISs and applications currently used by the government operate independently. The operation of different departments cannot be fully controlled and managed, and new data security issues arise. For example, the "gradient code" introduced by the Hangzhou Health Commission during the COVID-19 epidemic could leak private personal data, which highlights the potential for information leakage due to multiple information systems in the process of government data collection. Due to cross-departmental and cross-territory data barriers, the original decentralized data governance could not meet the demand for collaborative government data governance.

The core problem for the government data middle platform to solve is to make administrative processes data-based and use big data to guide government operations and decisions. With the development of digital governments, different departments face reorganization due to the improved informatization level. There is an urgent need to establish a cross-departmental and cross-regional e-government framework. The government service platform has actual needs of system integration and avoidance of redundant construction. The government DGF needs a middle platform concept. Based on the middle platform of government data, this paper constructed a government DGF (Figure 3).

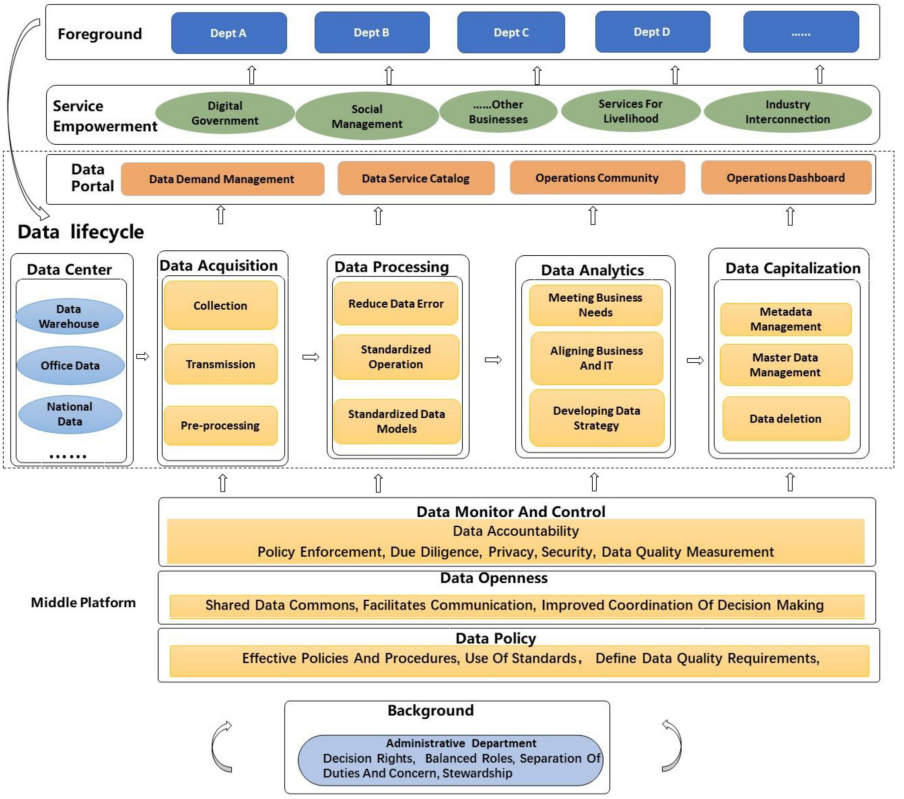

Figure 3.
The government data governance framework based on the middle data platform

5.1 Objective: service empowerment

In the implementation process, the DGF based on the data middle platform should pursue the goal of improving and enhancing the public service capability of big data and the avoidance of side effects, such as privacy leaks (Kim and Cho, 2018). By promoting the collection, development and application of big data, a governance framework for public services with full coverage, integrated utilization and unified access is formed. The government uses big data to empower administrative processes, and it supports cross-level, cross-regional, cross-system, cross-department and cross-business collaborative management by enhancing the service capacity of the information resource sharing system. These goals require the development of strategies based on the purpose of public services, accelerating data governance and promoting the development of public data applications and related industries.

5.2 Structure: define roles

The authority and responsibility of different stakeholders are clarified via the structural design to achieve collaborative decision-making among the responsible subjects (Mullon and Ngoepe, 2019; Al-Badi *et al.*, 2018). The foreground provides access for public service users, primarily via the constructed portals. The foreground must be constantly updated and iterated to meet the new demands of public services and social governance. The background is biased toward decision-making institutions, including office application systems for

departmental public officials (Cheong and Chang, 2007). The middle platform organically connects the foreground and background, realizing the coherence and collaboration between these areas. The middle platform contains several functions, such as portal support, authentication authorization, basic services, collaboration support, data exchange and data resources.

Using this structure, the executive sponsor, data governance leader, data governance council and data governance office are in a stable background and bear responsibility for the planning and coordinating of government data management plans and strategic directions (Weber *et al.*, 2009). The business units in the foreground are data producers, data owners and data consumers. They units aggregate and create and apply data. There are more technical staff such as data stewards, business data stewards and technical data stewards, in the middle platform who are responsible for the actual operation of data services and the development of technical specifications (Cheong and Chang, 2007). The Chief Data Officer (CDO), Chief Information Officer (CIO) and Chief Service Officer (CSO) assume the management decision structure in the data middle platform (Janssen *et al.*, 2020).

5.3 Procedure: policies and standards

The framework constructs standards, processes, policies and regulations for data governance (Alhassan *et al.*, 2019). First, it establishes a decision-making mechanism for data governance, introduces a data ownership list, delineates data rights and responsibilities, constructs effective data governance procedures and clarifies the division of roles and functions between the competent department and other departments. This construct helps to trace data responsibilities and regulate the individual behavior of organizational members. Second, standards and norms for data governance should be developed, unified operational standards and norms for improving data quality should be published and regulations for data security protection should be improved to enable data quality assessment.

Governments introduce evaluation regulations for data governance and establish corresponding assessment methods in conjunction with departmental and individual performance. It also publicizes to all departments the punishment for data transgression operations, using institutional norms to restrain the data governance process, which ensures data governance compliance (Tallon *et al.*, 2013).

5.4 Openness: communication and sharing

The openness of the framework is reflected in the communication and sharing levels. The communication aspect focuses on raising consistent and positive perceptions and building common awareness of data governance between stakeholders within and outside of government. Sharing is primarily achieved via interagency data sharing (IDS) and open government data (OGD). IDS emphasizes government departments as the main body and data as the carrier to achieve the interconnectivity and interoperability of government information and services, which builds an intelligent, responsive and efficient digital government (Barki and Pinsonneault, 2005). Through OGD, the government provides open data for public use (Nikiforova and McBride, 2021) uses data to create value, promotes the use of public sector information (Leviäkangas and Molarius, 2020), realizes cooperation with citizens, businesses and other subjects and enhances administrative transparency and social trust (Ubaldi, 2013).

5.5 Compliance: monitoring and control

Compliance aims to promote and implement data governance policies and oversee data governance processes by performing data monitoring and control. It focuses on data security

protection and data quality assessment (Al-Ruithe *et al.*, 2019). Data security protection specifically includes personal privacy protection, data security audits, data operation supervision, the organization of departmental assessment and due diligence. It promotes the orderly flow of data, achieves unified authority management at the security level and guarantees that data are securely protected in the process of transmission and governance. Data quality assessment specifically includes data integrity testing, data accuracy testing, data validity testing, data timeliness testing and data consistency testing (Khatri and Brown, 2010). Data quality evaluations are carried out based on test data quality indicators and evaluation rules (Strong *et al.*, 1997). The data are analyzed, scored, weighted and scored to obtain the data quality evaluation results.

5.6 Execution: data life cycle

Data governance is the organizational logic of the life cycle of collecting, processing, analyzing, storing, sharing and destroying data (Janssen *et al.*, 2020). First, data sources include data from government departments and data scattered across the Internet, individuals, groups and enterprises (Janssen *et al.*, 2017). Data collection may be decomposed into three subtasks, data collection, data transmission and data preprocessing (Munshi and Mohamed, 2017). Second, the data are processed via data cleaning, data fusion and other technical means. The basis of data processing is data planning, which primarily includes two aspects: indicators and dimensions. The indicators are selected based on specific business needs. The principle of selecting dimensions is to record the dimensions that affect the indicators. Third, the data analysis system uses data visualization services, report services and map services to prompt users to browse related information and analyze related business data for multiple users via multiple collaborative suites. Data analysis packages data services based on basic and thematic databases and customizes portals according to functional authority, usage scenarios and usage habits.

The data capitalization system contains metadata management, data queries and data life cycle management. A data asset warehouse is established to ensure that government managers can control the progress of data capitalization circulation in real time, trace the upstream and downstream flow of data via "blood relationship" in metadata management, track the sharing status of data assets, back up the flow information and clarify the open catalog of data assets and to correlate relevant departments and clarify the ownership and decision-making power of data. The data portal is responsible for the integration of applications and information, and it is the user portal of the big data platform. The portal includes system management, single sign-on, log management, multiapplication integration and other functions.

6. Case test: government data middle platform in Hainan Province

This paper constructed a DGF for government data based on a data middle platform that was applied for data governance in government administration processes. The data governance case of data middle platform construction in Hainan Province was used to verify the feasibility of the framework constructed in this study, and the advantages of this DGF were validated from the practical experience of data governance in Hainan Province (see Figure 4).

Hainan Province released the "Smart Hainan Master Plan (2020–2025)" in July 2020, which proposed building a smart data middle platform in Hainan Province to form cross-level, cross-regional, cross-sector, cross-business, cross-system data integration and intelligent application capabilities and establish a standardized public data resource pool in Hainan Province that continuously improves the data resource governance system in Hainan Province and provide a basis and specification for the full-lifecycle management of data

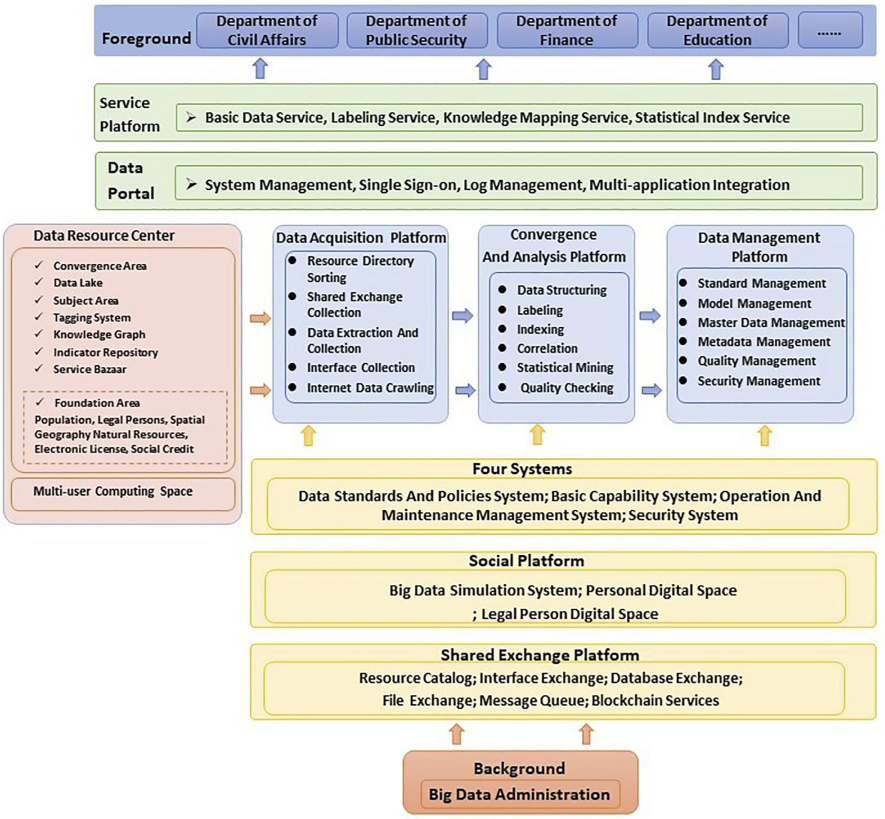

Figure 4.
One center, six platforms, four systems and a unified data portal

resources in Hainan Province (Hainan, 2020). The construction of Hainan Province's data middle platform was promoted so that all kinds types of data resources may be shared and mutually recognized on the province's integrated platform to realize the "standardization, intelligence and service" of data and enable full life cycle governance of data tracking.

The data middle platform in Hainan Province is more mature than the earlier data middle platforms in Shanghai and Zhejiang. The earlier data middle platforms provide technical services for administrative processes, but the current data middle platform transforms and expands administrative processes from technical and organizational levels. Past, data middle platform only constructed a system-level basic support layer to meet the data sharing and data opening needs of various departments. It primarily provided various instrumental basic data services for government public service departments and promoted the development of office automation. The data middle platform of Hainan Province is a reconfiguration of the whole framework from the administrative process and technical application level. It focuses on realizing the data governance of the data life cycle and expanding and developing government public services via the synergy of data and administrative processes. For example, the construction of the middle platform was accompanied by the elimination of data departments in each administrative business department in Hainan Province. The government uniformly adopted the Big Data Bureau as the management department of the data middle platform. From the organizational level, the conceptual reform of the data

middle platform achieved a shift to holistic data governance. For technology and management, Hainan Province realized the use of data-based thinking for administrative processes and technology applications, and it is representative and typical.

6.1 The data resource center

The construction of the government data middle platform in Hainan Province was performed around the data resource center. The main function of the data resource center is the construction of a data lake by integrating multiple data sources via the aggregation of government administration processes and social data. It dynamically collects structured and unstructured data from various departments, the Internet, the internet of Things, third parties, districts and counties. The data resource center is divided into two areas. The real-time area ensures that data are available in real time to meet the needs of applications and services. The offline area houses the basic database, tagging system, statistical indicators and relational mapping data.

6.2 The six platforms

The data collection platform is responsible for aggregating data from data sources, while providing data from the data resource center to the public in the form of files, databases, messages or application programming interfaces (APIs). The implementation content included resource directory sorting, shared exchange collection, data extraction and collection, interface collection and Internet data crawling. The convergence and analysis platform identified the underlying entities (e.g. the population, legal persons and electronic certificates) to which these data corresponded. Data from the same entity were correlated, and data attributes were compared in terms of values and logical relationships to discover the parts of data from different sources that corroborated (good data) and contradicted each other (bad data). Then, the inconsistent data were sent to the data source unit for repair. The data management platform contained metadata management. It tracked the status of data asset sharing, backed up the flow of information and clarified the open catalog of data assets. According to the needs of applications, the service platform encapsulated the basic data, statistical indicators, labels, relational mapping data, thematic data services and thematic data services in the basic library and provided services to the public via the data sharing and exchange platform to realize the service of data. Based on the completed trusted exchange platform, the social platform exchanged the population and legal person data in the data resource center to the big data simulation platform via the trusted exchange platform by constructing a big data simulation platform for each enterprise to build big data application system.

6.3 Four systems

The Big Data Bureau of Hainan Province, as the organizational management background of data governance, constructed the four systems, including standards and policy systems, basic capabilities systems, operations and maintenance (O&M) management systems and security systems. *The standard and policy system* establishes a comprehensive evaluation system of the province's informational and big data development level, and it regularly releases evaluation reports. The government incorporated data cooperation into the departmental assessment system and used performance appraisal. The Big Data Bureau regularly releases, tracks and evaluates the implementation of data governance policies and the effects of implementation. *The basic capability system* provides visual capability models for each phase of the data middle platform. *The operation and maintenance system* provides data governance-related knowledge queries, performance evaluations and service complaint functions for multiple business departments. *The security system* supports relevant research

institutions to evaluate the operation of government data centers, which improves data quality and strengthens data risk control. Data governance of the entire government administrative business process is managed and monitored centrally.

6.4 The unified data portal

The unified data portal performs in-depth mining and analysis using various display tools for each basic library and subject library. This portal provides data applications, including single sign-on construction, data demand management, base library query and data operation analysis. The unified data portal adds authentication and basic management functions to support the management and service of each application system of the platform.

7. Evaluation outcomes

This paper evaluated the design artifacts based on the data middle platform application in Hainan Province. Interviews with eight expert scholars and practitioners evaluated the utility of the design artifacts using the achievement of the four principles as criteria. The preliminary interview transcripts and summarized evaluation results are presented below. Appendix 2 shows the average scores of the eight experts on the design. Most of the scorers strongly endorsed the importance of the four principles in government data governance and scored each principle as "4 = very important" or "extremely important". Most of the experts rated the implementation level as "4 = moderately easy" and "5 = very easy" with the exception of "*O*3 = Use of standards, *O*6 = Stewardship, *A*6 = Effective policies and procedures, *C*6 = Security, and *U*6 = Facilitates communication." The relatively low scores were assigned because most experts believed that it was difficult to fully implement these elements via the design of a DGF alone, and further design and planning were needed from a technical and management perspective. Some experts also believed that the framework could be further enhanced by adapting it to the current situation of the Hainan government. Overall, the feedback on the new framework artifacts was very positive, with most of the raters agreed that the framework was able to achieve the principles. Therefore, the validation of a case study of a government data middle platform in Hainan Province showed that the framework was effective and mature, with the right components and skills.

7.1 Mechanisms for organization

To build a smarter, more responsive and efficiently governed digital government, the middle platform-based government DGF uses data as a carrier to achieve the interconnectivity and interoperability of government information by integrating cross-level and cross-departmental government data resources (Barki and Pinsonneault, 2005). Accountability for the division of roles and responsibilities within the organization promotes security awareness and behavior between government data stakeholders. The current return on investment (ROI) of big data operations is not immediate and difficult to estimate. A good and mature DGF can make the effectiveness of data governance an important measure of an organization's business performance.

7.2 Mechanisms for alignment

The data middle platform realized data refinement operations via data collection, cleaning, management and analysis, and it creates structured dynamic combinations to adapt and respond to new demands for government services and dynamically meet unpredictable public service expectations (Ciborra, 1996). Managers are expected to identify and classify data via data governance and perform guided and automated data usage assessments based on administrative processes. The data asset system constructed several thematic databases

based on the types of public services and key areas of social development, and it adopted the data supermarket approach for scenario-based approval, which reduced the difficulty of approval and improved operational efficiency.

7.3 Mechanisms for compliance

The government DGF sets data classification standards, quality assessment and policies based on government administrative processes and public service requirements. It established a unified data standard, simplified complex administrative business logic and lowered the threshold of data development and application. However, government DGF promoted IT security via the proactive monitoring of data quality and security standards. The framework clearly defined the authority and definition of data policies and procedures. Consistency in data policies and clarity in data quality assessment methods provide constraints on risks and transgressions in the government data governance process (Barki and Pinsonneault, 2005).

7.4 Mechanisms for common understanding

The DGF provides information related to IT operations and publishes actionable experiences and insights within the government on public service data management that are high quality, trusted and valuable. Collaboration within government is achieved via knowledge transfer within the organization. Government data management departments and executives have also mastered unified data governance processes based on the data life cycle and data-related administrative business terminology. The management of metadata must be standardized and evolve with the changing goals of government public services (Ciborra, 1996).

8. Conclusion and discussion

Following the DSR approach, this study designed and evaluated the artifacts of a DGF for government based on a data middle platform. This study summarized and analyzed the challenges of previous data governance research and the need for a government DGF using a systematic literature review. We constructed a government DGF with the middle platform concept as its core. Government public services departments were the business foreground in this framework, the administrative department was the background, and the data middle platform became the link between the foreground and background to enable reusable data governance. We evaluated the practicality and effectiveness of the design proposition using a case study of a big data middle platform for government services in Hainan Province, China, and expert ratings from actual practitioners and government staff. The results showed the government DGF based on the data middle platform was theoretically consistent with the four principles of data governance, and it was also effective and feasible in practice.

However, this study has some limitations. There are two major limitations to this study. First, the study focused solely on the peer-reviewed academic literature in the field of data governance during the systematic literature regression phase and ignored practitioner publications. In contrast, institutions such as Informatica and DAMA may provide further insight into data governance initiatives from a practitioner's perspective. Second, we did not describe all of the articles reviewed on DGFs in detail due to length limitations of the article. Although we worked very hard in designing DGFs in accordance with the requirements of digital government without sacrificing its prevalence in various government administrative processes and public services, a specific DGF can address specific situational government governance issues. Finally, the DGF constructed in this study needs further case studies to verify the effectiveness of the framework, which also needs to be continuously improved with technological and organizational development.

References

Abraham, R., Schneider, J. and Vom Brocke, J. (2019), "Data governance: a conceptual framework, structured review, and research agenda", *International Journal of Information Management*, Vol. 49, pp. 424-438.

Al-Badi, A., Tarhini, A. and Khan, A.I. (2018), "Exploring big data governance frameworks", *Procedia Computer Science*, Vol. 14, pp. 1271-1277.

Al-Ruithe, M., Benkhelifa, E. and Hameed, K. (2016), "A conceptual framework for designing data governance for cloud computing", *Procedia Computer Science*, Vol. 94, pp. 160-167.

Al-Ruithe, M., Benkhelifa, E. and Hameed, K. (2019), "A systematic literature review of data governance and cloud data governance", *Personal and Ubiquitous Computing*, Vol. 23 No. 5, pp. 839-859.

Alhassan, I., Sammon, D. and Daly, M. (2016), "Data governance activities: an analysis of the literature", *Journal of Decision Systems*, Vol. 25 Sup1, pp. 64-75.

Alhassan, I., Sammon, D. and Daly, M. (2019), "Critical success factors for data governance: a theory building approach", *Information Systems Management*, Vol. 36 No. 2, pp. 98-110.

Barki, H. and Pinsonneault, A. (2005), "A model of organizational integration, implementation effort, and performance", *Organization Science*, Vol. 16 No. 2, pp. 165-179.

Basmi, W., Boulmakoul, A., Karim, L. and Lbath, A. (2020), "Modern approach to design a distributed and scalable platform architecture for smart cities complex events data collection", *Procedia Computer Science*, Vol. 17, pp. 43-50.

Brous, P., Janssen, M. and Vilminko-Heikkinen, R. (2016), "Coordinating decision-making in data management activities: a systematic review of data governance principles", *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.

Cao, Q.H., Giyyarpuram, M., Farahbakhsh, R. and Crespi, N. (2020), "Policy-based usage control for a trustworthy data sharing platform in smart cities", *Future Generation Computer Systems*, Vol. 107, pp. 998-1010.

Chen, B., Haibin, F. and Zhao, W. (2020), "A preliminary study on the role of the government affairs central platform in the construction of digital government", *Journal of East China Normal University (Natural Science Edition)*, Vol. 5, pp. 131-136.

Cheong, L.K. and Chang, V. (2007), "The need for data governance: a case study", *ACIS 2007 Proceedings - 18th Australasian Conference on Information Systems*, pp. 999-1008.

Ciborra, C.U. (1996), "The platform organization: recombining strategies, structures, and surprises", *Organization Science*, Vol. 7 No. 2, pp. 103-118.

Cities, S.H. (2020), *'Zhejiang Government Nail': Online Government Affairs Collaboration Platform*, China Construction Information Technology, Beijing, Vol. 3, pp. 54-55.

Coleman, D.W., Hughes, A.A. and Perry, W.D. (2009), "The role of data governance to relieve information sharing impairments in the federal government", *2009 WRI World Congress on Computer Science and Information Engineering*, pp. 267-271.

Corbin, J. and Strauss, A. (1990), "Grounded theory research: procedures, canons and evaluative criteria", *Zeitschrift für Soziologie*, Vol. 19 No. 6, pp. 418-427.

Cordella, A. and Paletti, A. (2019), "Government as a platform, orchestration, and public value creation: the Italian case", *Government Information Quarterly*, Vol. 36 No. 4, p. 101409.

Costa, E., Soares, A.L. and de Sousa, J.P. (2020), "Industrial business associations improving the internationalisation of SMEs with digital platforms: a design science research approach", *International Journal of Information Management*, Vol. 53, p. 102070.

Dawes, S.S. (1996), "Interagency information sharing: expected benefits, manageable risks", *Journal of Policy Analysis and Management*, Vol. 15 No. 3, pp. 377-394.

Fan, J., Zhang, P. and Yen, D.C. (2014), "G2G information sharing among government agencies", *Information and Management*, Vol. 51 No. 1, pp. 120-128.

Feng, S., Luoying, C. and Jianhua, D. (2009), "Compatible and coexisting 'together' - analysis of the construction model of Pudong new area E-government platform with 'middle platform' as the core", *Shanghai Informatization*, Vol. 12, pp. 68-70.

Finney, S. and Corbett, M. (2007), "ERP implementation: a compilation and analysis of critical success factors", *Business Process Management Journal*, Vol. 13 No. 3, pp. 329-347.

Glaser, B. (1992), *Basics of Grounded Theory Analysis Emergence vs Forcing*, Sociology Press, Mill Valley, CA.

Guangqian, L. (2019), "What is a data middle platform?", *China Information Community*, Vol. 6, pp. 72-75.

Hainan, G. (2020), "Smart Hainan master plan (2020-2025)", available at: https://www.hainan.gov.cn/hainan/ztzc/202008/6758897ae22d4b6abb89387b2c41019d.shtml (accessed 26 June 2021).

Hevner, A. and Chatterjee, S. (2010), "Design science research in information systems", in Hevner, A. and Chatterjee, S. (Eds), *Design Research in Information Systems: Theory and Practice*, Springer, Boston, MA, pp. 9-22.

Janssen, M. and Estevez, E. (2013), "Lean government and platform-based governance-Doing more with less", *Government Information Quarterly*, Vol. 30 Suppl. 1, pp. S1-S8.

Janssen, M., Matheus, R., Longo, J. and Weerakkody, V. (2017), "Transparency-by-design as a foundation for open government", *Transforming Government: People, Process and Policy*, Vol. 11 No. 1, pp. 2-8.

Janssen, M., Brous, P., Estevez, E., Barbosa, L.S. and Janowski, T. (2020), "Data governance: organizing data for trustworthy artificial intelligence", *Government Information Quarterly*, Vol. 37 No. 3, p. 101493.

Ju, J., Liu, L. and Feng, Y. (2018), "Citizen-centered big data analysis-driven governance intelligence framework for smart cities", *Telecommunications Policy*, Vol. 42 No. 10, pp. 881-896.

Khatri, V. and Brown, C.V. (2010), "Designing data governance", *Communications of the ACM*, Vol. 53 No. 1, pp. 148-152.

Kim, H.Y. and Cho, J. (2018), "Data governance framework for big data implementation with NPS Case Analysis in Korea", *Journal of Business and Retail Management Research*, Vol. 12 No. 3, pp. 36-46.

Kim, D., Kim, S., Hwang, I. and Kim, T. (2017), "Why not comply with information security? An empirical approach for the causes of non-compliance", *Online Information Review*, Vol. 41 No. 1, pp. 2-18.

Leviäkangas, P. and Molarius, R. (2020), "Open government data policy and value added - evidence on transport safety agency case", *Technology in Society*, Vol. 63, p. 101389.

Lv, Z., Li, X., Wang, W., Zhang, B., Hu, J. and Feng, S. (2018), "Government affairs service platform for smart city", *Future Generation Computer Systems*, Vol. 81, pp. 443-451.

Magalhaes, G., Roseira, C. and Strover, S. (2013), "Open government data intermediaries: a terminology framework", *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance*, Association for Computing Machinery, Seoul, Republic of Korea, pp. 330-333.

Mcguirk, P.M., O'Neill, P.M. and Mee, K.J. (2015), "Effective practices for interagency data sharing: insights from collaborative research in a regional intervention", *Australian Journal of Public Administration*, Vol. 74 No. 2, pp. 199-211.

Mullon, P.A. and Ngoepe, M. (2019), "An integrated framework to elevate information governance to a national level in South Africa", *Records Management Journal*, Vol. 29 Nos 1-2, pp. 103-116.

Munshi, A.A. and Mohamed, Y.A.R.I. (2017), "Big data framework for analytics in smart grids", *Electric Power Systems Research*, Vol. 151, pp. 369-380.

Nikiforova, A. and McBride, K. (2021), "Open government data portal usability: a user-centred usability analysis of 41 open government data portals", *Telematics and Informatics*, Vol. 58, p. 101539.

Okoli, C. and Schabram, K. (2010), *A Guide to Conducting a Systematic Literature Review of Information Systems Research*, Concordia University, SSRN Electronic Journal, Rochester.

Panian, Z. (2010), "Some practical experiences in data governance", *World Academy of Science, Engineering and Technology*, Vol. 62 No. 1, pp. 939-946.

Paskaleva, K., Evans, J., Martin, C., Linjordet, T., Yang, D. and Karvonen, A. (2017), "Data governance in the sustainable smart city", *Informatics*, Vol. 4 No. 1, pp. 1-19.

Peffers, K., Rothenberger, M., Tuunanen, T. and Vaezi, R. (2012), *Design Science Research Evaluation*, Springer, Berlin, Heidelberg, pp. 398-410.

Rajagopalan, M.R. and Vellaipandiyan, S. (2013), "Big data framework for national E-governance plan", *2013 Eleventh International Conference on ICT and Knowledge Engineering*, pp. 1-5, doi: 10.1109/ICTKE.2013.6756283.

Ruijer, E., Grimmelikhuijsen, S., Hogan, M., Enzerink, S., Ojo, A. and Meijer, A. (2017), "Connecting societal issues, users and data. Scenario-based design of open data platforms", *Government Information Quarterly*, Vol. 34 No. 3, pp. 470-480.

Sarsfield, S. (2009), *The Data Governance Imperative*, IT Governance, Ely.

Soares, S. (2013), *IBM Infosphere: A Platform for Big Data Governance and Process Data Governance*, MC+Press, Boise Idaho, *Public.dhe.ibm.com*.

Strong, D.M., Lee, Y.W. and Wang, R.Y. (1997), "Data quality in context", *Communications of the ACM*, Vol. 40 No. 5, pp. 103-110.

Tallon, P.P., Ramirez, R.V. and Short, J.E. (2013), "The information artifact in IT governance: toward a theory of information governance", *Journal of Management Information Systems*, Vol. 30 No. 3, pp. 141-178.

Thompson, N., Ravindran, R. and Nicosia, S. (2015), "Government data does not mean data governance: lessons learned from a public sector application audit", *Government Information Quarterly*, Vol. 32 No. 3, pp. 316-322.

Tomusange, I., Yoon, A. and Mukasa, N. (2017), "The data sharing practices and challenges in Uganda", *Proceedings of the Association for Information Ence and Technology*, Vol. 54 No. 1, pp. 814-815.

Ubaldi, B. (2013), "Open government data: towards empirical analysis of open government data initiatives", *Opendataforum Info*, Vol. 3 No. 27, pp. 11-15.

Weber, K., Otto, B. and Österle, H. (2009), "One size does not fit all—A contingency approach to data governance", *Journal of Data and Information Quality*, Vol. 1 No. 1, Article 4.

Webster, J. and Watson, R.T. (2002), "Analyzing the past to prepare for the future: writing a literature review", *MIS Quarterly*, Vol. 26 No. 2, pp. xiii-xxiii.

Wei, L., Wenjian, X., Peng, L., Zhibin, Z. and Yonggang, Z. (2020), "Research on the evolution of information system architecture based on business middle platform", *Telecommunications Engineering Technology and Standardization*, Vol. 33 No. 11, pp. 8-12.

Xiao, Y. and Watson, M. (2017), "Guidance on conducting a systematic literature review", *Journal of Planning Education and Research*, Vol. 39 No. 1, pp. 93-112.

Yebenes, J. and Zorrilla, M. (2019), "Towards a data governance framework for third generation platforms", *Procedia Computer Science*, Vol. 151, pp. 614-621.

Zhang, S., Gao, H., Yang, L. and Song, J. (2017), "Research on big data governance based on actor-network theory and Petri nets", *2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pp. 372-377.

Zhiji, Q. (2020), "Talking about the way to build a data middleware system", *Information Technology and Informatization*, No. 10, pp. 93-95.

Appendix 1

C	Organization (O) O1	O2	O3	O4	O5	O6	Alignment (A) A1	A2	A3	A4	A5	A6	Compliance (C) C1	C2	C3	C4	C5	C6	C7	Common understanding (U) U1	U2	U3	U4	U5	U6	FR:PR:NA
F1	FR	FR	FR	NA	NA	FR	NA	NA	FR	FR	PR	FR	FR	FR	PR	FR	PR	FR	FR	FR	FR	FR	NA	FR	FR	17:3:5
F2	FR	NA	NA	FR	NA	NA	FR	FR	FR	FR	FR	FR	NA	FR	NA	NA	NA	FR	FR	NA	FR	FR	FR	FR	NA	15:0:10
F3	NA	FR	PR	NA	PR	NA	FR	FR	PR	FR	PR	FR	FR	NA	NA	FR	FR	FR	FR	FR	FR	FR	FR	FR	PR	15:5:5
F4	FR	NA	NA	NA	NA	FR	NA	FR	FR	FR	NA	PR	PR	PR	NA	FR	NA	FR	NA	PR	FR	FR	NA	NA	FR	10:4:11
F5	FR	NA	NA	NA	FR	NA	FR	PR	NA	NA	NA	NA	PR	NA	NA	NA	NA	NA	NA	PR	FR	NA	FR	FR	NA	6:3:16
F6	FR	FR	PR	NA	FR	FR	PR	PR	NA	FR	NA	FR	FR	FR	NA	FR	NA	FR	FR	NA	PR	FR	NA	PR	FR	13:5:7
F7	NA	NA	NA	NA	NA	NA	FR	FR	NA	NA	NA	NA	NA	NA	NA	NA	FR	NA	NA	NA	NA	FR	FR	NA	FR	6:0:19
F8	FR	FR	NA	NA	NA	PR	PR	FR	FR	PR	NA	PR	FR	FR	NA	FR	FR	FR	PR	NA	PR	NA	NA	NA	PR	9:7:9
F9	NA	NA	NA	NA	NA	NA	FR	FR	PR	NA	NA	NA	NA	NA	NS	NS	PR	NA	NA	NA	NA	NA	FR	NA	NA	3:2:18
F10	FR	PR	FR	NA	NA	FR	FR	NA	FR	FR	PR	PR	FR	FR	NA	FR	NA	FR	FR	NA	PR	FR	NA	PR	FR	13:5:7
F11	FR	FR	FR	FR	FR	FR	FR	FR	FR	FR	FR	PR	FR	PR	NA	NA	PR	FR	FR	NA	PR	FR	NA	FR	PR	16:5:4
F12	NA	FR	NA	FR	NA	FR	FR	FR	PR	FR	FR	FR	FR	FR	NA	NA	PR	FR	FR	FR	FR	FR	FR	FR	NA	17:2:6

Note(s): FR: fully expressing the principle; PR: partially reflecting the principle; NA: not reflecting the principle

Table A1. Scoring of the 12 frameworks in terms of the four principles

Appendix 2

Design principle	*O*1	*O*2	*O*3	*O*4	*O*5	*O*6	*A*1	*A*2	*A*3	*A*4	*A*5	*A*6
Importance	4	5	5	5	5	4	4	4	5	4	4	5
Implementation	4	5	3	5	4	3	4	4	5	4	4	3

Design principle	*C*1	*C*2	*C*3	*C*4	*C*5	*C*6	*C*7	*U*1	*U*2	*U*3	*U*4	*U*5	*U*6
Importance	5	5	5	4	5	5	5	5	5	4	5	4	5
Implementation	5	5	4	4	5	3	4	5	5	4	4	5	3

Note(s): Importance = Level of importance of the principles: (1 = not at all important; 2 = slightly important; 3 = moderately important; 4 = very important; 5 = extremely important)
Implementation = Level of implementation of the artifact: (1 = very difficult; 2 = moderately difficult; 3 = neither easy nor difficult; 4 = moderately easy; 5 = very easy)
O = Organization: (*O*1 = Decision rights; *O*2 = Balanced roles *O*3 = Ownership; *O*4 = Separation of duties and concern; *O*5 = Improved coordination of decision making; *O*6 = Stewardship)
A = Alignment: (*A*1 = Meeting business needs; *A*2 = Aligning business and IT; *A*3 = Developing data strategy; *A*4 = Defining data quality requirements; *A*5 = Reducing data error; A6 = Effective policies and procedures)
C = Compliance: (*C*1 = Accountability; *C*2 = Policy enforcement; *C*3 = Due diligence; *C*4= Privacy; *C*5 = Openness; *C*6 = Security; *C*7 = Data quality measurement)
U = Common understanding: (*U*1 = Shared data commons; U2 = Use of standards; U3 = Metadata management; *U*4 = Standardized data models; *U*5 = Standardized operations; *U*6 = Facilitates communication)

Table A2.
The evaluation outcomes of the design

About the authors
Zijun Mao is a professor of Nontraditional Security Research Center at Huazhong University of Science and Technology.

Jingyi Wu is a PhD student in College of Public Administration, Huazhong University of Science and Technology. His research interests include digital government, nontraditional security. Jingyi Wu is the corresponding author and can be contacted at: theresewu@hust.edu.cn

Yali Qiao is a postgraduate student in College of Public Administration, Huazhong University of Science and Technology. Her research interests include digital government and nontraditional security.

Hong Yao is a PhD student in College of Public Administration, Huazhong University of Science and Technology. His research interests include digital government and nontraditional security.