Postprint: System Dynamics Simulation Study on Factors Influencing Scientific Data Value-Added

Authors: Sun Lili, Wang Weijie, Sheng Jiefei

Date: 2024-04-03T00:00:00+00:00

Abstract

[Objective/Significance] To investigate the factors influencing scientific data value-added and reveal the intrinsic development patterns of scientific data value-added. [Method/Process] Based on expert interview materials and literature related to scientific data value-added, the grounded theory method was employed to develop a theoretical model of the influencing factors for scientific data value-added. On this basis, a system dynamics model for scientific data value-added was constructed to conduct dynamic simulation and analysis of the scientific data value-added process, revealing the functional relationships of various influencing factors on scientific data value-added. [Results/Conclusions] In the process of scientific data value-added, raw data quality serves as a prerequisite, the level of data deposit has a significant impact, data organization and integration are crucial for value-added formation, and scientific data sharing and development are essential for value-added realization.

Full Text

Preamble

Title: System Dynamics Simulation Study on Influencing Factors of Scientific Data Value Appreciation

Authors: SUN Lili[1], WANG Weijie[2], SHENG Jiefei[3]
[1]Institute of Information Management and Technology, Nanjing University of Technology, Nanjing 210009
[2]Nanjing University of Communication, Nanjing 211172
[3]School of Economics and Management, Nanjing University of Technology, Nanjing 211816

Abstract:
[Purpose/Significance] This study explores the influencing factors of scientific data value appreciation and reveals the inherent developmental laws governing this process. [Method/Process] Based on expert interview data and literature related to scientific data value appreciation, we employed grounded theory to develop a theoretical model of influencing factors. Building upon this foundation, we established a system dynamics model to dynamically simulate and analyze the value appreciation process, thereby elucidating the relationships between various factors and scientific data value appreciation. [Results/Conclusions] The findings indicate that raw data quality serves as a prerequisite for value appreciation, data storage and deposit levels exert significant influence, data organization and integration constitute the key to value formation, and scientific data sharing and development represent the critical pathway for value realization.

Keywords: scientific data; value appreciation; influencing factors; system dynamics simulation; data elements; FAIR

1 Introduction

In the context of data-intensive research paradigms, scientific data has become a crucial foundation for driving scientific progress and knowledge innovation, gradually emerging as a new strategic high ground for national scientific and technological competition. To enhance scientific data openness, numerous countries and regions have formulated various policies, regulations, and guidelines. For instance, the U.S. Office of Science and Technology Policy issued the "Memorandum on Promoting Access to Scientific Research Results,"while the European Union released the "Directive on Open Data and Reuse of Public Sector Information." These policies mandate that scientific data generated from public funding should follow the principle of "open by default, closed by exception." The "Opinions on Improving the Mechanism and System for Market-oriented Allocation of Data Elements" issued by the Central Committee of the Communist Party of China and the State Council further explicitly proposed accelerating the cultivation of data element markets and enhancing the value of data resources, providing direction for mining the value of scientific data and promoting the development of open science movements.

To advance standardized management of scientific data and thereby improve its application value, the international community proposed the FAIR (Findable, Accessible, Interoperable, and Reusable) principles draft at the academic conference held in Leiden, Netherlands, in 2014. Since their introduction, the FAIR principles have gained widespread international recognition, with many countries and organizations applying them to scientific data open sharing practices to promote quality improvement. From the perspective of data organization and processing, PETR et al. argued that the discoverability and metadata quality of big data influence its value, while others proposed that big data value stems from its embedded knowledge associations, with the core lying in characterizing these associations.

In terms of infrastructure, comprehensive scientific data platforms have been established, including the Open Science Center, Australia's Research Data Sharing Infrastructure, the European Open Science Cloud, and China' s disciplinary data sharing infrastructure, which has formed 20 national scientific data centers. These platforms provide strong support for realizing scientific data value. Regarding scientific data management services, many foreign universities have established centralized Research Data Management (RDM) services to support researchers. The University of Cambridge' s Data Champions program promotes good data management practices by establishing local data champions in departments or colleges.

Current domestic and international scientific data open sharing practices primarily focus on policy formulation, platform construction, and data management services. These practices have expanded the scale of scientific data openness and enhanced its application value, effectively promoting value realization. However, overall, scientific data management commonly exhibits phenomena of "emphasizing quantity over quality" and "emphasizing openness over value appreciation," facing increasingly severe challenges in scientific data reuse that seriously affect the role of scientific data as a production element and its value realization. Scientific data value appreciation runs through the entire process of scientific data open sharing. At this juncture of developing scientific data open sharing practices, exploring the core issue of scientific data value appreciation is timely and necessary.

2 Related Research

Existing research on data value primarily addresses data value formation or realization. From an economic perspective, Li Haijian et al. argued that data quality and timeliness affect data value formation. From a technical perspective, JOAO believed that using cloud computing and other open technologies to process data can enhance open data value. Regarding the realization of data element value, Ma Feicheng et al. considered data collection as the source, data organization as an important link for value realization, data circulation as the key, and data utilization as the "last mile." Xia Yikun et al., focusing on data elements, subjects, and environment, explored influencing factors of data element value and proposed pathways for value enhancement.

Research on scientific data value has mainly focused on theoretical discussions of value standards and realization strategies. Deng Jun et al. proposed principles and standards for scientific data value appraisal. Gu Liping et al. presented a theoretical framework for enhancing scientific data value from a library business practice perspective. Sun Jianjun et al. suggested that the scientific big data value chain should be oriented toward scientific big data resources and services.

Feng Yuan, based on value co-creation theory, proposed a value co-creation model for scientific data open sharing. Ren Ying et al. constructed a scientific data value co-creation system, arguing that system interaction, environment, culture, and technology jointly influence scientific data value co-creation.

Scientific data development and value realization have long been hot topics in academia. However, most existing research explores scientific data value realization from dimensions such as value subjects and environment. Scientific data value appreciation represents an extension of value realization, involving the entire lifecycle of scientific data. This study employs grounded theory to identify influencing factors of scientific data value appreciation, comprehensively analyzes its laws, and uses system dynamics simulation to reveal the internal mechanisms by simulating the effects of influencing factors.

3 Conceptual Framework

3.1 Conceptual Connotation of Scientific Data Value Appreciation

As a key element of open science, scientific data open sharing has become an international consensus. The issue of scientific data value is fundamental to data management. Existing literature reveals that Gu Liping et al. view data value enhancement as transforming scientific data from disorder to order through standardized organization and management such as establishing knowledge associations, thereby revealing hidden information and creating greater research value. Ren Fubing et al. propose that government open data value appreciation involves further integrating data resources to form personalized value-added products or services. The former emphasizes the data organization and management process, while the latter highlights data resource development and utilization.

In the current context of scientific data open sharing, value appreciation should focus on the entire lifecycle of scientific data. Scientific data value appreciation refers to the process where relevant stakeholders, through management means and data processing technologies, enhance the value effect of scientific data or generate value-added data products and services during the entire lifecycle from generation to development and utilization, thereby stimulating the potential value of scientific data.

3.2 Path Analysis of Scientific Data Value Appreciation

Scient data value appreciation is not merely a simple superposition of various data management activities but rather involves close collaboration among stakeholders to promote resource sharing, circulation, and innovative integration, thereby achieving value appreciation. Drawing on information value chain theory, the path of scientific data value appreciation can be analyzed as follows:

In the scientific data generation stage, researchers or research teams produce raw scientific data. The quality of raw scientific data forms the foundation for subsequent management and development. In the scientific data organization and integration stage, raw scientific data or institutionally aggregated data are further submitted to scientific data centers or publishing platforms. Professionals conduct quality reviews, perform standardized processing on approved data, assign high-quality metadata, establish knowledge associations, and then make the data publicly available. This stage significantly enhances data value.

In the scientific data development stage, after scientific data are opened for sharing, more social institutions and organizations can participate in innovative integration. Data research and development institutions can use intelligent means to conduct secondary development of scientific data according to user needs, forming value-added data products or services. Data operation institutions can create personalized scientific data services and derivatives based on market demand. As scientific data flow from raw scientific data to centrally stored data, to publicly available data, and finally to value-added scientific data and services, the value form of scientific data achieves a transformation from low value to high value.

[Figure 1: see original paper] The path of adding value of scientific data

4 Identification of Influencing Factors

4.1 Research Method

This study adopts grounded theory, a rigorous qualitative research method suitable for constructing theoretical models from complex qualitative materials. Unlike quantitative research, grounded theory proceeds directly from data, gradually abstracting and inducing theory through a bottom-up process. The study collected comprehensive materials on scientific data value appreciation through literature collection and expert interviews.

4.2 Data Collection

4.2.1 Interview Data Collection To avoid potential issues in interviews, we followed the normative interview framework proposed by MYERS et al. To ensure the interview outline was unambiguous and easily understood, we conducted pilot interviews to optimize the wording and address emerging issues. The final interview outline primarily included: (1) What are the main approaches for generating/collecting research data, and what difficulties are encountered? (2) Are you willing to open and share scientific data by submitting to data centers or data journals? (3) What are the key links in the scientific data processing and organization process, and what factors affect data value appreciation? (4) What factors may influence scientific data value appreciation during the open utilization process?

Before interviews, researchers explained the purpose and theme to participants, made confidentiality commitments, and obtained consent for recording. According to participants' data roles and specific contexts, question expressions were appropriately adjusted. Interview work was concentrated in October-November

2022, selecting 13 researchers with scientific data generation and utilization experience and 5 staff members from a provincial scientific data resource coordination center. A total of over 30,000 words of interview text were collected. With participants' consent, recordings were transcribed. Thirteen interview texts were randomly selected for coding analysis, with the remaining 5 reserved for saturation testing.

4.2.2 Literature Data Collection From August to December 2022, relevant literature was retrieved from CNKI and Web of Science Core Collection. Since direct research results on scientific data value appreciation influencing factors are scarce, lacking mature models for direct adoption, this study employed grounded theory to identify influencing factors. Forty-two relevant documents were obtained, with 35 randomly selected for coding and the remainder reserved for saturation testing.

4.3 Factor Identification

Open Coding: The obtained sample materials were repeatedly analyzed to familiarize with content. Sample materials were analyzed individually to form entry data, completing initial coding. The coding team, combining research themes and through repeated discussion, merged and streamlined initial entries based on principles such as relevance to scientific data value appreciation and distinguishability between entries, ultimately forming 76 entries. Coders were divided into Groups A and B, who independently categorized and named entries. After repeated discussion, 20 categories were formed by merging concepts with identical, similar, or interconnected meanings into the same category. Due to space limitations, partial open coding processes are shown in Table 1.

Some of the open coding processes

Axial Coding: By analyzing the content of initial categories, mutually independent categories were connected to develop deeper main categories. Based on path analysis of scientific data value appreciation and combining open coding results, the 20 categories were ultimately summarized into 7 main categories.

Selective Coding: The main categories obtained from axial coding were systematically analyzed to select overarching core categories and deeply explore connections between core and other categories. Following this approach, a theoretical framework reflecting the essence of the research phenomenon was constructed. Through comparative analysis of the 6 main categories, "scientific data value appreciation" was determined as the core category.

Saturation Testing: Theoretical saturation is a subjective concept without objective measurement indicators. This study used the reserved 5 interview texts and 7 theoretical documents for concept testing. No new concepts or theoretical connotations emerged, with reserved content basically covered by the previous 20 concepts. This indicates that the interview results and theoretical formation have reached saturation.

To verify the appropriateness of the categorization, this study employed the reverse classification method. Three graduate students not involved in the research were asked to classify the 76 entries into corresponding categories after understanding the main categories. Results showed that entries unanimously classified by all three judges into expected categories accounted for the majority, with 7 entries classified consistently by two judges and 0 entries completely inconsistent. After discussion, the main reason for inconsistent classification was the questionable distinguishability between "digital intelligence technology application ability" and "data processing ability." Since the content of "digital intelligence technology application ability" was already reflected in "data development ability," this main category and corresponding categories were deleted, ultimately yielding 6 main categories and 19 categories.

Three previous interviewees were randomly selected to verify the 归属 of influencing factors they described, and all could correctly correspond. Through reverse regression and feedback verification, the grounded theory results demonstrated certain credibility. The final influencing factors of scientific data value appreciation are shown in Table 2.

Influencing factors of scientific data value appreciation

5 System Dynamics Simulation

5.1 System Boundary Determination

Clear system boundaries are crucial for system dynamics modeling. This study takes the scientific data value appreciation process as the research object, focusing on relevant influencing factors. The influencing factors are divided into system boundaries, including six subsystems: raw data quality factors, data storage and deposit factors, data organization and integration factors, data sharing and development factors, funding guarantee factors, and data policy and supervision factors. Complex dynamic causal relationships exist among factors within each subsystem, with data storage and deposit factors significantly influenced by data deposit willingness. Therefore, perceived achievement related to data deposit willingness is included in this subsystem.

5.2 Causal Feedback Relationship Analysis

This study depicts the dynamic process of scientific data value appreciation and its relationships with influencing factors through causal feedback loop diagrams. Using the Vensim PLE tool, feedback loops of the scientific data value appreciation process were obtained, as detailed in Table 3.

Feedback loop of value-added scientific data

The causal relationship diagram reveals that scientific data value appreciation is not a simple linear process but involves multiple feedback loops. These loops demonstrate that improvements in data storage, organization, sharing, and development can enhance value appreciation, which in turn positively influences funding guarantees and policy supervision, creating a virtuous cycle.

[Figure 2: see original paper] Causal diagram of the influencing factors of scientific data value appreciation

5.3 System Flow Diagram

5.3.1 System Flow Diagram Construction Based on the causal relationships of scientific data value appreciation influencing factors, this study used the Vensim PLE tool to draw the system flow diagram, as shown in Figure 3. The diagram reflects the interaction patterns among four types of variables and reveals the relationships between various elements in the scientific data value appreciation process.

[Figure 3: see original paper] Flow diagram of the value-added system of scientific data

The system flow diagram includes 6 level variables, 6 rate variables, and 36 variables total. Specific variables and their meanings are shown in Table 4.

Variable names and types

5.3.2 Basic Assumptions To facilitate simulation, basic assumptions are proposed for the system dynamics model:

1. The scientific data value appreciation influencing factor system is a closed system with frequent interaction and feedback among subsystems, without considering external factors.
2. Raw data quality, data organization, and integration factors will change with accumulated practice.
3. The constructed system dynamics model has universality, capable of quantifying scientific data value appreciation levels.
4. System collapse caused by abnormal situations or emergencies such as major natural disasters or policy changes is not considered.

5.3.3 Simulation Equations and Parameter Settings The premise for running the scientific data value appreciation system model is setting relevant constants and initial values. Drawing on methods from scholars such as Gao Xiaoning and Yuan Hong, this study uses a questionnaire survey to obtain weights for equations and some variable initial values. The questionnaire employs a 10-point Likert scale, with 1-10 representing "strongly disagree" to "strongly agree." Survey participants were the 18 interviewees selected earlier. Questionnaires were distributed via QQ email with a 7-day recovery period and 100% recovery rate.

By organizing scores from each questionnaire and using the Analytic Hierarchy Process software YAAHP to calculate results, factor weights and overall consistency test results were obtained. The arithmetic mean method was used to process the 18 experts' scores, with consistency test indicator CR < 0.1, meeting research requirements. Final weights for influencing factors in each subsystem were obtained.

Main variable equations in the model include:

Scientific data value appreciation level = Raw data quality level $\times$ 0.25 + Data storage and deposit level $\times$ 0.13 + Data organization and integration level $\times$ 0.2 + Data sharing and development level $\times$ 0.14 + Funding guarantee level $\times$ 0.16 + Data policy and supervision level $\times$ 0.12

Raw data quality level = INTEG (Raw data quality level change amount) Raw data quality level change amount = SMOOTHI (0.2 + Researcher data literacy $\times$ 0.45 + Data software and hardware facilities $\times$ 0.3 + Institutional data service quality $\times$ 0.05)

Data storage and deposit level = INTEG (Data storage and deposit level change amount) Data storage and deposit level change amount = SMOOTHI (0.3 + Data deposit willingness $\times$ 0.25 + Data storage and maintenance mechanism $\times$ 0.2 + Data review mechanism $\times$ 0.1 + Funding guarantee level $\times$ 0.05 + Scientific data value appreciation level $\times$ 0.1)

5.4 Model Simulation and Analysis

5.4.1 Model Validity Testing To ensure model validity and smooth dynamic simulation, this study conducted validity testing on the scientific data value appreciation system dynamics model. The initial state of scientific data value appreciation level change was selected as the observation indicator. Using Vensim PLE for simulation, the time period was set to 12 months with a step size of 1 month, as scientific data value appreciation requires continuous promotion under various factors rather than being achieved overnight.

The change trend of scientific data value appreciation level is shown in Figure 4. The level maintained low growth in the early stage, but from the 8th month, the trend shifted from gradual to rapid increase. This means that if optimization is implemented across the six dimensions of raw data quality, data storage and deposit, data organization and integration, data sharing and development, data policy and supervision, and funding guarantee, the optimization effects will become apparent by the 8th month. This aligns with the actual operation of renowned scientific data centers such as China' s National Scientific Data Center and the European Open Science Data Center, where value appreciation mainly manifests in data scale and application value, and these activities require time to optimize.

[Figure 4: see original paper] Effectiveness analysis of the system dynamics model

The six influencing factors all positively affect scientific data value levels, with raw data quality having the greatest impact, followed by data sharing and development, and data storage and deposit. Data policy and supervision, data organization and integration, and funding guarantee also play positive roles as auxiliary support factors. The time lag before significant effects appear corresponds to the initially low value appreciation level. As process systems mature and policy frameworks improve, scientific data value appreciation efficiency becomes significant, creating a virtuous cycle.

5.4.2 Scenario Analysis This study conducted sensitivity analysis by changing six main parameters: raw data quality, data storage and deposit, data organization and integration, data sharing and development, data policy and supervision, and funding guarantee. Each parameter was set to increase by 20% while keeping other variables constant, simulating the change trend of scientific data value appreciation level under identical variation amounts to compare influence degrees.

The effect of each influencing factor on scientific data value level is shown in Figure 5. Raw data quality shows the most significant impact, with its influence proportion gradually decreasing as simulation time progresses.

[Figure 5: see original paper] The effect of each influencing factor on the value level of scientific data

Raw Data Quality Factors: In the raw data quality subsystem, researcher data literacy has the most significant impact, followed by data software and hardware facilities and institutional data service quality. As researcher data literacy improves and data sharing willingness increases, raw data quality is significantly enhanced. Figure 6 shows that enhancing researcher data literacy, improving institutional data service quality, and upgrading data software and hardware facilities can substantially improve raw data quality.

[Figure 6: see original paper] Trends of key factors in the raw data quality subsystem

Data Storage and Deposit Factors: In the data storage and deposit subsystem, data deposit willingness has the most obvious impact, followed by data review mechanisms and data storage and maintenance mechanisms. However, since enhancing researcher data deposit willingness and implementing storage mechanisms require an accumulation process, data storage level does not change significantly in the early stage. From the 8th month, as researcher willingness and related mechanisms mature, data storage and deposit level growth rate continuously increases (Figure 7).

[Figure 7: see original paper] Trends of key factors in the data storage subsystem

Data Organization and Integration Factors: In the data organization and integration subsystem, metadata quality demonstrates more significant positive promotion effects than data processing capacity and data integration capacity. As simulation time progresses, with cumulative enhancement of metadata quality and processing capabilities, scientific data organization and integration level grows rapidly from the 7th month (Figure 8).

[Figure 8: see original paper] Influencing trend of key factors in data sharing and development subsystem

Data Sharing and Development Factors: In the data sharing and development subsystem, data openness scale, data sharing platform quality, data development capacity, and data sharing service capacity all positively affect value appreciation, with data openness scale having the greatest impact. Data openness risk exerts reverse inhibition—higher risk leads to lower sharing and development levels. Figure 9 shows that factors are not significant in the first 8 months but become gradually apparent afterward, as these influences require accumulation time.

[Figure 9: see original paper] Trends of key factors in the data sharing and development subsystem

Data Policy and Supervision Factors: In the policy and supervision subsystem, key factors ranked by influence degree are data policy perfection, data evaluation and incentive mechanisms, and data supervision mechanisms. Data policy perfection has the most significant impact on data policy and supervision levels (Figure 10).

[Figure 10: see original paper] Influencing trend of key factors in the policy and regulatory subsystem

Funding Guarantee Factors: In the funding guarantee subsystem, fiscal support has faster effects, while data value appreciation revenue growth requires development accumulation. However, data value appreciation revenue is more sustainable (Figure 11).

[Figure 11: see original paper] The influencing trend of key factors in the fund guarantee subsystem

6 Conclusions and Discussion

After collecting comprehensive materials on scientific data value appreciation through literature review and expert interviews, this study employed grounded theory to inductively derive a theoretical model of influencing factors. System dynamics simulation was then conducted, yielding the following conclusions:

1. **Raw data quality is the prerequisite for scientific data value appreciation.** In the raw data quality subsystem, researcher data literacy has the most significant impact. Enhancing researchers' data collection, processing, analysis capabilities, and data ethics awareness can improve raw data quality.

2. **Data storage and deposit level significantly influences scientific data value appreciation.** When researchers' perceived effort and perceived risk of data deposit decrease, accompanied by scientific data sharing

policy pressure, deposit willingness increases, significantly improving the scale effect of data storage and deposit.

3. **Scientific data organization and integration is key to value formation.** Metadata quality has the most significant impact on organization and integration levels. High-quality metadata facilitates scientific data organization and integration.
4. **Scientific data sharing and development is key to value realization.** As the final step in value appreciation, data openness scale, sharing platform quality, and development capacity all positively promote sharing and development levels.

This study comprehensively and dynamically analyzed the scientific data value appreciation process using system dynamics, considering interactions among related factors and avoiding the limitation of empirical data measurement in existing research. The findings provide clear optimization targets for open scientific data governance practices. Although this study collected rich qualitative data through expert surveys and literature review, scientific data practices evolve rapidly with hardware and software environments. Future research should continuously track influencing factors using diversified methods to further validate and refine the conclusions.

7 Implications

Based on simulation results, this study proposes the following strategic implications for scientific data value appreciation:

Cultivate Data Literacy and Improve Data Environment: Researcher data literacy significantly affects raw scientific data quality. Data literacy education should cover the entire scientific data lifecycle, from generation to organization and utilization. Universities and libraries should serve as main venues for data literacy education, offering general and discipline-specific data literacy education through various convenient and efficient forms such as online consultations and free tools for metadata editing.

Improve Data Infrastructure: Scientific data infrastructure is crucial for value appreciation. Efforts should focus on improving scientific data platforms, enriching data processing functions, and following FAIR principles to ensure standardization, interoperability, and scalability.

Focus on Metadata Quality and Data Integration: Metadata quality is the foundation for scientific data organization and integration. Scientific data with high-quality metadata is more easily organized and utilized. Current metadata standards are often limited to specific research stages and lack universal standards. There is a need to plan scientific data metadata standards reasonably based on core metadata and disciplinary characteristics to meet open sharing and cross-domain interoperability requirements.

Expand Open Scale and Optimize Development Capacity: Scientific data management institutions should expand data openness scale, making more data discoverable and accessible. On this basis, they should optimize scientific data development services, build integrated platforms with openness and intelligence functions, and actively explore and cultivate mutually beneficial symbiotic models among stakeholders.

Explore Market-oriented Models: Under the premise of ensuring national data security and intellectual property rights, scientific data products and value-added services should be explored using emerging digital technologies such as cloud computing and artificial intelligence. Market-oriented models should be developed to align with the new trend of scientific data participating in market distribution as a production element. Regional pilot programs, such as in the Yangtze River Delta region, could integrate scientific data resources, build integrated platforms, and attract institutions and the public to participate in data product development through supporting incentive policies.

References

[1] HOLDREN J P. Increasing access to the results of federally funded scientific research [EB/OL]. [2023-07-06]. https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf.

[2] The European Parliament and the Council of the European Union. Directive (EU) 2019/1024 of the European parliament and of the council of 20 June 2019 on open data and the re-use of public sector information [EB/OL]. [2023-07-06]. https://www.legislation.gov.uk/eudr/2019/1024/introduction.

[3] Data FAIRport. Find, access, interoperate & re-use data [EB/OL]. [2023-12-06]. https://www.datafairport.org/.

[4] WEN L M, LI Y, GUO L. The practice progress and future exploration of open science at home and abroad[J]. Library and information service, 2021, 65(24): 109-122.

[5] University of Cambridge. Data champions[EB/OL]. [2021-04-03]. https://www.data.cam.Ac.uk/intro-data-champions.

[6] SONG S S, BAI W L. A review of big data governance research in China[J]. Journal of library and information science in agriculture, 2022, 34(4): 4-17.

[7] LIU L, SI L. Scientific data governance practices: Content systems and development trend[J/OL]. Information studies: Theory & application, 2023: 1-9. (2023-07-24). https://kns.cnki.net/kcms/detail/11.1762.G3.20230724.1603.002.html.

[8] LI H J, ZHAO L. Research of data valuation[J]. Finance & trade economics, 2023, 44(6): 5-20.

[9] ALBINO J P. An approach to value creation in open data for small and medium-sized enterprises using "R" ecosystem[C]// 2017 Computing Conference.

Piscataway, New Jersey: IEEE, 2018: 346-351.

[10] KŘEMEN P, NEČASKÝ M. Improving discoverability of open government data with rich metadata descriptions using semantic government vocabulary[J]. Web semantics: Science, services and agents on the world wide web, 2019, 55(C): 1-20.

[11] LI X H, FAN M H. Knowledge interaction in big data[J]. Information studies: Theory & application, 2019, 42(2): 68-73, 107.

[12] MA F C, WU Y S, LU H Z. Research on the path to realize the value of data elements[J]. Journal of information resources management, 2023, 13(2): 4-11.

[13] YANG Q Y, XIA Y K. The formation mechanism, influencing factors and implementation path of data factor value: An analysis based on grounded theory[J]. Library & information, 2023(2): 12-22.

[14] DENG J, JIA X Q, MA X J, et al. Study in appraisal standard of scientific data value[J]. Information science, 2013, 31(9): 37-41.

[15] GU L P, CHEN X L, ZHANG X Y, et al. A strategy for enhancing the value of open research data[J]. Library tribune, 2020, 40(9): 115-124.

[16] SUN J J, LI Y. Scientific big data: Paradigm remodeling and value realization[J]. Library & information, 2017(5): 20-26.

[17] FENG Y. Value co-creation model and operation mechanisms of open sharing of scientific data[J]. Library, 2022(9): 29-37.

[18] REN Y, LI N. Construction and simulation analysis of scientific data value co-creation system[J]. Digital library forum, 2023, 19(5): 42-53.

[19] REN F B, SUN M L. Research on the value-added process and mechanism of government open data based on value chain theory[J]. Information and documentation services, 2021, 42(4): 56-63.

[20] DONG C, HOU M. A three-layered model that sharing maximizes the value of scientific and technical resources (VAA)[J]. Science and technology management research, 2013, 33(11): 231-234.

[21] GLASER B G. The grounded theory perspective: conceptualisation contrasted with description[M]. Mill Valley: Sociology Press, 2001: 145.

[22] MYERS M D, NEWMAN M. The qualitative interview in IS research: Examining the craft[J]. Information and organization, 2007, 17(1): 2-26.

[23] GAO X N, HU W, ZANG G Q. System dynamics simulation analysis on factors affecting sharing efficiency of scientific research data[J]. Information studies: Theory & application, 2022, 45(8): 146-153, 103.

[24] YUAN H, WANG T. System dynamics analysis of the sustainable development path of the government open data ecosystem[J]. Library and information service, 2021, 65(17): 13-25.

[25] LI X. Social system dynamics: Principles, methods and applications of policy research[M]. Shanghai: Fudan Press, 2009.

[26] JIANG C H, LIU Z D, ZHAO Y S. Simulation study on safety experience based on system dynamics[J]. Safety and environmental engineering, 2012, 19(4): 97-101.

[27] GAO X N, et al. Research on influencing factors of scientific research data sharing efficiency based on system dynamics[J]. Information studies: Theory & application, 2022.

[28] YUAN H, et al. System dynamics analysis of sustainable development path of government open data ecosystem[J]. Library and information service, 2021.

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