
AI translation · View original & related papers at
chinaxiv.org/items/chinaxiv-202308.00396

Constructing a Data Capability Model for Knowledge Innovation-Oriented Research Teams (Postprint)

Authors: Du Xingye, Li He, Li Zhuozhuo

Date: 2023-08-26T00:00:00+00:00

Abstract

[Purpose/Significance] This study examines the core elements that influence knowledge innovation in research teams within data-intensive research environments, proposes a data capability model for research teams to facilitate knowledge innovation, and identifies the constituent factors of these capabilities. [Method/Process] Through literature review, this research synthesizes the research dimensions of knowledge innovation capabilities of research teams in traditional environments, analyzes the research environment shaped by the rapid advancement of science in data-intensive contexts, proposes six dimensions of data capabilities for research teams and their attendant factors in the new environment, and constructs a data capability model for research teams aimed at promoting team knowledge innovation. [Results/Conclusion] In data-intensive environments, the data capabilities of research teams encompass: data awareness within the research team; data retrieval and discovery capabilities founded on such awareness; the ability to organize, integrate, and manage data; the ability to describe and preserve data; the ability to analyze and utilize data and to discover knowledge and patterns therefrom; and data discrimination and evaluation capabilities. These capabilities foster knowledge innovation in research teams through knowledge discovery, knowledge management, knowledge integration, knowledge sharing, knowledge evaluation, and related dimensions.

Full Text

Preamble

Vol. 62 No. 4 February 2018
ChinaXiv Cooperative Journal

Research on Data Capability Model Construction for Knowledge Innovation in Scientific Research Teams

Du Xingye^{1, 2}, Li He¹, Li Zhuozhuo³

¹ School of Management, Jilin University, Changchun 130022

² National Science Library, Chinese Academy of Sciences, Beijing 100190

³ Department of Archives and E-Government, School of Social Science, Soochow University, Suzhou 215123

Abstract

[Purpose/Significance] This study examines the core elements influencing knowledge innovation in research teams within data-intensive scientific environments and proposes a data capability model to promote knowledge innovation in research teams, identifying the capability structures within these factors. **[Method/Process]** Through literature review, this paper summarizes research dimensions of knowledge innovation capabilities in research teams under traditional environments, analyzes the scientific research environment formed by rapid development in data-intensive contexts, proposes six dimensions of research team data capabilities and their constituent factors in the new environment, and constructs a data capability model aimed at promoting team knowledge innovation. **[Result/Conclusion]** In data-intensive research environments, research team data capabilities include: data awareness, data retrieval and discovery capabilities based on that awareness, abilities to organize, integrate, and manage data, capabilities to describe and preserve data, abilities to analyze and utilize data to discover knowledge and patterns, and data discrimination and evaluation capabilities. These capabilities promote knowledge innovation in research teams through knowledge discovery, knowledge management, knowledge integration, knowledge sharing, and knowledge evaluation.

With the rapid development of information technology, the knowledge foundation of scientific research has undergone significant changes. Scientific research across multiple fields is acquiring massive amounts of raw scientific data from miniaturized, embedded, intelligent sensor devices and their full integration with the Internet. Various types of observational data, sensory data, experimental data, computational data, simulation data, and modeling data are growing exponentially and have become important data resources that the scientific research and discovery community increasingly values and depends upon. The utilization of these massive, heterogeneous scientific datasets for knowledge discovery constitutes the data-intensive scientific paradigm [1]. Under this new paradigm, researchers need a capability different from traditional experimental, organizational, and communication skills—data capability, defined as the ability to organize, aggregate, analyze, discover, and utilize massive, dynamic scientific data to conduct knowledge exploration and discovery in a data-driven manner [2]. How to extract information and discover knowledge from intensive scientific data to achieve knowledge innovation is a new challenge facing researchers and research teams. Wang Xiaowen et al. point out that facing the large volumes of data continuously generated and created during scientific research, researchers must possess data organization, management, and analysis skills to truly con-

duct research [3].

Research teams and individual researchers are distinct concepts. A research team is a scientific organization composed of two or more researchers working toward common research goals. Individual researchers' data capabilities complement and strengthen the team's overall data capability. Building a data capability framework for researchers can enhance data capabilities, enabling research teams to rapidly adapt to new research environments and conditions. Current research shows that in data-intensive research environments, data has become new infrastructure and a necessary tool for scientific research. Studies on individual researchers' data capabilities can be applied to research teams. This paper aims to establish a data capability framework for research teams to enhance overall knowledge innovation capacity. It first analyzes the dimensions of knowledge innovation capabilities in research teams under traditional environments (from a research perspective), then examines characteristics of knowledge innovation in research teams in data-intensive environments, constructs a knowledge innovation capability model for research teams in data-intensive contexts from a data capability perspective, and provides reference for knowledge innovation activities in the new environment. Specific research questions addressed are: What are the characteristics of research on knowledge innovation capabilities in research teams under traditional environments? What constitutes research team data capability? How can a core capability model for knowledge innovation in research teams be constructed based on researcher data capabilities?

2. Related Research

2.1 Main Dimensions of Knowledge Innovation Capabilities in Research Teams Under Traditional Environments

Knowledge innovation refers to the process of acquiring basic scientific and technological knowledge through scientific research, aiming to pursue new discoveries, explore new patterns, establish new theories, create new methods, and accumulate new knowledge [4]. The essence of knowledge innovation can be summarized into two aspects: dimensions of knowledge innovation and methods of knowledge innovation. Knowledge innovation promotes the generation, flow, and dissemination of new ideas and accelerates their application in products and services for societal benefit [5]. The most widely recognized model internationally is I. Nonaka's SECI spiral model of knowledge creation, which posits that knowledge creation involves the revolutionary emergence of new knowledge through the spiral transformation of explicit and tacit knowledge [6].

Research team knowledge innovation capability refers to a team's ability to discover, accumulate, absorb, understand, master, and apply knowledge to ultimately produce new knowledge [7]. Numerous scholars have studied this topic, primarily from dimensions including team learning capability, technological factors, internal knowledge integration, and knowledge potential theory, as shown

in .

TABLE:1 Main Dimensions and Perspectives in Research on Research Team Innovation Capabilities

Dimension	Representative Studies	Main Perspective
Team Learning Capability	A. Khedhaouria et al. [8]	Analyzes factors influencing research team knowledge innovation from a knowledge acquisition perspective, emphasizing the key role of learning orientation
Team Technology Capability	J. Hong et al. [9]	Examines mixed effects of social (learning culture, team quality, knowledge complexity) and technical (IT support) factors on team knowledge creation and performance from a socio-technical perspective
Team Knowledge Integration	Zhao Limei et al. [10]; Liu Yanfang et al. [11]	Discusses knowledge integration as a critical link to achieving knowledge innovation [10]; Views knowledge integration as an important component of knowledge creativity, with team innovation depending on internal knowledge integration [11]

Dimension	Representative Studies	Main Perspective
Team Knowledge Transfer	Guo Yanli et al. [12]; Yang Bin et al. [4]; Li Gang et al. [13]; Wang Xinchun et al. [14]	Constructs task-based team knowledge innovation models based on knowledge potential theory [12]; Argues that improving knowledge transfer efficiency is key to enhancing innovation capability [4]; Suggests reducing knowledge stickiness to improve transfer performance [13]; Develops a SECI-based knowledge creation capability framework including sensing, exploration, absorption, and adaptation [14]
Team Coordination Mechanisms & Communication	Wu Yang et al. [15-16]	Introduces coordination mechanisms [15] and communication management [16] into research team knowledge innovation systems to achieve more effective organizational knowledge innovation

Reviewing these studies reveals that previous research on team knowledge innovation capabilities has focused primarily on knowledge-level integration, absorption, transformation, and team collaboration abilities. Few studies have examined capabilities from the perspective of research environments and team data, yet with scientific development and the accumulation of scientific data in network environments, the research environment now places greater emphasis on data's role in knowledge innovation. Understanding, organizing, mining, integrating, and utilizing data have become increasingly important and constitute one of research teams' key capabilities. Therefore, data capability dimensions cannot be omitted from studies of research team knowledge innovation capabilities.

2.2 New Requirements of Data-Intensive Research Environments for Scientific Research

The term “data-intensive” first originated from “data-intensive computing” [17]. The book *The Fourth Paradigm: Data-Intensive Scientific Discovery* notes that scientific research paradigms have evolved from empirical science, theoretical science, and computational science to data science [18]. The fourth paradigm of scientific research addresses data-intensive science, representing a shift from traditional hypothesis-driven to data exploration-based scientific methods, with the goal and task of advancing current technological frontiers in managing, analyzing, and discovering knowledge from large, high-velocity data [19].

In data-intensive research environments, the collection, analysis, storage, sharing, and visualization of scientific datasets have become central themes of research activities. Researchers face important questions about how to obtain infrastructure support suitable for data-intensive science, conveniently access and manage scientific datasets, and establish new scientific communication systems [20]. Liang Na et al. analyzed the new inspirations that data-intensive scientific discovery brings to research activities, proposing challenges researchers face regarding scientific data management, interoperability, data rights management, and data literacy in the new paradigm. They note that in the current “big data” environment, data has transformed from mere research results to the foundation and tools of scientific research. Researchers focus not only on data organization, description, preservation, access, analysis, and utilization, but also on how to use interactive, open networks to objectify and computationally enable massive data for knowledge discovery and collaborative research [20].

Data-intensive research environments clearly impose high demands on researchers’ and teams’ data capabilities. Researcher or team data capability refers to data awareness and the abilities to retrieve and discover data, organize, integrate, and manage data, describe and preserve data, and analyze and utilize data to discover knowledge and patterns from it. These capabilities are undoubtedly related to other knowledge innovation abilities (such as knowledge integration and learning capabilities) but differ in their focus on data itself and its organizational management and application, particularly in utilizing massive data in open network environments for knowledge discovery [21].

3. Research Team Data Capabilities

3.1 Research Team Data Awareness

Data awareness is the first capability research teams should possess in data-intensive research environments. It refers to the recognition of team data needs by every team member and the team as a whole, specifically manifested as: the team and its members fully recognizing data’s importance in research work, being able to use data acquisition and utilization to solve problems, having keen perception and insight into data in relevant fields, efficiently and quickly identifying valuable data, tracking and organizing data long-term, and being adept at

finding solutions or approaches from acquired data. Data awareness determines the degree of consciousness in acquiring, evaluating, and utilizing data. Within research teams, the strength of data awareness determines whether researchers can effectively process and utilize data, whether they can mine useful data information in a timely manner, and whether they can efficiently disseminate and share data outcomes, directly affecting the processes and effectiveness of data discovery, management, organization, and sharing.

The main factors and capability structures for evaluating research team data awareness are shown in .

TABLE:2 Factors in the Data Awareness Dimension

Factor	Capability Structure Level
Understanding of data's importance in research	Very familiar and knows how to obtain
Awareness of datasets in relevant fields and data IP	Familiar with some and knows how to obtain
Habit of collecting and organizing valuable data	Aware but has not attempted to obtain
Experience in data analysis	Has habit, intermittent
Philosophy of data sharing	Has habit, uses tools proficiently

3.2 Research Team Data Retrieval, Discovery, and Production Capabilities

Retrieving and discovering data is a crucial part of research work. For researchers, using appropriate search terms and techniques to obtain needed data from vast databases or datasets and discovering useful data is the starting point of every research project. When researchers identify data needed for their work, they can retrieve existing research data or produce relevant data through research methods or experiments. This requires researchers to possess certain data collection, retrieval, and production capabilities. By cultivating these capabilities, research teams can establish standardized knowledge acquisition abilities.

Researchers must understand different data sources and apply corresponding data knowledge to select the most relevant data from massive datasets to solve team research problems. They need to master basic data acquisition methods,

be adept at obtaining various types of public data, and be able to discover data from non-public sources that can solve specific problems through legitimate methods. Additionally, researchers must possess the ability to produce data through research methods, including using various forms of survey methods and tools, and be able to produce derived data through deep mining. Specific capabilities are shown in .

TABLE:3 Capabilities in Data Retrieval, Discovery, and Production Dimensions

Capability	Capability Structure Level
Traditional data acquisition and discovery	Familiar and continuous
Note-taking, real-time audio recording, data preservation	Familiar but intermittent
Survey data organization	Possesses and is familiar
Data coding	Possesses and implements
Online dataset retrieval and discovery	Possesses and implements
Data source identification, cross-disciplinary source identification	
Data retrieval techniques and skills	
Data retrieval methods	
Data discovery methods	
Data retrieval tool usage	
Data discovery techniques and skills	
Experimental data collection and organization	
Survey data collection and organization	
Data discovery software and tool usage	

Capability	Capability Structure Level
Data production capability	

3.3 Research Team Data Management Capability

Data requires management and has its own lifecycle. Organizations abroad have proposed several data lifecycle models describing data's entire lifecycle from generation, collection, description, storage, discovery, analysis to reuse, such as the Digital Curation Centre (DCC) Curation Lifecycle Model [22] and the UK Data Archive Data Lifecycle Model [23]. The Royal Society's 2012 report *Science as an Open Enterprise* noted the need for more experts to manage and support the use of digital data [24], and its 2014 *Science 2.0 Consultation* emphasized the importance of data management standards [25].

What is data? T. Koltay argues that data can be any computer file used to organize data composed of text, numbers, dates, links, etc., manageable by computer programs, transmittable via computers and networks, and enabling people to understand data's meaning [26].

Data management or curation for scientific research refers to researchers' ability to standardize collection, classification, storage, utilization, updating, and management of data throughout its lifecycle. Research team data management capability refers to the degree to which teams implement research data management processes or standards, as shown in .

TABLE:4 Research Team Data Management Capability Dimensions

Data Management Capability	Capability Structure Level
Research team data policy or management standards/plan	
Data management system	
Data format processing and classification standards	
Data storage space allocation and management	
Data copyright management	
Data utilization records	
Data addition, update, preservation, deletion records	
Data sharing policy	
Data storage in public databases or platforms	
Research team data responsibility system	
Whether team has data specialists	
Whether team has data experts	
Research team data sharing policy	
Whether public data is shared within team	
Whether private data is shared among members	

3.4 Research Team Data Description and Preservation Capability

Data comes in categories, including qualitative and quantitative data, and in different stages requiring different description and representation methods. Metadata is primarily used for data description and preservation. For research teams, processing data from different sources, types, and forms using appropriate methods is a major challenge. Researchers need to clearly classify and organize team-owned data. [Figure 1: see original paper] shows a framework for data description.

The demand dimensions for research team data description and preservation capabilities are shown in .

TABLE:5 Research Team Data Description and Preservation Capability Dimensions

Data Description and Preservation Capability	Capability Structure Level
Understanding of metadata description standards in the field	
Ability to properly open, edit, preserve, and organize multimedia data	
Data stage and dimension management	
Whether data at different stages and dimensions is organized and managed	
Data management tools	
Understanding of data management tools	

3.5 Research Team Data Analysis and Knowledge Discovery Capability

After collecting, describing, organizing, and managing data, researchers must use appropriate data analysis methods and tools based on data attributes and research purposes to derive research conclusions. In data-intensive research environments, the dramatic increase in data volume and complexity of data problems require researchers to employ professional data analysis tools. Data operation and analysis are crucial processes for mining data content to discover its value. Researchers' data understanding, analysis tools, methods, and understanding of data relationships all affect data operation and analysis capabilities and ultimately influence analysis results. Through data operation and analysis, researchers can interpret data distribution meanings, explain trend causes, understand measurement errors through sample-whole relationships, estimate overall data representation possibilities, and draw reliable conclusions. The dimensions of research team data analysis and knowledge discovery capabilities are shown in .

TABLE:6 Research Team Data Analysis and Knowledge Discovery Capability Dimensions

Data Analysis and Knowledge Discovery Capability	Capability Structure Level
Database usage and management capability	Possesses
Chart software usage capability	Basically possesses
Prediction capability for missing data	Barely possesses
Data synthesis and comparison capability	
Data correlation and analysis capability	
Data fusion and reuse capability	
Visualization tools usage and analysis capability	
Understanding of visualization tools	
Analysis of visualization results	

3.6 Research Team Data Discrimination and Evaluation Capability

Data discrimination and evaluation capability refers to researchers' and teams' ability to critically understand data meaning and value, correctly assess the feasibility of data acquisition and analysis methods, and identify potential data errors and omissions. In data-intensive research environments, facing massive data and increasingly complex data operation and analysis methods, researchers must maintain a discriminating attitude toward analysis results, critically evaluate data, and reasonably question data based on collection and production contexts to recognize result limitations and develop the thinking habit and ability to judge data value in specific contexts. The factors to consider for research team data discrimination and evaluation capabilities are shown in .

TABLE:7 Research Team Data Discrimination and Evaluation Capability Dimensions

Data Discrimination and Evaluation Capability	Capability Structure Level
Data identification capability	Possesses
Dirty data identification capability	Basically possesses
Data cleaning strategy formulation capability	Possesses
Data discrimination and evaluation standards	Basically possesses
Understanding of data limitations	
Data discrimination capability	
Data error identification capability	
Data evaluation capability	
Critical data evaluation ability	

4. Data Capabilities for Knowledge Innovation in Research Teams

Research team knowledge innovation capability is essentially a collective ability based on individual researcher capabilities, organized and coordinated through

team collaboration mechanisms. In data-intensive research environments, research team knowledge innovation capability heavily depends on data capabilities. Constructing a data capability model based on knowledge innovation is crucial for analyzing and enhancing team knowledge innovation capabilities. The data capability model addresses how research teams acquire, utilize, organize, and analyze data, using massive scientific data as research resources to discover knowledge and conduct knowledge innovation activities.

[Figure 2: see original paper] presents the research team data capability model constructed based on the above analysis. This model takes researcher data capability as its core and promoting team knowledge innovation capability as its extension, focusing on how researcher data capabilities form and transform into team knowledge innovation capability. The model has two characteristics: first, it uses data capability as the foundation for team knowledge innovation capability; second, it has unlimited boundaries—that is, knowledge innovation capability boundaries. The dashed lines in the model represent differences between capabilities, while other factors affecting knowledge innovation capability are not considered.

4.1 Data Awareness Promotes Knowledge Innovation Awareness Formation

Data awareness is an expression of innovation awareness and its starting point. Research team data awareness can promote the formation of knowledge innovation awareness. Knowledge innovation awareness in the new research paradigm includes keen sensitivity to data, persistent tracking ability, and insight and judgment of data value. Data awareness determines how researchers acquire data, and its strength determines whether they can effectively process and utilize data, whether they can mine useful data information promptly, and whether they can efficiently disseminate and share data outcomes. This directly affects the processes and effectiveness of data management, thereby influencing knowledge production and innovation outcomes.

4.2 Data Retrieval and Discovery Capabilities Determine Knowledge Acquisition Ability

Data retrieval and discovery capabilities are the foundation of team knowledge innovation. When researchers identify needed data, they can retrieve existing research data through Internet search engines, modern networks, and communication technologies, or produce relevant data through questionnaires, experiments, qualitative and quantitative methods. This requires certain data retrieval and discovery capabilities. With rapid development of information sensing technology, miniaturized, embedded, and intelligent sensor devices linked with the Internet enable acquisition of more precise and larger-scale sensory data in shorter timeframes. Issues such as whether data can be objectively and accurately retrieved, whether attributes like retrieval techniques, methods, and environments are systematically described, whether distortion occurs during sys-

tem processing, and whether traceability is possible have become unavoidable challenges. Cultivating researchers' data retrieval and discovery capabilities enables teams to establish standardized knowledge acquisition abilities, providing solid knowledge reserves and reliable sources through digitalization and cloud storage, thus laying the foundation for the quantitative-to-qualitative transformation of knowledge innovation.

4.3 Data Description and Storage Capabilities Reflect Team Knowledge Sharing Capability

Data description involves cognition and capabilities in analyzing data characteristics, media, representation forms, and structures. Accurately describing structured and semi-structured data types is fundamental to data analysis. As data volumes explode, storage has shifted from centralized to distributed systems. To improve storage efficiency, researchers can use data access and cloud storage to build data warehouses, integrating massive data to serve team operations and promote knowledge sharing, usage, and innovation. Data storage is a long-term, continuously improving process.

At the individual level, accurately analyzing results based on data description and appropriately representing trends and variations helps develop strong data analysis skills. Good data analysis and storage capabilities enable researchers to obtain better-matched data content or forms during collection. At the team level, robust data description and storage capabilities enable effective internal data resource sharing, improving knowledge sharing efficiency. In data-intensive environments, knowledge innovation depends on processing and analyzing massive data. Only through effective expression and communication of data analysis results can teams ensure effective sharing of intellectual achievements, reduce knowledge sharing costs, and ultimately enhance knowledge innovation capabilities.

4.4 Data Management Enables Better Knowledge Integration and Promotes Knowledge Sharing

Data management refers to researchers' ability to effectively collect, store, process, and apply data, standardize classification and organization, update data, and establish management standards—most importantly, the ability to use metadata. In data-intensive environments, while massive research data is generated, data format diversification is also noteworthy. Even similar data in the same field may have different formats due to various reasons, yet academia lacks detailed metadata for scientific data, especially regarding collection, processing, and transformation processes [1]. Metadata is the foundation for standardized data management and an important component of data management plans, used to better describe data, reveal content characteristics, identify and evaluate data, track changes during usage, and enable efficient massive data management for effective discovery, retrieval, organization, and management.

For research teams, standardized data management not only helps organize internal knowledge, discard useless knowledge, and organically integrate individual researchers' knowledge into more systematic and coherent forms, but also enables teams to absorb external knowledge during data exchange, integrate it into existing knowledge systems, thereby improving knowledge integration capabilities and promoting knowledge innovation.

4.5 Data Analysis and Utilization Promote Knowledge Discovery

Data analysis and utilization are crucial processes for mining data content to discover value. Researchers' data understanding, analysis tools, methods, and understanding of data relationships all affect data analysis capabilities and final discovery results. In big data environments, research teams can use machine algorithms, cloud computing, batch processing, and query analysis technologies to accurately discover needed knowledge and improve analysis efficiency.

In data-intensive environments, research teams can capture, process, and manage various types of data, mine scientific phenomena and patterns from massive datasets, and support interactive collaborative research on these phenomena to achieve knowledge discovery. Improving knowledge discovery capabilities helps teams find embedded knowledge in data, greatly benefiting knowledge innovation capability enhancement.

4.6 Data Discrimination Capability Facilitates Knowledge Evaluation Capability Formation

Data discrimination capability refers to researchers' ability to critically understand data meaning and value, correctly judge the feasibility of data acquisition and analysis methods, and identify potential errors and omissions. In data-intensive environments, facing massive data and complex analysis methods, researchers must maintain discriminating attitudes toward analysis results, critically evaluate data, and reasonably question data based on collection and production contexts to recognize result limitations and develop the ability to judge data value in specific situations.

In data-intensive research environments, improving scientific research and decision-making capabilities based on sufficient, reliable data and scientific methods requires improved perspectives for discriminating and evaluating data. In research teams, having appropriate data discrimination and evaluation standards—such as understanding data source authority and the rationality and effectiveness of data acquisition and analysis methods—can help improve teams' ability to correctly utilize data, thereby facilitating knowledge evaluation capability formation and ultimately promoting knowledge innovation.

5. Conclusion

This study constructs a research team data capability model in the data-intensive research environment, focusing on six dimensions: data awareness, data retrieval and discovery capability, data organization/integration/management capability, data description and preservation capability, data analysis and knowledge discovery capability, and data discrimination and evaluation capability. It also proposes how these capabilities promote related knowledge innovation capabilities. However, the model does not further investigate the actual effects of these promotional relationships or whether all capabilities positively promote team knowledge innovation. Therefore, the model needs refinement through further empirical research. Future work will establish hypotheses based on this model, test them, and refine the model accordingly.

References

- [1] HEY T, TANSLEY S, TOLLE K. *The fourth paradigm: data-intensive scientific discovery* [M]. Washington: Microsoft Research, 2009.
- [2] Li Lirui, Deng Zhonghua. Research on data literacy education for scientific big data from an “Internet +” perspective [J]. *Library*, 2016(11): 92-96.
- [3] Wang Xiaowen, Shen Si. Review and implications of foreign researcher data literacy education [J]. *Information and Documentation Services*, 2017(3): 102-106.
- [4] Tang Wuxiang. *On Innovation* [M]. Beijing: China Braille Press, 1999.
- [5] ROGERS D M A. Knowledge innovation system: the common language [J]. *The journal of technology studies*, 1993: 19(2): 2-8.
- [6] NONAKA I, TOYAMA R, KONNO N. SECI, Ba, and leadership: a unified model of dynamic knowledge creation [J]. *Long range planning*, 2000, 33(1): 5-34.
- [7] Yang Bin, Xiong Wanling, You Jing. Empirical research on knowledge innovation capability enhancement paths in university research teams based on knowledge transfer [J]. *Information Studies: Theory & Application*, 2011, 34(8): 60-64.
- [8] KHEDHAOURIA A, JAMALA. Sourcing knowledge for innovation: knowledge reuse and creation in project teams [J]. *Journal of knowledge management*, 2015, 19(5): 932-948.
- [9] HONG J, LEE O K, SUH W. Creating knowledge within a team: a socio-technical interaction perspective [J]. *Knowledge management research & practice*, 2017, 15(1): 23-33.
- [10] Zhao Limei, Sun Yanhua. Research on characteristics and connotations of internal knowledge integration in university research teams for knowledge

innovation [J]. *Science and Technology Management Research*, 2015, 35(1): 171-176.

[11] Liu Yanfang, Yuan Yongjiu. Research on organizational internal knowledge integration levels for knowledge innovation [J]. *Information Science*, 2012(12): 119-122.

[12] Guo Yanli, Yi Shuping. Research on task-based team knowledge innovation models based on knowledge potential [J]. *Information Studies: Theory & Application*, 2013, 36(1): 20-24.

[13] Li Gang, Ba Zhichao. Research on factors influencing knowledge stickiness in research teams [J]. *Journal of Library Science in China*, 2017, 43(1): 89-106.

[14] Wang Xinchun, Qi Guijie, Liang Yikai, et al. Research on improving knowledge creation capabilities in open innovation communities [J]. *Journal of Intelligence*, 2016, 35(3): 203-206.

[15] Wu Yang, Su Jun. Research on complex characteristics of research team knowledge innovation systems and coordination mechanism effects [J]. *Science of Science and Management of S. & T.*, 2012, 33(1): 156-165.

[16] Wu Yang, Li Xiaoqiang, Xia Di. Research on feedback mechanisms of communication management in research team knowledge innovation processes [J]. *Science & Technology Progress and Policy*, 2012, 29(1): 7-10.

[17] Dong Chunyu, Xue Yonghong. Data-intensive, big data and the “fourth paradigm” [J]. *Studies in Dialectics of Nature*, 2017, 33(5): 74-80, 86.

[18] Deng Zhonghua, Li Zhifang. Evolution of scientific research paradigms—the fourth paradigm of scientific research in the big data era [J]. *Information and Documentation Services*, 2013, 34(4): 19-23.

[19] Huang Xin, Deng Zhonghua. Requirement analysis and support for data-intensive scientific research [J]. *Information Studies: Theory & Application*, 2017, 40(2): 66-70.

[20] Liang Na, Zeng Yan. Promoting data-intensive scientific discovery to enhance technological innovation capabilities: new models, methods, and challenges—publication of the translation of *The Fourth Paradigm: Data-Intensive Scientific Discovery* [J]. *Bulletin of Chinese Academy of Sciences*, 2013(1): 115-121.

[21] CARLSON J, JOHNSTON L, WESTRA B, et al. Developing an approach for data management education: a report from the data information literacy project [J]. *The international journal of digital curation*, 2013, 8(1): 204-217.

[22] DCC curation lifecycle model [EB/OL]. [2017-08-26]. <http://www.dcc.ac.uk/resources/curation-lifecycle-model>.

[23] UK Data Archive research data lifecycle [EB/OL]. [2017-08-14]. <http://www.data-archive.ac.uk/create-manage/life-cycle>.

[24] BOULTON R, CAMPBELL P, COLLINS B, et al. *Science as an open enterprise* [R]. London: Royal Society, 2012.

[25] Science 2.0 consultation: Royal Society response [EB/OL]. [2017-06-14]. <https://royalsociety.org/~media/policy/Publications/2014/science-2-0-consultation-response-290914.pdf>.

[26] KOLTAY T. Data literacy for researchers and data librarians [J]. *Journal of librarianship & information science*, 2017, 49(1): 3-14.

Author Contributions: Du Xingye: research design, literature review, drafting; Li He: research design, revision; Li Zhuozhuo: data collection and organization.

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

Source: ChinaXiv — Machine translation. Verify with original.