

Knowledge Management Processes, Technologies, Models, and Strategies in the Big Data Era: Postprint

Authors: Ye Yingping, Chen Haitao, Chen Hao

Date: 2023-07-26T00:00:00+00:00

Abstract

[Purpose/Significance] The era of big data has brought transformation to traditional knowledge management. To adapt to the advent of the big data era, it is necessary to re-conceptualize and reposition the knowledge management process, apply big data technological tools to construct a new knowledge management model, and address the evolving challenges in knowledge management.

[Method/Process] By examining the formation process from data to knowledge, integrating the 4V characteristics of big data, and understanding the big data technological tools required for knowledge management, this study divides knowledge management in the big data era into four stages: knowledge production, knowledge accumulation, knowledge exchange, and knowledge application. Combined with the utilization of big data technological tools, a new knowledge management model is constructed, and countermeasures are proposed to address issues such as fragmentation integration, application value, hardware support, and privacy ethics.

[Results/Conclusion] The development of big data has driven the transformation of the knowledge management process from traditional models to technology-driven models. The knowledge management model in the big data era is closely intertwined with the application of big data technology, emphasizing the extraction of knowledge value from massive fragmented data to more effectively support organizational decision-making. To achieve this, it is essential to ensure hardware infrastructure support and information security safeguards, integrate big data technology with small sample analysis, and advance knowledge management to a new level.

Full Text

Introduction

While existing research has extensively examined the decision-support functions and value-added processes brought by big data development, studies on the knowledge 增值 (value-added) process remain insufficient. To address this gap, this paper first reviews the formation process of “data-information-knowledge,” analyzing the transformation from traditional knowledge management to big data knowledge management. It then introduces knowledge management tools and approaches in the big data era based on the 4V characteristics of big data, constructs a new knowledge management model grounded in big data contexts, and proposes corresponding countermeasures from four perspectives: data fragment integration, information security assurance, decision-support assistance, and hardware infrastructure support. These efforts aim to better achieve the objectives of knowledge management in big data environments—enhancing organizational competitiveness and innovation capabilities through knowledge management, thereby promoting improvements in social production efficiency.

Based on the literature review above, we can see that in the big data context, massive amounts of fragmented data from diverse channels undergo extraction, cleaning, and transformation before entering data warehouses to form specific data collections. Multiple data collections are then interconnected through online networks to build multidimensional data cubes, enabling dynamic analysis and integration into information resources. Through descriptive and predictive analysis, this information is ultimately visualized to externalize knowledge, presenting it as intuitive visual profiles and forming knowledge graphs through inter-knowledge associations.

The knowledge formation process is illustrated in Figure 1 [Figure 1: see original paper]. The transformation from data to knowledge involves four stages: data warehousing, online analytical processing, information mining, and knowledge presentation, accompanied by the formation of data flow, information flow, knowledge flow, and value flow. The value of all data, information, and knowledge is ultimately realized through the explicit expression of knowledge graphs. Moreover, considering the cross-integration of massive data in the big data era, the knowledge formation process is not a one-to-one relationship from data to knowledge but rather a many-to-many relationship. Consequently, in the construction of knowledge graphs, corresponding data networks, information networks, knowledge networks, and value networks also exist.

Regarding the relationship between data and knowledge, P. Cooper [2] argues that the descriptive meaning of data in a specific environment constitutes information, and structuring and organizing this information forms knowledge. In essence, data represents basic information, while information can be processed and refined into knowledge. In the big data context, large volumes of fragmented raw data are scattered and isolated. Through processing and interpretation, associations are established between data to form information, which is then

processed and analyzed into knowledge for application in production and daily life. Simply hoarding data and information is meaningless; only through data mining and knowledge management can data and information serve economic and social development, thereby unleashing the true value of big data. Zhao Rongying et al. [3] contend that knowledge involves extracting, integrating, and solidifying data and information.

Knowledge management refers to the management of internal knowledge assets by enterprises (or other organizations), with its core task being the integration of explicit and tacit knowledge to achieve knowledge innovation and value creation through processes such as knowledge acquisition, absorption, and transfer. To this end, Japanese scholar Ikujiro Nonaka et al. [9] proposed the SECI model, which posits that tacit knowledge completes a spiral ascent through four processes: socialization, externalization, combination, and internalization. Knowledge management hinges on knowledge transfer, and numerous scholars have researched ways to enhance it by increasing the transfer willingness of knowledge providers, strengthening the absorptive capacity of knowledge recipients, and appropriately shortening knowledge distance, striving to expand knowledge search scope and optimize knowledge encoding methods to improve transfer efficiency and integration effectiveness.

With the advent of the big data era, scholarly focus in knowledge management has shifted from transforming tacit knowledge into explicit knowledge to how to mine knowledge value from massive, multi-source heterogeneous data. Consequently, domestic and international scholars have conducted a series of studies on the theoretical development and technical applications of big data in knowledge management. Theoretically, A. Simonet et al. [10] introduced the concept of active knowledge based on the agility of big data increments, arguing that big data's flexibility can effectively enhance the expressive capability of knowledge management. Qin Dianqi [11] applied big data concepts to knowledge integration theory, discussing knowledge management models, characteristics, and specific steps at the individual and organizational levels. In applied research, C. S. Liew et al. [12] explained how to leverage mobile intelligent terminal platforms in big data environments to mine effective information from data for knowledge management. Huang Tian'en et al. [13] designed a big data-based simulation model using knowledge management methods for power grid early warning, addressing integrated security issues in energy internet environments. He Jun [14] systematically elaborated on how big data changes enterprise knowledge management and strategic decision-making based on evolving decision-making environments.

Scholars generally agree that knowledge management in big data environments requires rapid processing of larger datasets and more optimized methods to mine knowledge value from massive low-density data. Ye Yingping et al. [15] studied network embedding-based knowledge management models. Compared with traditional knowledge management research, big data differs in various knowledge management processes: knowledge acquisition, storage, integration, and use.

First, knowledge acquisition has shifted from traditional knowledge management's targeted search and proactive acquisition from external organizations to the emergence of massive fragmented data, expanding the scope and volume of knowledge acquisition and changing its direction. Second, knowledge storage has transitioned from private knowledge repositories to shared distributed storage, with this cloud-based approach expanding storage capacity and breaking ownership restrictions. Third, knowledge integration has evolved from integrating structured, codifiable data in traditional knowledge management to logical integration of multi-source heterogeneous data that is structured, semi-structured, and unstructured, increasing integration difficulty while optimizing integration methods. Fourth, the key prerequisite for knowledge use has shifted from explicating tacit knowledge to data cleaning and desensitization. Information sharing in the big data era has transformed the application of data processing results from serving knowledge 增值 (value-added) for a single entity to being used by multiple parties.

Based on the transformation from traditional to big data knowledge management, the knowledge management process in the big data context can be summarized as shown in Figure 2 [Figure 2: see original paper]. Over time, the degree of knowledge processing deepens gradually. Knowledge acquisition is completed through collecting massive fragmented data, knowledge is stored in a distributed manner, knowledge integration is achieved by processing multi-source heterogeneous data, and data analysis results are ultimately applied to realize business value.

3. Technical Tools for Knowledge Management in the Big Data Context

Unlike traditional knowledge management processes, knowledge management in the big data context requires more high-tech tools and software. The application of various big data practical technologies promotes the transformation of knowledge management processes, thereby enhancing knowledge management levels and organizational operational efficiency. Meng Xiaofeng [4], while studying computer technology development in the big data era, introduced numerous big data processing tools, including user-generated content, perceptual systems, stream processing and batch processing models, human-computer interaction technology, Caffeine indexing, and the Pregel graph computing model. According to the 4V characteristics of big data—vast volume, dynamic growth, multi-source heterogeneity, and low value density—rapidly increasing large datasets require data cleaning tools for knowledge acquisition; massive semi-structured and unstructured data demand data warehousing tools for knowledge storage; complex and uneven-quality data need cloud computing tools for knowledge integration; and dynamically increasing raw data require visualization tools for data mining to extract knowledge value for use.

Accordingly, the data cleaning, storage, analysis, and mining tools involved in the four knowledge management processes of knowledge acquisition, storage,

integration, and use are shown in Figure 3 [Figure 3: see original paper].

3.1 Data Cleaning Techniques and Tools

The vast volume of big data does not increase data value but rather signifies more data noise and information garbage proliferation. Therefore, after acquiring massive amounts of data, the first step is data cleaning as a preprocessing task to transform noisy, non-clean data into high-quality clean data for better subsequent data analysis and knowledge management. This non-cleanliness of data manifests in two aspects: inaccuracy and incompleteness of data itself, and redundancy and inconsistency between data. Peng Yu et al. [16] define data cleaning as the process of extracting and integrating required data sources, extracting relational entities, and uniformly defining structures after association and aggregation, aiming to improve data quality and lay a foundation for data analysis. Cheng Xueqi et al. [17] also point out that data cleaning is a prerequisite for data analysis, including four methods: materialization or ETL-based, federated database or middleware-based, data stream-based, and search engine-based. In the big data era, subtle yet useful information is mixed within massive data volumes, so the data cleaning process requires a trade-off between quality and quantity—overly fine granularity may filter out useful information, while overly coarse granularity fails to achieve the cleaning effect. Using ETL (extraction, transformation, load) tools, data from different operating platforms, programming languages, and physical locations can be extracted in a unified format, then cleaned, transformed, integrated, and loaded into data warehouses.

3.2 Cloud Distribution Techniques and Tools

The disruptive impact of big data on data storage is not merely a matter of technical evolution and capacity expansion. From data sources to data thinking, it brings revolutionary changes to data storage. Due to enormous storage space requirements, cloud-based distributed data warehouses are needed for data storage. The dynamic growth of big data poses challenges to data warehouse response times. Unlike the concept of static relational databases in traditional knowledge management, data warehouses also involve dynamic access to big data in decision support systems, requiring data processing within data warehouses. Jin Cheqing et al. [18] compare traditional databases with emerging data storage technologies, arguing that data warehouses possess data processing functions beyond storage, specifically divided into stream processing and batch processing models. Stream processing relies more on summary data structures in memory, requiring storage-class memory (SCM) devices such as PCM (phase-change memory), exemplified by Twitter's Storm, Yahoo's S4, and LinkedIn's Kafka. The representative of batch processing models is Google's MapReduce programming model, which divides data sources into blocks, assigns them to different Map tasks, executes Map functions on parsed key/value pairs, writes intermediate results to local hard drives, and then Reduce tasks sort by key values and output final results.

3.3 Online Processing Techniques and Tools

Qiu Dong [19] analyzed the challenges posed by big data environments to data processing from a statistical perspective. Online processing provides elastic expansion, low cost, rich patterns, and massive computing capabilities for big data, enabling simultaneous processing of heterogeneous systems in structured, semi-structured, and unstructured forms. It organizes computing capabilities through networks to achieve scaled and specialized data processing. With the rise of 4G networks and mobile terminals, the concept of mobile cloud online processing has emerged, using application software to collectively process and collaboratively compute networked storage data, thereby achieving the purpose of extracting data value through online processing.

3.4 Visualization and Presentation Techniques and Tools

Visualization technology is a necessary means for data mining and knowledge discovery, an important tool for promoting knowledge dissemination and transformation, and a fundamental technical basis for knowledge management systems. Through visualization technology, knowledge representation is enriched, knowledge comprehension difficulty is reduced, and people's willingness to receive and learn knowledge is stimulated, promoting knowledge sharing both within and outside organizations. Ren Lei et al. [20] argue that visual analytics, through summarization expression, better enables knowledge base invocation and control, facilitates knowledge iteration and updating, and introduce common visualization technologies including tag clouds, history flows, and spatial information flows, as well as visualization software such as RShiny, Flot, Raphael, Google Charts, and SAS Visual Analytics. In practical applications, appropriate visualization technologies and programs can be selected based on data mining and knowledge discovery needs to realize data value utilization. Guan Sifa et al. [21] point out that both online processing and visualization presentation require high data authenticity—only relatively accurate and trustworthy high-quality data yield meaningful and valuable computation and visual analysis results. If data itself contains errors or problems, data analysis and mining results will also be incorrect. Additionally, visualization technology requires metadata management. Metadata is data about data, reflecting associations between data and inherent data attributes. Metadata representation facilitates visualization and presentation of data.

4. Construction of Knowledge Management Model in the Big Data Era

In the big data era, data has transformed from the processing object of traditional knowledge management to the foundational resource of knowledge management. These real-time dynamic data, unstructured data, and even machine-generated data create massive volume and multi-source heterogeneity complexity, posing challenges to traditional knowledge management's knowledge stor-

age and integration. Knowledge management in big data environments exhibits characteristics of surging knowledge volume, faster processing speeds, greater knowledge diversity, lower knowledge value density but higher total value, and more informatized and intelligent processing. The growth in data volume cannot be solved by simply scaling up knowledge management systems—quantitative changes lead to qualitative changes. Traditional knowledge management concepts and technologies can no longer cope with massive big data. The update and processing speeds of knowledge management systems have exceeded the acceleration methods and logical loads of traditional knowledge management, requiring new models.

The knowledge management model in the big data era consists of four stages: knowledge production, knowledge accumulation, knowledge exchange, and knowledge application. In the knowledge production stage, massive fragmented data is collected through knowledge discovery and acquisition. In the knowledge accumulation stage, the collected data is stored in knowledge warehouses. As knowledge production and accumulation progress, the relationships between knowledge become increasingly complex, causing difficulties for subsequent group feedback processes and enabling targeted precision 推送 (push). For instance, Baidu's search 词条 (entries) purify knowledge value during accumulation and establish inter-knowledge relationships, using knowledge graphs to provide precise 推送 (push) to customers from massive databases. Unlike static knowledge repositories in traditional knowledge management, knowledge warehouses feature dynamic storage characteristics to enhance real-time indexing capabilities. During accumulation, continuously accumulating customer and market knowledge complements organizational knowledge reserves, effectively remedying knowledge shortages and perfecting the knowledge system.

When facing knowledge exchange needs, knowledge in knowledge warehouses can be shared. However, to protect knowledge source property rights and enable multiple possibilities for knowledge development, knowledge exchange platforms must be established to share desensitized knowledge in independent exchange spaces, supplementing and updating original knowledge systems based on shared new knowledge. To provide personalized search services, Baidu associates and stores data information from multiple terminal devices using the same IP address and login account, sharing and cross-analyzing it with browsing records and transaction information from other platforms. In the knowledge exchange stage, knowledge sharing enables organizational knowledge innovation—existing knowledge is updated when available, and new knowledge is created when absent. New knowledge integrates with original knowledge and connects with actual organizational operations, advancing innovation in products, services, and business models.

In the knowledge application stage, knowledge evaluation in big data contexts differs from traditional business research by providing comprehensive, objective, and detailed evaluation results through knowledge analysis and mining to assist organizational managers in making correct strategic decisions. Additionally,

based on knowledge evaluation data, knowledge value and relationships can be assessed to filter out useless, repetitive, and outdated knowledge, constructing orderly knowledge networks. Knowledge evaluation services can also provide knowledge products and decision-making consulting to customers, achieving external knowledge service purposes. For example, Baidu competitively ranks entries based on index click-through rates. After organizations complete the four knowledge management processes—production, accumulation, exchange, and application—evaluation data from knowledge application is recaptured in knowledge discovery and acquisition, forming knowledge feedback. This enables knowledge to achieve spiral 增值 (value-added) through repeated cycles. During organizational decision-making and operational activities, associated data and behavioral data (metadata about data) are generated and flow back to organizations along with the big data torrent, also opening to all internet entities. Through knowledge management, these participate repeatedly in various stages.

5. Strategies for Knowledge Management in the Big Data Era

Knowledge management in the big data era uses data analysis and mining techniques to transform massive fragmented data into valuable knowledge, providing knowledge services for socio-economic development. Thus, raw data, technical platforms, and application value are three key elements for realizing knowledge management in the big data era. Additionally, to ensure orderly knowledge management, information security issues warrant attention. To address these challenges, this section proposes relevant strategies.

5.1 Fragment Integration of Raw Data and Countermeasures

Massive fragmented data in the big data era divides information into discontinuous, illogical 零碎 (fragmented) units, driving fragmentation in knowledge dissemination across thinking, expression, content, and subjects. Ren Fubing [22] and Zhang Wende et al. [23] both note that redundant information makes knowledge acquisition temporally and spatially random and uncertain—knowledge acquisition and transmission occur in fragmented time and variable spaces. Meanwhile, rapid data generation and updates demand higher timeliness and responsiveness from knowledge management. To cope with big data era knowledge management, organizations need greater transformative thinking, with timely change in communication, analysis, and decision-making capabilities to adapt to the “fragmented” environment and better leverage knowledge management. Knowledge dissemination forms are shifting toward “concise and interesting” user-generated content, yet similar viewpoints are expressed broadly and fragmentarily, making fragmented information acquisition more complex and time-consuming. Knowledge integration involves cross-fragmented information across multiple types and domains. Therefore, establishing knowledge warehouses enables rapid and complete acquisition of massive associated data, organizing redundant fragmented information under big data through system-

atic analysis principles to build required knowledge frameworks and human-computer interaction platforms. Systematic, structured knowledge bases create unified fragmentation processing rules, add query modules with access functions, and set user demand permissions, thereby improving fragmented data acquisition speed and information integration accuracy.

5.2 Application Value of Decision Support and Countermeasures

The purpose of knowledge management lies in knowledge services. Big data knowledge management extends beyond 梳理 (sorting) historical experiences to establishing sufficiently large sample libraries covering all production and life domains through massive data. Based on Markov chain calculations, event occurrence probabilities can be inferred and future development directions deduced. This information prediction based on big data analysis is objective and scientific, and with deep machine learning development, its prediction accuracy will continuously improve. Big data knowledge management is gradually breaking free from incomplete information constraints, making it applicable in organizational decision support to provide decision-support services for customers. Gu Xitong et al. [24] introduced data mining into electronic health services and medical diagnosis decision support. Hu Xiaorong et al. [25] applied data analysis to risk identification and decision support for listed companies. B. Gu et al. [26] used data-driven approaches for customer satisfaction evaluation to replace traditional questionnaires and customer 回访 (follow-ups). Knowledge management can assist decision-making, and correspondingly, decision support systems can also be applied to organizational knowledge management behavior.

5.3 Hardware Support for Technical Platforms and Countermeasures

With internet development, more software tools are applied to knowledge management processes, such as Microsoft's OneNote and knowledge management systems like ServiceHot. In contrast, academia and industry pay less attention to hardware support facilities for knowledge management, with relatively insufficient infrastructure construction. Knowledge management platforms are comprehensive systems composed of expert systems, knowledge bases, and other technical tools, along with various infrastructures supporting knowledge collection, processing, storage, transmission, and utilization, enabling rapid knowledge structure analysis, sharing, and categorized processing of complex knowledge content. Constructing knowledge management platforms requires not only advanced data analysis and knowledge management software but also crucial infrastructure. Damaged storage devices cause platform information loss and leakage; processor failures paralyze system operations; during online processing, faster devices waste time waiting for slower devices, causing computational capacity and energy consumption waste, with potentially immeasurable economic losses. Therefore, to enhance knowledge management capabilities in the big data context, efficient infrastructure matching software must be introduced, with timely hardware updates and maintenance, and emergency measures for

crisis situations to ensure basic equipment operation and safeguard data and information security.

5.4 Privacy Ethics in Information Security and Countermeasures

While big data profoundly changes social development and human life, it also brings non-negligible data privacy issues and information security risks. Huang Guobin et al. [27] argue that privacy leakage risks arise from combinations of different data sources and superposition of multi-channel information, increasing inference probabilities of user identity characteristics. Wang Shiwei [28] also believes that widespread application of big data analysis techniques like clustering and sentiment analysis is gradually obtaining personal user characteristics and 窥视 (peeping into) private lives, with user privacy information leakage potentially causing social problems like misleading and fraud, affecting social stability. Big data era knowledge content is 庞杂 (vast and complex), and much acquired information cannot be fully utilized. Therefore, in practice, acquiring only data information needed for academic research and industrial development becomes key to solving privacy protection issues. During data cleaning and desensitization, comprehensive analysis of demand objectives and expected results is needed to plan information scale, structure, and hierarchy. When acquiring or exchanging information, precise 对接 (matching) with required data structures is necessary to expand effective information proportions. Additionally, technical measures remain essential for information security, such as firewall and anonymity protection technologies to effectively monitor and detect malicious information acquisition behaviors and maintain data security, while hiding user identity information and behavioral characteristics to enhance user privacy protection awareness.

The big data era signifies more precise knowledge management and more intelligent knowledge services, bringing opportunities: (1) Systematization of knowledge management—although machine-automated massive data collection reduces value density compared to traditional manual knowledge searching, the geometric growth in information volume still yields enormous total value with richer and more complete value systems; (2) Intelligentization of knowledge management—combining machine learning wisdom can liberate humans from mid- to low-end knowledge management tasks, focusing on advanced knowledge production and application; (3) Scientification of knowledge management—organizations relying on big data analysis for operational decisions become more objective and accurate, with data mining assisting in predicting future development directions for more scientific strategic decisions.

However, big data also exposes potential hazards and challenges: (1) Data security issues—massive data openness brings data leakage and privacy security challenges, yet excessive information narrowing affects analysis accuracy. Balancing data security and value, and balancing data openness and privacy protection during data cleaning and desensitization, are critical concerns; (2) Processing large amounts of heterogeneous data—knowledge management must

address the diversified explosive growth of data volume, especially unstructured information like images and videos, challenging storage and processing. Reconstructing organizational business processes with big data thinking to improve analysis accuracy and timeliness is essential; (3) Hardware infrastructure issues—increasingly complex and heterogeneous hardware environments cause time waste during online processing waiting for slower servers, creating 木桶效应 (barrel effect) challenges for infrastructure construction. Improving overall hardware cluster performance is crucial; (4) Energy consumption—as data volume grows, data center storage scale expands, making high energy consumption a bottleneck for big data development. How to optimize knowledge management and services while reducing energy consumption requires future solutions. However, this has not yet attracted sufficient social attention, with scholars focusing more on practical application research. Therefore, integrating big data with other technologies to improve knowledge resource production and supply capabilities, while combining big data technology with small sample analysis, represents important 融合手段 (integration methods) for big data thinking and industry.

Big data is not a cold world filled with algorithms and machines. By strengthening knowledge management and service awareness, we look forward to a knowledge society entering the era of universal big data.

References

- [1] Liu Zhihui, Zhang Quanling. Review of big data technology research [J]. Journal of Zhejiang University (Engineering Science), 2014, 48(6): 957-972.
- [2] COOPER P. Data, information, knowledge and wisdom [J]. Anaesthesia & intensive care medicine, 2014, 15(1): 44-45.
- [3] Zhao Rongying, Wei Xuqiu. Gathering knowledge into wisdom: A framework model of knowledge management in big data environments [J]. Information Studies: Theory & Application, 2017, 40(9): 20-23.
- [4] Meng Xiaofeng, Cixiang. Big data management: Concepts, techniques and challenges [J]. Journal of Computer Research and Development, 2013, 50(1): 146-169.
- [5] Xu Libo, Pan Xuwei, Yuan Ping, et al. Knowledge intelligence emergence innovation: Concept, system and path [J]. CAAI Transactions on Intelligent Systems, 2017, 12(1): 47-54.
- [6] Liu Jiexuan. Data governance in university library knowledge management [J]. Information Science, 2018, 36(1): 108-129.
- [7] Liu Jie. Research on optimization of state-owned enterprise document management based on knowledge management in big data environment [J]. Sci-Tech Information Development & Economy, 2014, 24(17): 121-123.

- [8] Zeng Runxi, Wang Lin, Du Hongtao. Research on big data research network and structure from the perspective of knowledge management [J]. Journal of the China Society for Scientific and Technical Information, 2016, 35(11): 1173-1184.
- [9] Nonaka Ikujiro, Takeuchi Hirotaka. The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation [M]. Beijing: China Water & Power Press, 2006.
- [10] SIMONET A, FEDAK G, RIPEANU M. Active data: A programming model for managing big data lifecycle across heterogeneous systems and infrastructures [J]. Future generation computer systems, 2015, 53(9): 25-42.
- [11] Qin Dianqi. Personal knowledge organization under the concept of integration and big data [J]. Information Studies: Theory & Application, 2014, 37(2): 19-22.
- [12] LIEW C S, WAH T Y, SHUJA J, et al. Mining personal data using smartphones and wearable devices: A survey [J]. Sensors, 2015, 15(2): 4430-4469.
- [13] Huang Tian'en, Sun Hongbin, Guo Qinglai, et al. Knowledge management and advanced security warning based on power grid operation simulation big data [J]. Power System Technology, 2015, 39(11): 3080-3087.
- [14] He Jun. Analysis of big data's impact on enterprise management decisions [J]. Science & Technology Progress and Policy, 2014, 31(4): 65-68.
- [15] Ye Yingping, Lu Yanqiu, Xiao Yanhong. Construction of knowledge innovation model based on network embedding [J]. Library and Information Service, 2017, 61(7): 102-110.
- [16] Peng Yu, Pang Jingyue, Liu Datong, et al. Big data: Connotation, technology system and prospect [J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(4): 469-482.
- [17] Cheng Xueqi, Jin Xiaolong, Wang Yuanzhuo, et al. Survey on big data systems and analytical techniques [J]. Journal of Software, 2014, 25(9): 1889-1908.
- [18] Jin Cheqing, Qian Weining, Zhou Minqi, et al. Benchmarking data management systems: From traditional databases to emerging big data [J]. Chinese Journal of Computers, 2015, 38(1): 18-34.
- [19] Qiu Dong. Challenges of big data era to statistics [J]. Statistical Research, 2014, 31(1): 16-22.
- [20] Ren Lei, Du Yi, Ma Shuai, et al. Survey on visual analysis of big data [J]. Journal of Software, 2014, 25(9): 1909-1936.
- [21] Guan Sifa, Chao Lemen. Key issues, challenges and countermeasures of information analysis in big data era [J]. Library and Information Service, 2015, 59(3): 12-18.

- [22] Ren Fubing. Fragmentation and jigsaw puzzle: Diffusion and integration of network communication [J]. *Information and Documentation Services*, 2014, 35(3): 18-24.
- [23] Zhang Wende, Cheng Han, Liu Tian. User decision-oriented fragmentation integration model for university information [J]. *Information Studies: Theory & Application*, 2018, 41(3): 64-67.
- [24] Gu Xitong, Zhang Xiaofei, Liu Xiaoxiao, et al. Data-driven electronic health service management research: Challenges and prospects [J]. *Journal of Management Science*, 2017, 30(1): 3-14.
- [25] Hu Xiaorong, Yao Changqing, Gao Yingfan. Research on risk identification method and visualization for listed companies based on automatic risk phrase extraction [J]. *Journal of the China Society for Scientific and Technical Information*, 2017, 36(7): 663-673.
- [26] GU B, YE Q. First step in social media measuring the influence of online management responses on customer satisfaction [J]. *Production and operations management*, 2014, 23(4): 570-582.
- [27] Huang Guobin, Zheng Lin. Research on big data information security risk framework and coping strategies [J]. *Library Science Research*, 2015, 33(13): 24-29.
- [28] Wang Shiwei. On new characteristics and requirements of information security in big data era [J]. *Library and Information Service*, 2016, 60(6): 5-14.

Author Contributions

Ye Yingping: Wrote the paper content; Chen Haitao: Designed the research framework; Chen Hao: Responsible for data collection and proofreading.

Knowledge Management Process, Technical Tools, Models and Strategies in the Big Data Era

Ye Yingping, Chen Haitao, Chen Hao

Management School of Jilin University, Changchun 130022

Abstract: [Purpose/significance] The big data era has brought changes to traditional knowledge management. To adapt to the advent of big data, it is necessary to understand the knowledge management process, use technical tools to build new knowledge management models, and overcome knowledge management challenges. [Method/process] After summarizing the formation process from data to knowledge, together with the 4V features of big data and analysis of big data technology tools needed for knowledge management, knowledge management in the big data era is divided into four stages: knowledge production, knowledge accumulation, knowledge exchange, and knowledge application.

With the help of technology tools, a new model of knowledge management is built. Solutions are proposed to problems such as debris consolidation, application value, hardware support, and ethics on privacy. [Result/conclusion] Big data promotes the development of the transformation of the knowledge management process from the traditional model. The model of knowledge management in the period of big data is closely related to the use of technology tools. It emphasizes the value of knowledge extracted from mass fragmentation data and more effectively assists organizational decision-making. Therefore, hardware support, information security, and combining big data technology with small sample analysis will push knowledge management to a new level.

Keywords: big data era; technical tool; knowledge management model

Call for Papers: *Knowledge Management Forum*

Knowledge Management Forum (ISSN 2095-5472, CN 11-6036/C) has been approved by the State Administration of Press, Publication, Radio, Film and Television of China for formal online publication qualifications. It was newly revised in 2016 and selected into the internationally renowned open access journal directory (DOAJ) in 2017. This journal focuses on research results regarding knowledge production, creation, organization, integration, mining, sharing, analysis, utilization, and innovation. Any knowledge management issues concerning governments, enterprises, universities, libraries, and other physical and virtual organizations—including theories, methods, tools, technologies, applications, policies, solutions, and best practices—fall within the scope of this journal. The journal publishes by article, with accepted manuscripts entering a rapid publication process and achieving immediate full open access.

The 2019 issues will focus on: Internet + knowledge management, big data and knowledge organization, communities of practice and knowledge operation, content management and knowledge sharing, knowledge creation and open innovation, data mining and knowledge discovery. We now solicit manuscripts from domestic and international academic and industrial circles:

1. Manuscript themes should relate to knowledge, discussing knowledge management, knowledge services, knowledge innovation, and related issues. Articles may focus on theory, application, technology, methods, models, or best practices.
2. Articles must be substantive, integrating theory with practice, with clear research purposes, appropriate research methods, original academic insights, and reference or guidance value for theory or practice.
3. All submissions must undergo similarity detection, peer expert review, and pass the editorial department's initial review, re-review, and final review.
4. Article length is not limited, but generally 4,000-20,000 words are appropriate.
5. Submissions will be notified of acceptance within one month.

Please submit manuscripts at www.kmf.ac.cn. Contact: Liu Yuanying, Tel: 010-82626611-6638.

The journal primarily publishes online through our website (www.kmf.ac.cn) and authorized databases, practicing open access, publishing by article, and print-on-demand.

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

Source: ChinaXiv — Machine translation. Verify with original.