Recent Advances and Trends in Knowledge Fusion Research: Postprint

Authors: Zhu Xiang, Zhang Yunqiu

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

Abstract

[目的/意义]To review and evaluate research on knowledge fusion in recent years, aiming to provide reference for future related research.[方法/过程]First, the concept of knowledge fusion is analyzed. Then, the framework, process, and methods of knowledge fusion are reviewed. Subsequently, research trends in knowledge fusion are summarized. Finally, research prospects are presented.[结果/结论]Knowledge fusion research presents new characteristics in the big data environment, but still fails to meet the requirements of the big data environment. Future research should focus on four aspects: constructing a hierarchical, multi-dimensional, and three-dimensional big data knowledge fusion framework; improving the efficiency of knowledge fusion; constructing real-time dynamic fusion mechanisms; and conducting big data empirical application research.

Full Text

Preamble

ChinaXiv Collaborative Journal
Volume 63, Issue 16, August 2019

Recent Advances and Trends in Knowledge Fusion Research

Zhu Xiang, Zhang Yunqiu
Department of Medical Informatics, School of Public Health, Jilin University, Changchun 130021

Abstract

[**Purpose/Significance**] This paper reviews and evaluates recent research on knowledge fusion to provide references for future studies. [**Method/Process**] We first analyze the concept of knowledge fusion, then systematically review its frameworks, processes, and methods, summarize research trends, and finally propose future research directions. [**Result/Conclusion**] Knowledge fusion research has exhibited new characteristics in the big data environment, yet still fails to meet the demands of this environment. Future research should focus on four aspects: constructing a hierarchical, multi-dimensional, and comprehensive big data knowledge fusion framework; improving knowledge fusion efficiency; building real-time dynamic fusion mechanisms; and conducting empirical big data application studies.

Keywords: knowledge fusion; big data; review
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2019.16.015

With the advent of the big data era, data sources have become more extensive and data structures more complex and diverse. Multi-source heterogeneity represents the most significant challenge for "full data" analysis in the big data era, and knowledge fusion offers an effective solution to this problem. It enables deep semantic-level processing of data, generates new knowledge, and provides novel services. Knowledge fusion has become increasingly important and has emerged as a new focus in information science research. However, knowledge fusion research remains in its infancy with few mature studies. Therefore, it is necessary to systematically review and evaluate current knowledge fusion research and propose future directions to guide subsequent studies.

Using "knowledge fusion" and "知识融合" as search terms, we conducted topic searches in Web of Science Core Collection and CNKI from January 2012 to September 2018 to retrieve relevant literature published since the beginning of the big data era. After screening, we obtained 55 closely related papers. Further backward tracing yielded 13 additional relevant papers, totaling 68 papers that form the basis of this study. Following a conceptual analysis of knowledge fusion, we systematically reviewed three key aspects—framework, process, and method—according to the logical sequence of "framework-process-method," then summarized research trends and proposed future research directions oriented toward the big data environment.

1 Conceptual Analysis of Knowledge Fusion

Knowledge fusion (KF) evolved from data fusion and information fusion. Data fusion (DF) originated in the 1970s in the military domain, primarily integrating data from multiple sensors to form comprehensive judgments about targets [1]. In the 1990s, the concept of information fusion (IF) emerged, extending processing objects from data to multi-source information. It involves comprehensive processing of multi-source information to obtain more accurate and reliable reasoning decisions than single information sources [2]. By the late 1990s, knowledge fusion began to attract attention. From the perspective of knowledge science in library and information science, management, and computer science, knowledge fusion objects are no longer limited to sensor information but include existing knowledge bases, knowledge extracted from databases, networks, and business systems, and even various methods and expert experiences. Some scholars argue that the three concepts intersect significantly in research objects, processing units, fusion processes, and results [3], while others believe knowledge fusion encompasses data fusion and information fusion [4]. Since data and information are both processing objects and products of knowledge fusion in the big data environment, and the three have transformational relationships, this paper considers knowledge fusion to include data fusion and information fusion.

In the development of knowledge fusion, two definitional perspectives exist: the first emphasizes that knowledge fusion is the process of converting multi-source knowledge into a unified knowledge schema, while the second emphasizes fusion results, indicating that knowledge fusion is not simple integration of multi-source heterogeneous data sources but is characterized by the generation of new knowledge—a view widely accepted in recent big data environments. Synthesizing definitions proposed by researchers [5-8], this paper defines knowledge fusion in the big data environment as: a process oriented toward knowledge services and decision support, based on massive multi-source heterogeneous data, information, and knowledge, utilizing fusion algorithms and rules to combine, reason, and create new knowledge at the semantic level.

2 Knowledge Fusion Frameworks

Knowledge fusion frameworks provide the foundation for implementing knowledge fusion. Framework research has always been a key focus, generally integrating the latest theoretical methods and guiding method and application research. This section reviews knowledge fusion frameworks from macro and micro perspectives: macroscopically, they can be divided into method-oriented and service-oriented frameworks; microscopically, they can be viewed from horizontal (multi-dimensional) and vertical (layered) perspectives. Finally, we review frameworks specifically proposed for the big data environment.

2.1 Macro-Level Framework Research

International knowledge fusion framework research began early, with the most notable being the KRAFT project completed in 2001 [9]. The key to the KRAFT framework lies in converting multi-source heterogeneous resources into a unified knowledge model, influencing subsequent research. Macro-level frameworks can be categorized into two types: method-oriented frameworks (e.g., ontology and rule-based frameworks) and service-oriented frameworks (e.g., decision support systems for intelligent assistance and early warning), as shown in Table 1 .

Table 1 Orientations of Knowledge Fusion Frameworks

Orientation	Examples
Method-oriented	Ontology and rule-based frameworks [10]; multi-clue fusion with background knowledge [11]; classification rule-enhanced systems [12]; context-based frameworks [13]; linked data-based frameworks [14]; probability theory-based frameworks [15]
Service-oriented	Decision support systems for intelligent assistance/early warning [16]; digital reference consultation frameworks [17]; cloud manufacturing efficiency frameworks [18]; precision agriculture frameworks [19]; health assessment management frameworks [20]; edge domain knowledge verification frameworks [21]

2.2 Micro-Level Framework Research

Almost all knowledge fusion frameworks exhibit hierarchical divisions. Vertically, frameworks are divided into different layers; horizontally, fusion is divided into different dimensions, as shown in Table 2 .

Table 2 Layers and Dimensions of Knowledge Fusion Frameworks

Perspective	Examples
Layered fusion (vertical)	Pre-fusion, during fusion, post-fusion [22]; base library, association-level, feature-level, requirement-level fusion [23]; data layer, model layer, application layer [24]; knowledge resource layer, knowledge organization layer, knowledge service layer [25]

Perspective	Examples
Multi-dimensional fusion (horizontal)	Knowledge layer, method layer, thought layer fusion [26]; attribute, instance, and concept fusion [27]; data layer, model layer, parameter layer fusion [12]; simple, extended, instance, parameter, adaptive, faceted, and historical fusion [13]; content, structure, application dimensions [28]; instance, domain set, relation, attribute, and concept fusion [29]

Layered fusion represents a vertical perspective showing different stages of the fusion process. Although layer names vary, they share essential similarities and can be summarized into three layers: data layer, fusion layer, and application layer. The data layer aims to associate data or information resources using methods like ontology and knowledge elements to form unified knowledge representation. The fusion layer identifies conceptual relationships, forms effective organization systems, and creates new knowledge through reasoning and fusion algorithms. The application layer matches fusion results with user needs to achieve effective knowledge services.

If layered fusion is vertical, multi-dimensional fusion is horizontal. Knowledge includes instances, attributes, concepts, relations, etc. Multi-dimensional fusion reflects different knowledge aspects, constituting fusion across different dimensions.

2.3 Big Data-Oriented Framework Research

Current big data-oriented knowledge fusion framework research is primarily domestic. Wang Yuefen et al. [30] designed framework structures and components for big data environments, including functional/application systems, operation/guarantee systems, and evaluation/optimization systems. Tang Xiaobo et al. [31] proposed a big data knowledge fusion framework consisting of data acquisition/knowledge representation, unified schema construction, fusion processing, and derived knowledge processing. Zhang Xinyuan et al. [32] designed a framework aimed at providing more precise knowledge services, including pre-fusion, fusion process, and fusion service components.

The big data environment features real-time changes, making dynamic knowledge fusion a new research hotspot. Su Xinning et al. [33] improved fusion algorithms to establish dynamic knowledge link databases corresponding to knowledge base entries, enabling rapid generation of dynamic knowledge points. L.

Tan et al. [34] proposed dynamic functional model tools and knowledge fusion methods to achieve effective and efficient online consumer participation.

3 Knowledge Fusion Process

Due to knowledge complexity, direct fusion is difficult. Unified formal knowledge representation is a prerequisite, so the basic approach involves first representing knowledge uniformly. Current representation forms mainly include knowledge elements and knowledge networks. Accordingly, we categorize knowledge fusion into knowledge element fusion and knowledge network fusion.

3.1 Knowledge Element Fusion

Most knowledge fusion research in information science uses knowledge elements (KE) for unified representation. The main idea is to convert various data sources into knowledge elements as the fusion data foundation, then fuse them using algorithms and rules supported by ontologies to obtain usable new knowledge. The basic process is shown in Figure 1 [Figure 1: see original paper].

While many scholars have studied knowledge elements, no unified definition or formal representation exists. Typically, a knowledge element is the indivisible minimal unit that can semantically express knowledge completely [37]. Knowledge elements are "knowledge about knowledge" that can describe structural knowledge features [38]. Assuming knowledge element A:

$$A = \{(P_1, R_1), (P_2, R_2), \dots, (P_j, R_j), \dots, (P_n, R_n)\}$$

where P represents knowledge element attribute features and R represents attribute values.

Based on fusion objects, knowledge fusion can be divided into three levels: decision-level, feature-level, and data-level [35]. Decision-level and feature-level are advanced fusion, while data-level is basic fusion. As levels decrease, data volume and difficulty increase. Early research focused on advanced fusion due to computational limitations, but recent big data developments have shifted focus toward basic fusion, reflecting full-data rather than small-sample analysis. In network big data fusion frameworks, Zhou Liqin et al. [29] collected, extracted, and represented network big data, while Zhu Juan et al. [36] fused multi-source information in a three-layer knowledge model for personalized product recommendations.

3.2 Knowledge Network Fusion

In big data multi-source heterogeneous knowledge, complex relationships exist between different knowledge entities. Knowledge networks can structurally display these relationships, making network fusion a new research hotspot. A knowledge network (KN) is a network structure composed of knowledge nodes/elements and associations, with knowledge bases as primary carriers. As shown in Figure 2 [Figure 2: see original paper], the main idea is to fuse acquired knowledge elements/networks from knowledge bases or networks with existing ones to form new knowledge networks for knowledge discovery.

Microsoft's Probase [44] and Google's Knowledge Vault [45] have proposed models for acquiring and fusing knowledge from network big data to form knowledge networks. Lin Hailun et al. [46] summarized methods for fusing new knowledge into existing networks from three aspects: entity expansion, relation expansion, and classification expansion. Zhou Liqin et al. [29] extracted knowledge elements from unstructured data sources and fused them into the hypertension ontology in DiseaseOntology to build domain problem-solving knowledge bases. Tian Pengwei et al. [47] categorized heterogeneous information network fusion methods into five types: meta-path extraction, multi-relational networks, and hypergraph/hypernetwork modeling.

4 Knowledge Fusion Methods

Knowledge fusion methods are key to implementation. Current methods fall into three categories: semantic-based, information fusion algorithm-based, and graph model-based methods.

4.1 Semantic-Based Methods

A key difference between knowledge and information is knowledge's reasoning capability, with semanticization being the main approach. Semantic-based methods can be subdivided into semantic rule-based and ontology-based methods.

4.1.1 Semantic Rule-Based Methods

These methods rely on defined logical operation rules to constrain fusion conditions. J. Gou et al. [48] described comparison rules to clarify fusion conditions and filter knowledge objects through result formats and semantic rules, guiding the fusion process to avoid illogical results. Z. Feng et al. [49] used semantic similarity for semantic disambiguation to achieve knowledge fusion between primary lexical chains. G. Wang et al. [50] proposed a multi-source process innovation knowledge fusion algorithm based on semantic element reconstruction with corresponding semantic conflict resolution rules.

4.1.2 Ontology-Based Methods

Ontology has good conceptual hierarchical structures, supports logical reasoning, and can describe knowledge at the semantic level through conceptual models, making it suitable for knowledge expression. Many knowledge fusion frameworks are ontology-based, often belonging to the data layer along with knowledge elements. Ontology mapping and fusion achieve knowledge fusion using knowledge organization theory. The German Ontobroker project [51] first introduced ontology into knowledge fusion frameworks. N. Xie et al. [52] implemented multi-source agricultural information fusion based on agricultural domain ontologies. J. Liu et al. [53] proposed a dynamic ontology construction method for more effective knowledge fusion. A. Smirnov et al. [54] built RDF ontologies for robots and used ontology matching for semantic interoperability. H. Fan et al. [55] studied knowledge fusion implementation patterns based on ontology knowledge representation.

4.2 Information Fusion Algorithm-Based Methods

Knowledge fusion inherits from information fusion, so its algorithms have been adapted for knowledge fusion, mainly Bayesian networks and D-S evidence theory.

4.2.1 Bayesian Networks

Bayesian networks are graphical networks for probabilistic reasoning that can discover potential relationships between data. They are widely used for uncertainty and incompleteness problems with simple, direct calculations. E. Santos et al. [56] represented probabilistic models as Bayesian knowledge bases and proposed Bayesian knowledge fusion algorithms. K. Coussement et al. [57] proposed a Bayesian decision support framework formally fusing subjective expert opinions with objective organizational information. L. Zhang et al. [58] designed knowledge category preference contexts using Bayesian methods. W.C. Yue et al. [59] proposed a fuzzy Bayesian network method for multi-source knowledge fusion reasoning.

4.2.2 D-S Evidence Theory

D-S evidence theory [60] addresses evidence and possibility reasoning to eliminate uncertainty, representing a development beyond Bayesian theory with weaker conditions. R. Yan et al. [61] applied D-S theory to software fault diagnosis. G. Peng et al. [62] developed a D-S-based method for product development. L. Sun [63] extended D-S theory for multi-attribute fusion. L.Q. Sun et al. [64] applied D-S-based knowledge fusion to study H5N1 avian influenza spatial distribution.

4.3 Graph Model Methods

Graph models [65] have become popular, including probabilistic graphs, topic graphs, or relation graphs. By obtaining prior knowledge from other data types, they assign probabilities to knowledge as link prediction problems on graphs. Y.J. Wu et al. [66] created a large topic graph for the Gulf oil spill incident, validating topic graphs for knowledge fusion. G. Levchuk et al. [67] proposed probabilistic graph fusion algorithms for multi-source text media. G. Koumoutsos et al. [68] proposed a scalable framework for graphically presenting structured and unstructured resource knowledge to achieve fusion.

5 Knowledge Fusion Research Trends

Based on the above review, current knowledge fusion research exhibits these characteristics: (1) With big data development and technological advances, interdisciplinary tools like ontology, semantic web, multi-Agent, linked open data, XML, RDF, Hadoop, and MapReduce have been applied. Most frameworks can process multi-source heterogeneous data and depend on adaptable knowledge extraction and unified representation. (2) Knowledge elements and knowledge networks, as structured representation methods, have become mainstream, with knowledge network fusion becoming a new trend. (3) Semantic-based methods dominate, with ontology methods gaining prominence. Knowledge fusion has also inherited information fusion algorithms and recently developed graph model methods.

However, limitations exist: (1) No unified, general-purpose framework exists; method-service integrated, hierarchical, multi-dimensional frameworks are lacking. (2) Knowledge element creation still requires manual participation, affecting efficiency. Knowledge network modeling and fusion for multi-source, heterogeneous, large-scale data remains difficult, constraining progress. For example, hypergraph/hypernetwork methods mainly remodel rather than fuse large-scale heterogeneous data. (3) Semantic rule-based methods require manual rule definition, limiting efficiency and applicability for big data, especially web data. Many algorithms depend on ontology interoperability, whose complexity reduces efficiency, and automatic ontology construction remains challenging. Information fusion algorithms also struggle in the big data era—Bayesian networks become too large-scale, causing storage and efficiency issues, while D-S theory is unsuitable for large-scale data fusion. Graph model methods based on external closed knowledge sets lack scalability.

6 Research Prospects

Big data has brought abundant data sources, new technologies, and challenges including diversity, dynamics, fragmentation, and uncertainty. Natural language diversity causes synonymous and polysemous expressions. Real-time updates in networks, business systems, and databases create conflicts between old and new values. Knowledge fragmentation and complex relationships make systematic knowledge collection difficult. Randomness and uncertainty are more pronounced in dynamic big data environments, making multi-source heterogeneous data processing more challenging than ever.

Future research should focus on four aspects:

6.1 Constructing Hierarchical, Multi-Dimensional, and Comprehensive Big Data Knowledge Fusion Frameworks

Frameworks must adapt to big data environments, process large-scale multi-source heterogeneous data, and meet fine-grained, precise, deep fusion needs. As shown in Figure 3 [Figure 3: see original paper], big data technology and demands continuously evolve. Layered fusion should be more refined, multi-dimensional fusion more comprehensive, and fusion levels more basic and full-data oriented. A "layered + multi-dimensional + full-data" architecture will lead to deep fusion—comprehensive, multi-angle fusion that mines deep relationships, makes implicit knowledge explicit, discovers new knowledge, and meets users' deep needs.

6.2 Improving Knowledge Fusion Efficiency

Future research should develop more automated, structured knowledge representation methods, including automated construction of knowledge elements and networks, while reducing ontology dependency. A key challenge is leveraging ontology advantages while overcoming its inherent difficulties. More efficient and accurate fusion algorithms are needed that can evaluate knowledge authenticity and reliability and analyze knowledge structure deeply and clearly.

6.3 Building Real-Time Dynamic Fusion Mechanisms

In big data environments, especially web data, knowledge changes constantly with inconsistent update frequencies across sources. This dynamicism reflects temporal changes and variations caused by different people, environments, backgrounds, and experiences. Current methods mainly target static knowledge; dynamic knowledge fusion remains exploratory. Future research should incorporate temporal information, develop dynamic fusion methods, and build real-time dynamic feedback adjustment mechanisms to address big data dynamic challenges.

6.4 Conducting Big Data Empirical Application Research

Current research focuses more on frameworks and methods than empirical studies. Existing empirical studies use small-scale, less heterogeneous data, leaving the effectiveness of fusion methods for real big data environments unproven. Urgent need exists for genuine big data empirical research on knowledge fusion.

References

[1] WALTZ EE, LINAS J. Intelligence fusion pushed[J]. Aviation week and space technology, 1979(1): 205-211.

[2] TAHANI H, KELLER JM. Information fusion in computer vision using the fuzzy integral[J]. IEEE transactions on systems man & cybernetics, 1990, 20(3): 733-741.

[3] LIU Xiaojuan, LI Guangjian, HUA Bolin. Knowledge fusion: conceptual analysis and definition[J]. Library and information service, 2016, 60(13): 13-19, 32.

[4] ZHU Zhenyuan, LI Guangjian. Preliminary exploration of knowledge fusion from a "data-information-knowledge" holistic perspective: correlation and comparison of data fusion, information fusion, and knowledge fusion[J]. Information studies: theory & application, 2017, 40(2): 12-18.

[5] SMIRNOV A, PASHKIN M, CHILOV N, et al. Multi-agent architecture for knowledge fusion from distributed sources[C]//International workshop of central and eastern Europe on multi-agent systems. Heidelberg: Springer, 2001: 293-302.

[6] MARTENS D, DEBACKER M, HAESEN R, et al. Ant-based approach to the knowledge fusion problem[C]//International workshop on ant colony optimization and swarm intelligence. Heidelberg: Springer, 2006: 84-95.

[7] HU X, HU J, SEKHARIA, et al. A fuzzy knowledge fusion framework for terms conflict resolution in concurrent engineering[J]. Concurrent engineering, 2011, 19(1): 71-84.

[8] QIU Junping, YU Houqiang. Progress and trends of international knowledge fusion research from the perspective of knowledge science[J]. Library and information service, 2015, 59(8): 126-132, 148.

[9] REECE A, HUI K, GRAY A, et al. Designing for scalability in a knowledge fusion problem[C]//International workshop on ant colony optimization and swarm intelligence. Heidelberg: Springer, 2006: 84-95.

[10] HU X, HU J, SEKHARIA, et al. A fuzzy knowledge fusion framework for terms conflict resolution in concurrent engineering[J]. Concurrent engineering, 2011, 19(1): 71-84.

[11] FONTANI M, ARGONES-RUA E, TRONCOSO C, et al. The watchful forensic analyst: multi-clue information fusion with background knowledge[C]//IEEE International Workshop on Information Forensics and Security. Guangzhou: IEEE, 2013: 120-125.

[12] FISCH D, KALKOWSKI E, SICK B. Knowledge fusion for probabilistic generative classifiers with data mining applications[J]. IEEE transactions on knowledge and data engineering, 2014, 26(3): 652-666.

[13] SMIRNOV A, LEVASHOVA T, SHILOV N. Patterns for context-based knowledge fusion in decision support systems[J]. Information fusion, 2015, 21: 114-129.

[14] GAO Jinsong, LIANG Yanqi. Research on knowledge fusion model in linked data environment[J]. Information science, 2016, 34(2): 50-54.

[15] TAIC T, CHANG CT, CHANG YS. Hybrid knowledge fusion and inference on cloud environment[J]. Future generation computer systems, 2018, 87: 568-579.

[16] KRIEGELER U, PFENNIGSCHMIDT S, ZIEGLER HG. Practical aspects of the use of a knowledge fusion toolkit in safety applications[C]//IEEE eleventh international symposium on autonomous decentralized systems. Mexico: IEEE, 2013: 1-4.

[17] SHEN Wang, LI Yafeng, HOU Haochen. Research on digital reference consultation knowledge fusion framework[J]. Library and information service,

2013, 57(19): 139-143.

[18] LIU J, XU W, ZHAN H. Multi-source and heterogeneous knowledge organization and representation for knowledge fusion in cloud manufacturing[C]//Proceedings of international conference on soft computing techniques and engineering applications. New Delhi: Springer, 2014: 55-61.

[19] KAWTRAKULA, KHUNTHONG V, SUKTARACHAN M, et al. Development of an information integration and knowledge fusion platform for spatial and time based advisory services: precision farming as a case study[C]//Annual SRII global conference. San Jose: IEEE, 2014: 241-248.

[20] YUE K, QIAN W, FU X, et al. Qualitative probabilistic network-based fusion of time-series uncertain knowledge[J]. Soft computing, 2015, 19(7): 1953-1972.

[21] LI F, DONG XL, LANGEN A, et al. Knowledge verification for long-tail verticals[J]. Proceedings of the VLDB endowment, 2017, 10(11): 1370-1381.

[22] WANG F, FAN H, LIU G. Big data knowledge service framework based on knowledge fusion[C]//Proceedings of the international joint conference on knowledge discovery, knowledge engineering and knowledge management. Porto: Lda, 2016: 116-123.

[23] LI Lirui, DENG Zhonghua. Research on small data fusion for scientific users under the "Internet+" background[J]. Library and information service, 2016, 60(6): 58-63.

[24] ZHU Juan, TANG Xiaobo. Personalized product recommendation based on three-layer knowledge fusion model[J]. Research on library science, 2017(5): 24-30.

[25] HUANG Xinping. Research on multi-dimensional semantic knowledge fusion for government website information resources[D]. Changchun: Jilin University, 2017.

[26] ZHOU Fang, LIU Yuzhan, HAN Liyan. Research on knowledge fusion method based on fuzzy set theory[J]. Journal of Beijing Institute of Technology (social sciences edition), 2013, 15(3): 67-72.

[27] XIE Nengfu. Research on agricultural knowledge fusion framework based on agricultural ontology and fusion rules[J]. Anhui agricultural sciences, 2013, 41(1): 395-397.

[28] CHEN Weidong, WANG Ping, WANG Yicheng, et al. Multi-dimensional semantic knowledge fusion for government website information resources[J]. Information studies: theory & application, 2017, 40(6): 111-116.

[29] ZHOU Liqin, FAN Hao, PAN Jianpeng. Research on knowledge fusion framework in network big data[J]. Journal of intelligence, 2018, 37(1): 145-150, 197.

[30] WANG Yuefen, CEN Yonghua. Research on architecture design of knowledge fusion system in big data era[J]. Digital library forum, 2016(10): 16-24.

[31] TANG Xiaobo, ZHU Juan, YANG Fenghua. Research on knowledge fusion framework model in big data environment[J]. Research on library science, 2016(1): 32-35.

[32] ZHANG Xinyuan, QIU Junping. Research on knowledge fusion framework in big data environment[J]. Research on library science, 2016(8): 66-70.

[33] SU Xinning. Knowledge organization theory and methods for knowledge services[M]. Beijing: Science Press, 2014.

[34] TAN L, TANG D, WANG Q, et al. Open design pattern, method, and its self-organization mechanism[J]. Procedia CIRP, 2016, 56: 34-39.

[35] LIU Weibin, ZOU Zhiyuan, XING Weiwei. Feature fusion methods in pattern classification[J]. Journal of Beijing University of Posts and Telecommunications, 2017, 40(4): 1-8.

[36] ZHU Juan, TANG Xiaobo. Personalized product recommendation based on three-layer knowledge fusion model[J]. Research on library science, 2017(5): 24-30.

[37] SUO Chuanjun, GAI Shuangshuang. Research on the connotation, structure, and description model of knowledge elements[J]. Journal of library science in China, 2018, 44(4): 54-72.

[38] WEN Youkui, JIAO Yuying. Knowledge discovery based on knowledge elements[M]. Xi'an: Xidian University Press, 2011.

[39] GOU Jin. Research on key technologies in knowledge fusion[D]. Hangzhou: Zhejiang University, 2005.

[40] SUN L. Knowledge element-based competitive intelligence analytics for SWOT situation assessment[C]//The 8th international symposium on computational intelligence and design. Hangzhou: IEEE, 2015, 1: 576-579.

[41] TANG Xiaobo, ZHU Juan. Review of key issues in knowledge fusion under big data environment[J]. Library journal, 2017, 36(7): 10-16.

[42] TANG Xiaobo, WEI Wei. Knowledge fusion: growth point of knowledge services in big data era[J]. Research on library science, 2015(5): 9-14, 8.

[43] SUN Lin, WANG Yanzhang. Research on multi-source competitive intelligence fusion method based on knowledge elements[J]. Journal of intelligence, 2017, 36(11): 65-71.

[44] WU W, LI H, WANG H, et al. Probase: a probabilistic taxonomy for text understanding[C]//Proceedings of the 2012 ACM SIGMOD international conference on management of data. Scottsdale: ACM, 2012: 481-492.

[45] DONG X, GABRILOVICH E, HEITZ G, et al. Knowledge vault: a Web-scale approach to probabilistic knowledge fusion[C]//Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining. New York: ACM, 2014: 601-610.

[46] LIN Hailun, WANG Yuanzhuo, JIA Yantao, et al. Survey on knowledge fusion methods for network big data[J]. Chinese journal of computers, 2017, 40(1): 1-27.

[47] TIAN Pengwei, ZHANG Xian, HU Zhengyin, et al. Survey on heterogeneous information network fusion methods[J]. Library and information service, 2017, 61(7): 137-144.

[48] GOU J, JIANG Y, WU Y, et al. A new knowledge fusion method based on semantic rules[C]//The 8th international conference on signal processing. Beijing: IEEE, 2006: 1865-1868.

[49] FENG Z, LI-JUAN Z, XIAO-MING F, et al. The research of Vietnamese language news clue extraction method based on converged network semantic knowledge[C]//The 27th Chinese conference on control and decision. Qingdao: IEEE, 2015: 3287-3289.

[50] WANG G, HU Y, TIAN X, et al. An integrated open approach to capturing systematic knowledge for manufacturing process innovation based on collective intelligence[J]. Applied sciences, 2018, 8(3): 340.

[51] FENSEL D, DECKER S, ERDAMANN M, et al. Ontobroker: how to make the WWW intelligent[C]//Proceedings of the 11th Banff knowledge acquisition for knowledge-based systems. Banff: AAAI, 1998: 36-42.

[52] XIE N, WANG W, YANG X, et al. Rule-based agricultural knowledge fusion in Web information integration[J]. Sensor letters, 2012, 10(1/2): 635-638.

[53] LIU J, XU W, JIANG H. Research on dynamic ontology construction method for knowledge fusion in group corporation[J]. Advances in intelligent systems & computing, 2014, 278: 289-298.

[54] SMIRNOV A, KASHEVNIK A, MIKHAILOV S, et al. Multi-level robot self-organization in smart space: approach and case study[C]//Conference on smart spaces. Cham: Springer, 2015: 68-79.

[55] FAN H, WANG F, ZHENG M. Research on knowledge fusion connotation and process model serving for knowledge service demand[C]//China conference on knowledge graph and semantic computing. Singapore: Springer, 2016: 184-195.

[56] SANTOS JR E, WILKINSON JT, SANTOS EE. Fusing multiple Bayesian knowledge sources[J]. International journal of approximate reasoning, 2011, 52(7): 935-947.

[57] COUSSEMENT K, BENOIT DF, ANTICO M. A Bayesian approach for incorporating expert opinions into decision support systems: a case study of

online consumer satisfaction detection[J]. Decision support systems, 2015, 79: 24-32.

[58] ZHANG L, QIAN W, ZHANG Z, et al. Research on knowledge demand information acquisition for product design[C]//The 12th international conference on computational intelligence and security. Wuxi: IEEE, 2016: 677-680.

[59] YUE W, CHEN X, GUI W, et al. A knowledge reasoning fuzzy-Bayesian network for root cause analysis of abnormal aluminum electrolysis cell condition[J]. Frontiers of chemical science and engineering, 2017, 11(3): 414-428.

[60] DEMPSTER AP. Upper and lower probabilities induced by a multivalued mapping[J]. Annals of mathematical statistics, 1967, 38(2): 325-339.

[61] YAN R, LI G, LIU B. Knowledge fusion based on DS theory and its application on expert system for software fault diagnosis[C]//IEEE international conference on big data. Boston: IEEE, 2017: 1991-2000.

[62] PENG G, MAO H, WANG H, et al. BOM-based design knowledge representation and reasoning for collaborative product development[C]//Symposium on computational intelligence for security and defense applications. Verona: IEEE, 2015: 1-10.

[63] SUN L, WANG Y. A multi-attribute fusion approach extending Dempster-Shafer theory for combinatorial-type evidences[J]. Expert systems with applications, 2018, 96: 218-229.

[64] SUN L, WARD MP, LI R, et al. Global spatial risk pattern of highly pathogenic avian influenza H5N1 virus in wild birds: a knowledge-fusion based approach[J]. Preventive veterinary medicine, 2018, 152: 32-39.

[65] LAO N, MITCHELL T, COHEN WW. Random walk inference and learning in a large scale knowledge base[C]//Proceedings of the conference on empirical methods in natural language processing. Stroudsburg: Association for Computational Linguistics, 2011: 529-539.

[66] WU Y, LEHMAN A, DUNAWAY DJ. Evaluations of a large topic map as a knowledge organization tool for supporting self-regulated learning[J]. Knowledge organization, 2015, 42(6): 386-398.

[67] LEVCHUK G, BLASCH E. Probabilistic graphical models for multi-source fusion from text sources[C]//Symposium on computational intelligence for security and defense applications. Verona: IEEE, 2015: 1-5.

[68] KOUMOUTSOS G, FASLI M, LEWIN I, et al. Graph-based information exploration over structured and unstructured data[C]//IEEE international conference on big data. Boston: IEEE, 2017: 1991-2000.

Author Contributions:

Zhu Xiang: Proposed research ideas and framework, wrote and revised the paper;

Zhang Yunqiu: Determined research topics, provided revision suggestions, improved and revised the research.

Abstract: [Purpose/significance] Knowledge fusion, as an effective method to deal with multi-source heterogeneous data and generate new knowledge semantically, has become a new research point of information science in the big data environment, but it is still in its infancy. The paper aims to sort out, evaluate the current research on knowledge fusion, and provide reference for future research. [Method/process] Firstly, the concept of knowledge fusion was analyzed. Then the framework, process and method of knowledge fusion were combed. Then the research trend of knowledge fusion was summarized. Finally, the research prospect was made. [Result/conclusion] Knowledge fusion research presented new research characteristics in the big data environment, but it can't meet the requirements of the big data environment. In the future, we should build a hierarchical and multi-dimensional big data knowledge fusion framework, improve the efficiency of knowledge fusion, build real-time dynamic fusion mechanism, and carry out big data empirical research based on knowledge fusion.

Keywords: knowledge fusion; big data; review