Knowledge Co-clustering: Knowledge Aggregation Patterns from a Domain Analysis Perspective Postprint

Authors: Chen Guo, Wu Wei, Xiao Lu

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

Abstract

[Purpose/Significance] Current research on knowledge aggregation patterns focuses on "what kinds of knowledge associations are used to conduct knowledge aggregation." This paper supplementarily explores the subsequent question of "what forms knowledge units are aggregated into using knowledge associations," aiming to improve knowledge aggregation pattern research and guide in-depth practice.

[Method/Process] Drawing on the classification method of two major types of polymerization reactions in chemistry, this paper proposes to categorize knowledge aggregation into knowledge homopolymerization and knowledge copolymerization based on "whether to preserve the differences among knowledge units and their associations," and explores the basic implementation forms of knowledge copolymerization.

[Results/Conclusion] Domain knowledge is the foundation for implementing knowledge copolymerization; with documents and terms as the basic granularities of knowledge units, and user demand entry points and aggregation target resources as dimensions, knowledge copolymerization can be implemented through four basic forms: facet-based navigation, multi-dimensional concept association recommendation, knowledge element linking, and discovery of latent associations among resources.

Full Text

Knowledge Copolymerization: A Knowledge Aggregation Mode from the Perspective of Domain Analysis

Chen Guo[1], **Wu Wei**[1], **Xiao Lu**[2] [1]Department of Information Management, School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094 [2]School of Journalism, Nanjing University of Finance and Economics, Nanjing 210046

Abstract:

[**Purpose/Significance**] Current research on knowledge aggregation modes focuses on "what types of knowledge associations should be used for knowledge aggregation." This paper complementarily explores the subsequent question of "what forms knowledge units can be aggregated into using these associations," aiming to improve knowledge aggregation mode research and guide practical implementation. [**Method/Process**] Drawing on the classification of polymerization reaction types in chemistry, we propose dividing knowledge aggregation into knowledge homopolymerization and knowledge copolymerization based on "whether the differences between knowledge units and their associations are preserved," and explore the basic implementation forms of knowledge copolymerization. [**Result/Conclusion**] Domain knowledge serves as the foundation for knowledge copolymerization. Taking documents and words as the basic granularities of knowledge units, and using user demand portals and aggregation target resources as dimensions, knowledge copolymerization can be implemented through four basic forms: faceted navigation, multidimensional conceptual association recommendation, knowledge element linking, and discovery of potential resource associations.

Keywords: knowledge organization; knowledge aggregation mode; knowledge copolymerization; domain knowledge analysis; domain conceptual relation
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2018.08.015

Knowledge aggregation has become a research hotspot in the field of library and information science. The term "aggregation" corresponds to the English word "Aggregation," referring to a combination of heterogeneous components or a whole formed by combining multiple small units [1]. At the application level, a typical example is "polymerization reaction" in chemistry, where monomer molecules connect with each other to form new polymer materials [2]. In information science and related fields, concepts such as "aggregation," "fusion," and "integration" have emerged with knowledge, data, information, and resources as objects. Li Yating distinguished and defined these concepts: aggregation focuses on discovering multiple associations among objects and establishing corresponding systems; fusion emphasizes new changes produced after processing objects; integration focuses on comprehensively managing objects based on their common rules [3]. Knowledge aggregation addresses the problem that valuable knowledge resources are highly dispersed and disorderly distributed due to the rapid growth of information resources. Compared with data integration and information integration, knowledge aggregation represents a qualitative improvement in aggregation intensity and granularity. Knowledge aggregation is the prerequisite for achieving diversified, deep-level, and open knowledge organization and utilization based on user needs.

In recent years, numerous studies on "aggregation" have emerged in China's library and information science field, with most projects and research results focusing on academic resources (with library collections as typical representatives) as objects [5-6]. Research on knowledge aggregation modes mostly distinguishes modes based on the knowledge associations used for aggregation. For example, He Defang et al. summarized semantic-based deep aggregation methods for library collections into three types: concept association-based, citation relationship-based, and research ontology-based [5]. Zhao Rongying et al. categorized collection resource aggregation modes into traditional and deep aggregation modes, with the former including catalog-based and metadata-based modes, and the latter including ontology-based, linked data-based, topic model-based, and bibliometric analysis-based knowledge aggregation modes [7]. Subsequent research on knowledge aggregation can mostly be classified into these types. Notably, after reviewing the concepts and components of information aggregation, Cao Shujin et al. proposed that information aggregation modes include not only relation-based and granularity-based aggregation but also context-based aggregation [8].

Although current research has thoroughly explored "what associations should be used for aggregation," this only addresses the first half of the knowledge aggregation mode question. After obtaining rich and deep-level knowledge associations (such as semantic associations), the question of "what forms knowledge units can be aggregated into using these associations" should also be answered by knowledge aggregation mode research, yet this aspect has received relatively little attention. In quite a few knowledge aggregation studies, substantial effort is spent obtaining deep-level knowledge associations, but they are ultimately used only for clustering knowledge units—for example, optimizing similarity calculation to divide clusters or constructing relationship networks to identify subgroups. In fact, clustering is only a primary method of knowledge aggregation. With full utilization of the fine-grained and diverse characteristics of knowledge associations, multiple knowledge aggregation methods can be implemented according to user needs and knowledge resource features.

---

2. Knowledge Homopolymerization and Knowledge Copolymerization

In chemistry, polymerization reactions can be divided into two modes based on whether the monomer types are consistent (i.e., whether the polymerization forces are singular): homopolymerization, where similar monomer molecules form polymers through the same force, and copolymerization, where multiple types of monomer molecules form polymers through different forces [2]. Similarly, knowledge aggregation also has analogous mode differences: if the type differences of knowledge units are not considered and their associations are simplified into a single "relevance relationship" (typically similarity), the resulting aggregation is called knowledge homopolymerization; if the type differences of knowledge units are considered and aggregation is performed based on fine-grained knowledge associations formed from the perspective of knowledge unit differences, it is called knowledge copolymerization.

2.1 Knowledge Homopolymerization

Knowledge homopolymerization is the most primitive, direct, and widely applied knowledge aggregation mode. Since type differences of knowledge units are not considered in aggregation, the quantitative calculation of "relevance relationships" is the core. Common relevance relationships include co-occurrence relationships and similarity relationships. Co-occurrence relationships refer to knowledge units appearing in the same time, space, or knowledge scenario, with quantification methods including proximity index, equivalence coefficient, and improved TF-IDF algorithms. Similarity relationship quantification mainly includes three categories: similarity calculation based on individual co-occurrence relationships; similarity transformation based on semantic knowledge bases; and semantic similarity calculation that fuses the first two results [9].

The implementation forms of knowledge homopolymerization mainly include two types: one places knowledge units in a corresponding feature space and divides clusters based on distance, typical examples being clustering and classification of words or texts; the other represents knowledge units and their relationships as networks and divides clusters based on specific network indicators and network subgroup discovery techniques, typical examples being word cluster division in co-word networks. Thus, clustering is only one implementation form of knowledge homopolymerization. Homopolymerization results map multi-dimensional knowledge associations onto a single dimension of relevance relationships, which is not conducive to achieving multi-dimensional combination of knowledge resources and therefore has obvious defects in practical applications. The typical problem is "knowing what is but not knowing why"—that is, knowing that associations exist between clustering results but not knowing what associations they are, which is not conducive to the expansion, refinement, and guidance of user needs.

2.2 Knowledge Copolymerization

Knowledge copolymerization mode does not emerge from nowhere. The main task of current knowledge organization has shifted from sorting and attribution to knowledge association and linking construction [11]. In knowledge organization systems, ontologies and concept maps are more advanced than classification systems and thesauri because they preserve the semantic type differences of knowledge units and knowledge associations [10]. Similarly, knowledge copolymerization focuses on aggregating knowledge units into knowledge copolymer bodies that can solve specific problems or tasks through diverse associations as forces in specific knowledge scenarios, making it a deeper knowledge aggregation method.

In practice, a typical application of knowledge copolymerization is faceted navigation. When users search for or click on a topic, faceted navigation systems display multi-dimensional associations between relevant results and search objects, a mode that has obvious advantages in web information organization [12]. Taking "mobile phone" as an example, Baidu's related search term guidance and JD.com's product navigation correspond to knowledge homopolymerization and knowledge copolymerization modes respectively, with the latter containing multiple facets related to mobile phone attributes. In addition, besides knowledge copolymerization at the word (concept) granularity, knowledge aggregation at the sentence, document, and user granularity can also be deepened from knowledge homopolymerization to knowledge copolymerization. For example, Tang Xiaobo et al. proposed aggregating Weibo posts about an event according to sentence logical relationships, from "cause" to "event" and then to "result" [13].

Thus, the result of knowledge copolymerization is a three-dimensional structure containing multiple facets (determined by the type differences of knowledge units and their associations), from which different knowledge unit clusters can be obtained through interpretation from different facets. After obtaining knowledge copolymerization results, users can select and utilize corresponding facets according to their specific needs. Although the results of knowledge homopolymerization can also be regarded as a three-dimensional space, they lack facets supported by knowledge unit differences and can only be divided into clusters by distance. Therefore, the knowledge copolymerization mode ensures the integrity and application flexibility of aggregation results, better meeting users' needs for knowledge exploration and utilization compared to knowledge homopolymerization.

---

3. Foundation of Knowledge Copolymerization: Domain Conceptual Relations

3.1 Domain Analysis Perspective in Knowledge Organization

In knowledge production, communication, and utilization, both user needs and knowledge content typically have domain characteristics, making domain knowledge essential for knowledge organization. Domain knowledge corresponds to general knowledge (such as common sense, logical rules, and mathematical operation knowledge) and refers to important concepts in specific domains and their associations [14]. In information services and intelligence work, merely pursuing optimization of information organization methods while ignoring the domain characteristics of knowledge leads to the problem of "emphasizing form over content" [15]. Therefore, the classic "domain analysis paradigm" has emerged in intelligence research methods, with its core idea being to examine human information activities from the perspective of specific knowledge domains [16].

In existing resource systems, the organizational structure of knowledge content is inconsistent with user demand structures. User knowledge needs often arise around the solution of specific tasks, determining that target knowledge has certain domain characteristics—that is, the required knowledge is constrained by the knowledge structure (such as concepts and conceptual associations) of the domain corresponding to the task. Due to the unknown nature of target knowledge, user needs are also latent and vague, making their knowledge-seeking behavior an iterative process of continuous learning, optimization, and feedback. However, in current resource systems, knowledge content is dispersed across different documents represented by "words." Without domain knowledge as a reference, knowledge associations must be established based on "relevance" to infer the direction of user need transformation, but this approach cannot effectively guide user needs or match them with knowledge resources. Research shows that mining conceptual associations without domain knowledge guidance produces many meaningless results that cannot meet users' specific needs [21].

Essentially, regardless of how vague user needs are or how distributed knowledge resources are, there is consistency between the two against the background of domain knowledge. In domain knowledge scenarios, knowledge units in knowledge resources collectively constitute knowledge copolymer bodies for solving specific problems through various forms of associations. The explicit and rich concept types and fine-grained conceptual associations contained in domain conceptual association systems serve as effective bases for revealing multi-dimensional knowledge associations between knowledge units such as words and documents. Through multi-dimensional, fine-grained aggregation of knowledge units based on these associations, users can be guided interactively and progressively to expand and refine their knowledge needs horizontally and vertically within the domain knowledge space, thereby effectively acquiring and utilizing knowledge.

3.2 Domain Conceptual Relations and Their Role in Knowledge Copolymerization

Relationships between knowledge units include equivalence relationships, hierarchical relationships, and associative relationships. General knowledge organization systems are built based on these relationships—for example, the "use, replace, broader term, narrower term, related term" relationships in the Chinese Thesaurus correspond to equivalence, hierarchical, and associative relationships. However, domain-specific knowledge organization pays more attention to concrete and differentiated "associative relationships" because different domains emphasize different features of the same concept. For instance, the concept of "water" has significantly different essential features emphasized in chemistry, physics, biology, and environmental science [20]. Correspondingly, after introducing domain background knowledge, the general "associative relationship" can be distinguished into rich and diverse fine-grained conceptual associations, such as method-application relationships and disease-symptom relationships.

Domain conceptual relations have two obvious characteristics compared to general conceptual associations: concepts have clearly defined connotations, and the types of conceptual associations are concrete, refined, and diversified [19].

As the essential level of knowledge associations, domain conceptual relations are the foundation for achieving fine-grained, multi-dimensional knowledge copolymerization. In knowledge copolymerization mode, the role of domain conceptual relations can be summarized as ensuring consistency between knowledge resource organizational structures and user knowledge demand structures—that is, the identification and utilization of knowledge units and their associations in existing resources should refer to domain conceptual association systems, and the transformation, expansion, and refinement of user needs should also be guided within the space corresponding to domain conceptual association systems.

---

4. Four Basic Implementation Forms of Knowledge Copolymerization

The manifestation of knowledge aggregation is closely related to its elements. Cao Shujin et al. summarized information aggregation elements as "aggregation object granularity, context, and relationship," and based on context analysis, explained aggregation mechanisms and introduced aggregation applications to sort out information aggregation modes and corresponding manifestations [8]. Based on this theoretical foundation, the basic implementation forms of knowledge copolymerization are as follows:

- **Aggregation object granularity**: Using words (such as keywords, tags, subject terms, concept terms, knowledge elements, etc.) and documents (such as web pages, short texts, academic literature, etc.) as the two most basic knowledge units in knowledge resource systems.
- **Aggregation context**: Using the switching paths between two types of knowledge units during users' knowledge exploration and utilization as the aggregation context—for example, obtaining articles through keywords and exploring knowledge related to terms of interest after reading articles. Overall, this includes four basic contexts: from word to word, from word to document, from document to word, and from document to document.
- **Aggregation relationship**: Using domain conceptual relations and the "document-word" inclusion relationship as aggregation bases. Domain conceptual relations are mainly used to build multi-dimensional association systems between words, while the "document-word" inclusion relationship maps the coarse-grained relationships of document resources to fine-grained conceptual associations at the word level, achieving multi-dimensional, multi-granularity aggregation of knowledge resources and targeted guidance for users.

Correspondingly, using words and documents as basic granularities, the switching paths between words and documents in users' knowledge exploration and utilization as contexts, and domain conceptual relations and "document-word" inclusion relationships as aggregation bases, several basic implementation forms of knowledge copolymerization can be summarized as shown in Figure 1 [Figure 1: see original paper]. The four quadrants in this figure indicate that when users explore corresponding aggregation targets (vertical axis) from corresponding demand portals (horizontal axis), they can utilize four basic implementation forms of knowledge copolymerization.

4.1 From Word to Document: Knowledge Aggregation Based on Faceted Navigation

Searching for corresponding document resources based on specific words is the most common knowledge exploration path for users. Both searching documents by keywords and browsing documents by navigation terms belong to this form, which can be collectively referred to as "navigation" for guiding from words to resources (mainly documents). Therefore, navigation is an important implementation form that must be considered in knowledge resource aggregation.

Traditional knowledge service platforms generally have problems such as flat navigation structures and lack of semantic associations between navigation terms, making it difficult to meet users' continuous and progressive knowledge exploration needs [22]. Therefore, many researchers have explored navigation optimization from the perspective of knowledge aggregation. Li Yating argued that navigation services based on knowledge aggregation should not rely solely on strict hierarchical structures but should focus on extracting aggregation categories from content [3]. Hu Yuan et al. proposed a framework for digital library community knowledge navigation systems based on knowledge aggregation [23]. Zhang Yunzhong et al., based on the complementary fusion of expert classification and folksonomy, proposed a tax-folk hybrid navigation model that achieves resource aggregation with a tree structure of "strong trunk, flourishing branches, and lush leaves" [24]. On the premise of fully utilizing knowledge association structures and resource attribute features, faceted navigation systems can be constructed [25].

Relative to traditional navigation systems based on similarity or classification/thesauri, faceted navigation systems fully preserve the diverse relationships between result documents and navigation vocabulary, while also better conforming to users' resource navigation habits. This is also the resource navigation implementation form under the knowledge copolymerization mode.

After introducing domain conceptual association systems, the original navigation systems in knowledge systems can be transformed into faceted navigation. The specific approach is shown in Figure 2 [Figure 2: see original paper]: First, domain concept terms are screened from the original navigation terms to exclude navigation terms with poor resource positioning effects and streamline navigation directions (as shown in Figure 2, terms b, f, and g are eliminated from the original navigation system); second, navigation facets are set based on domain concept types and fine-grained associations (as shown in Figure 2, three facets are established). Each navigation term represents a document collection, thereby aggregating and organizing the existing large number of document resources through faceted navigation. Based on this approach, the authors constructed a faceted navigation system using the cardiovascular forum in the DXY medical community as an example to achieve knowledge copolymerization of UGC resources, with the system prototype shown in Figure 3 [Figure 3: see original paper] [26].

4.2 From Word to Word: Knowledge Aggregation Based on Multidimensional Conceptual Association Recommendation

When users search for or browse documents related to a certain word, other words related to that word constitute part of their target knowledge. Users' knowledge-seeking and utilization behaviors are continuous processes that usually require constant horizontal expansion and vertical deepening. Query recommendation on search result pages and tag recommendation on content pages all belong to knowledge guidance "from word to word." Therefore, effective implementation of knowledge aggregation needs to fully consider this form.

Traditional word recommendation usually relies on corpora such as logs and documents to calculate and rank word similarity [27], then lists recommendations linearly. The semantic ambiguity makes users know that terms are related but not how they are related, which is not conducive to further expansion and refinement of user needs. Therefore, it is necessary to add semantic information to make it more consistent with users' knowledge structures. Lu Wei et al.'s experiments showed that topic analysis can significantly improve query recommendation accuracy [28]. Hong Jie et al. argued that in domain-specific knowledge systems, relying on semantic resource bases to build query recommendation systems is a good choice, and demonstrated through experiments that using domain ontologies can achieve better results than Baidu and Google query recommendations [29]. Jin Yan'an systematically discussed tag recommendation technologies based on semantic granularity, topic sensitivity, and user motivation [30]. In fact, the core goal of related word recommendation is to construct a space for subsequent changes in users' knowledge scenarios (such as vertical deepening or horizontal expansion). The connotation of a word is usually expressed through its conceptual relationships with other words, and words related to the current word are highly likely to be the targets for users' next need expansion or refinement. Therefore, multi-dimensionally aggregating related words based on conceptual associations to carry out recommendations is an important way to achieve knowledge copolymerization in the "from word to word" scenario.

As shown in Figure 4 [Figure 4: see original paper], after introducing domain conceptual association systems, concepts related to users' entry words and their fine-grained associations are introduced into word recommendation, forming multidimensional structured related word recommendations. On the one hand, this can exclude interference words unrelated to domain knowledge; on the other hand, by indicating the fine-grained association types between related words and entry words (as shown on the right side of Figure 4, word a has related words with different recommendation dimensions), users can selectively choose the direction for their next knowledge search. This enables effective aggregation of words in knowledge systems. Based on this approach, the authors constructed a multidimensional recommendation system based on conceptual associations using the cardiovascular forum in the DXY medical community as an example to achieve knowledge copolymerization of UGC resources, with the system prototype shown in Figure 5 [Figure 5: see original paper].

4.3 From Document to Word: Knowledge Aggregation Based on Knowledge Element Linking

Conceptual units with semantic connotations in documents can be called knowledge elements. The knowledge value of a document lies in organizing multiple knowledge elements through several associations. When users browse a specific document, they are usually interested in certain knowledge elements within it. Using knowledge elements in documents as carriers for related content aggregation is an important way to solve the "fragmented" distribution of knowledge in documents and achieve precise knowledge aggregation from massive text aggregation.

The knowledge element linking system is an important form for implementing knowledge copolymerization in the document browsing scenario. Zeng Jianxun believes that building knowledge networks based on knowledge element linking is an important direction for knowledge linking [31]. Sun Zhen et al. proposed a new scientific measurement paradigm based on knowledge elements, with its core being using professional problems and disciplinary knowledge points as measurement 口径 and basic knowledge units [32]. Bi Chongwu et al. argued that linking knowledge elements can aggregate the finest-grained knowledge elements into knowledge sets of different granularities, and proposed a multi-granularity knowledge set organization method based on knowledge elements [33]. Chen Guo et al. proposed a scheme for constructing a knowledge element linking system in online communities that integrates domain knowledge bases and co-occurrence analysis, achieving deep connectivity of fragmented knowledge in online communities through building knowledge element linking systems [34].

The implementation approach for knowledge copolymerization based on knowledge element linking is shown in Figure 6 [Figure 6: see original paper]. First, knowledge elements in documents are annotated with reference to concept terms in domain conceptual association systems, directing them to corresponding knowledge element content pages; second, knowledge element content pages are constructed, generally including the basic connotation of the knowledge element, other knowledge elements related to it, and document resources related to it. By constructing an independent knowledge element linking system, content aggregation of related knowledge elements can be achieved within documents without changing the original resource organizational architecture. Based on this approach, the authors constructed a knowledge element linking system using the cardiovascular forum in the DXY medical community as an example to achieve knowledge copolymerization [34], with examples of knowledge element annotation results in documents shown in Figure 7 [Figure 7: see original paper] and knowledge element content page examples shown in Figure 8 [Figure 8: see original paper].

4.4 From Document to Document: Knowledge Aggregation Based on Discovery of Potential Resource Associations

Documents are the ultimate carriers for users to acquire knowledge. From document to document is the most direct aggregation form in knowledge systems. Common text aggregation forms calculate similarity based on external attribute information or content annotations of documents, then perform text clustering and classification. The main problem with this aggregation method is excessively high result redundancy, while ignoring potential semantic associations between documents, thus having obvious defects in meeting task-oriented knowledge needs. For example, after browsing a document introducing "hypertension," users may need to understand content related to certain symptoms rather than more similar documents introducing "hypertension."

To solve this problem, researchers have conducted studies on discovering document resource associations from the semantic level. Chen Lanjie and Hou Pengjuan divided methods for revealing associations among digital literature resources into mining-based and construction-based methods, with the latter including semantic and ontology methods, linked data methods, etc., and pointed out their respective advantages and disadvantages [35]. Zhao Yiping and Bi Qiang proposed using semantic analysis and vector space models to calculate document content similarity, then merging semantic association information and document metadata into linked data to achieve subsequent similar literature discovery; they also noted that literature association discovery based on objective knowledge systems and knowledge structures should better reflect the relevance of related literature [36]. Hong Yunjia and Xu Xin proposed a multi-level text clustering method based on domain ontology suitable for tree structures in knowledge bases, achieving multi-level clustering from coarse to fine granularity [37].

In the knowledge copolymerization mode, relying on domain conceptual associations, the approach of knowledge discovery from non-related literature can be referenced to establish resource associations [38]. That is, associations are established between two relatively independent documents whose corresponding concepts have associations, further achieving document aggregation. The basic implementation method is shown in Figure 9 [Figure 9: see original paper]: If domain concept a_1 appears in document A and domain concept a_2 appears in document B, and there is a certain fine-grained relationship between a_1 and

a_2 in the domain knowledge system, then this association can be preliminarily annotated between A and B; the final association type between A and B can be established through statistics of more preliminary associations. In practice, the authors conducted document aggregation experiments based on association discovery in the cardiovascular forum of the DXY medical community. Taking "coronary angiography" as an example, only 61% of discussion posts were aggregated with "coronary heart disease" discussions through direct associations, while the remaining 39% required aggregation through the diagnostic relationship between coronary angiography and coronary heart disease based on the domain conceptual association system, demonstrating significant improvement in effectiveness [19].

5. Conclusion

Research on knowledge aggregation modes should not only focus on what knowledge associations to use for aggregation but also explore what forms knowledge units can be aggregated into using these associations. Given the lack of research on the latter, this paper refers to the two types of polymerization reactions in chemistry and divides knowledge aggregation modes into knowledge homopolymerization and knowledge copolymerization, with the distinction being whether the diversity of knowledge units and their associations is preserved in aggregation. From the perspective of domain knowledge analysis, the foundation for implementing knowledge copolymerization is the effective utilization of domain concepts and their fine-grained associations. In the "document-word" knowledge resource system, using documents and words as basic knowledge unit granularities and user demand portals and aggregation target resources as dimensions, knowledge copolymerization modes can be implemented through four basic forms: faceted navigation-based knowledge aggregation, multidimensional conceptual association recommendation-based knowledge aggregation, knowledge element linking-based knowledge aggregation, and potential resource association discovery-based knowledge aggregation.

Knowledge aggregation is simultaneously influenced by resource organizational structures and user demand structures. Domain knowledge composed of domain concepts and their fine-grained associations is the key to ensuring consistency between the two. Currently, research and practice on using specific domain knowledge backgrounds to conduct knowledge aggregation are relatively lacking, and knowledge aggregation for specific domains differs greatly in mode from general knowledge aggregation. As research and practice on knowledge aggregation gradually deepen, knowledge aggregation for specific domains will face new challenges. The knowledge copolymerization mode and its implementation forms proposed in this paper provide a basic approach that will be further improved in subsequent research and practice.

References

[1] THOMPSON D. Oxford Modern English-Chinese Dictionary [M]. Beijing: Foreign Language Teaching and Research Press, 2003: 38.
[2] CLAYDEN J, GREEVES N, WARREN S G. Organic chemistry [M]. Oxford, UK: Oxford University Press, 2001.
[3] Li Yating. Review of knowledge aggregation research [J]. Library and Information Service, 2016, 60(21): 128-136.
[4] Zhao Xueqin. Research status and future research suggestions on knowledge aggregation and service [J]. Information Theory and Practice, 2015, 38(2): 132-135.
[5] He Defang, Zeng Jianxun. Research on semantic-based deep aggregation of collection resources [J]. Journal of Library Science in China, 2012, 38(4): 79-87.
[6] Bi Qiang, Yin Changyu, Teng Guangqing, et al. Theoretical basis and methods of digital resource aggregation [J]. Journal of Library Science in China, 2016, 42(3): 4-19.
[7] Zhao Rongying, Wang Song, Dong Ke. Review of domestic collection resource aggregation mode research [J]. Library and Information Service, 2014, 58(18): 138-143.
[8] Cao Shujin, Ma Cichang. Research on the composition of information aggregation concept and aggregation modes [J]. Information Science, 2015(1): 9-14.
[9] Liu Ping, Chen Ye. Review of research progress on word similarity [J]. New Technology of Library and Information Service, 2012, 28(7): 82-89.
[10] Ma Wenfeng, Du Xiaoyong. Several theoretical issues on knowledge organization systems [J]. Journal of Library Science in China, 2007(2): 13-17.
[11] Teng Guangqing, Bi Qiang. Analysis of the evolution path of knowledge organization systems and development trends of related research [J]. Journal of Library Science in China, 2010(5): 49-53.
[12] Zhang Zhensen, Cheng Hao, Li Li. Research on theory and application of faceted classification systems for network information [J]. Journal of Intelligence, 2009, 28(8): 62-66.
[13] Tang Xiaobo, Xiao Lu. Research on microblog theme mining based on single-sentence granularity [J]. Journal of Intelligence, 2014, 33(6): 623-632.
[14] Jing Liping, Yun Jiali, Yu Jian. Opportunities and challenges of domain knowledge in text clustering applications [J]. Computer Engineering & Science, 2010, 32(6): 88-91.
[15] Wang Lin. Research on several issues in knowledge organization from the perspective of domain analysis paradigm [J]. Library and Information Service, 2011, 55(4): 90-105.
[16] HJØRLAND B, ALBRECHTSEN H. Toward a new horizon in information science: domain analysis [J]. Journal of the Association for Information Science and Technology, 1995, 46(6): 400-425.
[17] BUCKLAND M K, CHEN A, GEBBIE M, et al. Variation by subdomain in indexes to knowledge organization systems [C]// BEGHTOL C, HOWARTH L C, WILLIAMSON N J. Dynamism and stability in knowledge organization: proceedings of the 6th international ISKO conference. Wuerzburg: Ergon Ver-

lag, 2000: 48-53.
[18] VICKERY B C. Faceted classification: a guide to construction and use of special schemes [J]. Library quarterly information community policy, 1960(3): 289.
[19] Chen Guo. Research on network community knowledge aggregation based on domain conceptual association [D]. Wuhan: Wuhan University, 2015.
[20] State Bureau of Quality and Technical Supervision. Terminology work—Principles and methods: GB/T 10112-1999 [S]. Beijing: Standards Press of China, 1999: 1-6.
[21] Liang Kaiqiang, Lu Jukang. Association rule mining based on domain ontology and concept lattice [J]. Computer Engineering and Design, 2007, 28(13): 3033-3035.
[22] Bi Qiang, Zhou Shanshan, Ma Zhiqiang, et al. Research on tag cloud optimization mechanism for knowledge association [J]. New Technology of Library and Information Service, 2014(5): 33-40.
[23] Hu Yuan, Chen Lin, Ai Wenhua. Digital library community integrated push service organization based on knowledge aggregation [J]. Research on Library Science, 2017(19): 9-17.
[24] Zhang Yunzhong, Yang Meng. Tax-folk hybrid navigation: a new model for resource aggregation in social tagging systems [J]. Journal of Library Science in China, 2014(3): 78-89.
[25] HEARST M A. Clustering versus faceted categories for information exploration [J]. Communications of the ACM, 2006, 49(4): 59-61.
[26] Chen Guo, Xiao Lu, Sun Jianjun. Construction of faceted navigation system for online communities—taking the DXY cardiovascular forum as an example [J]. Information Theory and Practice, 2017, 40(10): 112-116.
[27] Wu Sizhu. Research progress on tag recommendation methods in social tagging systems [J]. Library Journal, 2010, 29(3): 48-52.
[28] Lu Wei, Zhang Xiaojuan. Research on query recommendation based on theme and user preference analysis [J]. Journal of the China Society for Scientific and Technical Information, 2012, 31(12): 1252-1258.
[29] Hong Jie, Zhang Jian, Hu Liang. Research on query recommendation algorithm for professional search engines based on domain ontology knowledge base—taking the salt lake chemical industry as an example [J]. Journal of the China Society for Scientific and Technical Information, 2014(10): 1091-1098.
[30] Jin Yan'an. Research on social tag recommendation technology and methods [D]. Wuhan: Huazhong University of Science and Technology, 2011.
[31] Zeng Jianxun. Research status and development trends of knowledge linking [J]. Information Theory and Practice, 2011(2): 119-123.
[32] Sun Zhen, Leng Fuhai. Analysis of new scientific measurement paradigm based on knowledge elements [J]. Journal of the China Society for Scientific and Technical Information, 2017, 36(6): 555-564.
[33] Bi Chongwu, Wang Zhongyi, Song Hongwen. Research on multi-granularity integrated knowledge service of digital libraries based on knowledge elements [J]. Library and Information Service, 2017, 61(4): 115-123.
[34] Chen Guo, Xiao Lu. Research on knowledge element linking system con-

struction in online communities [J]. Data Analysis and Knowledge Discovery, 2017, 1(11): 75-83.
[35] Chen Lanjie, Hou Pengjuan. Research on methods for revealing associations among digital literature resources [J]. Library, 2015(2): 41-45.
[36] Zhao Yiping, Bi Qiang. Research on application of linked data in similar literature discovery for academic resource networks [J]. New Technology of Library and Information Service, 2016, 32(3): 41-49.
[37] Hong Yunjia, Xu Xin. Research on multi-level text clustering for knowledge base based on domain ontology—taking Chinese culinary culture knowledge base as an example [J]. New Technology of Library and Information Service, 2014(12): 19-26.
[38] SWANSON D R. Undiscovered public knowledge [J]. Library quarterly, 1986, 56(2): 103-118.

Author Contributions

Chen Guo: Proposed the idea and completed the experiments.
Wu Wei: Wrote the initial draft.
Xiao Lu: Reviewed literature and revised the paper.

Knowledge Copolymerization: A Knowledge Aggregation Mode Under the Perspective of Domain Analytic Paradigm

Chen Guo[1], Wu Wei[1], Xiao Lu[2]
[1]Department of Information Management, School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094
[2]School of Journalism, Nanjing University of Finance and Economics, Nanjing 210046

Abstract: [Purpose/significance] Current research on knowledge aggregation mode focuses on "what types of knowledge connection that knowledge aggregation is based on." As an indispensable complementary, this paper explores the follow-up question of "what form knowledge elements can be aggregated based on those connections," it aims at perfect the research of knowledge aggregation mode and guide related practice. [Method/process] Taking the classification of polymerization in chemistry, this paper put forward to divide knowledge aggregation mode into homopolymerization and copolymerization according to "whether or not remain the differences between knowledge units and their correlations," and then explored the basic realization forms of knowledge copolymerization. [Result/conclusion] Domain knowledge is the basis of knowledge copolymerization. Utilizing documents and words as two basic knowledge elements, and considering users' demand portals and target knowledge elements as two dimensions, knowledge copolymerization can be achieved through the following four basic forms: based on faceted navigation, multidimensional recommendation based on conceptual relation, based on knowledge element linking, and knowledge discover based on the potential connection of resources.

Keywords: knowledge organization; knowledge aggregation mode; knowledge copolymerization; domain analytic paradigm; domain conceptual relation

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