ChinaRxiv

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202304.00439

Knowledge Metrology in Library and Information Science from the Perspective of Citation Content Analysis: A Preliminary Discussion (Postprint)

Authors: Bi Chongwu, Ye Guanghui, Xia Lixin

Date: 2023-04-01T16:02:58+00:00

Abstract

[Purpose/Significance] Currently, knowledge measurement research in the library and information science field has remained at the levels of knowledge carrier measurement and knowledge attribute measurement due to its inability to use knowledge units as the measurement unit. Proposing a research approach that achieves knowledge content measurement from the perspective of citation content analysis represents a worthwhile research direction to explore. [Method/Process] Based on reviewing the current status of knowledge measurement research across different fields, this study summarizes the core advantages and key issues of conducting knowledge measurement research in the library and information science field, proposes the main ideas and implementation paths for conducting such research from the citation content analysis perspective, and prospects the innovative applications and development directions of this research approach. [Result/Conclusion] Conducting knowledge measurement research from the citation content analysis perspective can break through the long-standing research bottlenecks in the library and information science field, truly extend the measurement object of knowledge from knowledge carriers to knowledge content, and achieve measurement of the quantity, quality, value, and relationships of knowledge units.

Full Text

Rethinking Knowledge Measurement Research in Library and Information Science from the Perspective of Citation Content Analysis

Bi Chongwu[1], **Ye Guanghui**[2], **Xia Lixin**[2] [1] School of Information Management, Zhengzhou University, Zhengzhou 450001 [2] School of Information Management, Central China Normal University, Wuhan 430079

Abstract: [Purpose/Significance] Current knowledge measurement research in library and information science remains constrained at the level of knowledge carriers and knowledge attributes because it cannot adopt knowledge elements as the unit of measurement. This paper proposes a research approach to achieve knowledge content measurement from the perspective of citation content analysis, representing a promising direction worth exploring. [Method/Process] Based on a review of knowledge measurement research across different fields, this study summarizes the core advantages and key challenges of conducting knowledge measurement research in library and information science, outlines the main ideas and implementation pathways for such research from a citation content analysis perspective, and envisions innovative applications and future development directions for this research approach. [Result/Conclusion] Knowledge measurement research from the citation content analysis perspective can break through long-standing bottlenecks in library and information science, truly penetrating the measurement unit from knowledge carriers to knowledge content, and enabling quantitative, qualitative, value-based, and relational measurement of knowledge elements.

Keywords: knowledge measurement; citation content; knowledge element; knowledge evaluation

Introduction

Knowledge metrology was first proposed as a new disciplinary concept by Professor Liu Zeyuan at the "International Symposium on Research Evaluation, Scientometrics, and Informetrics" in 1998 [1]. According to its definition, knowledge metrology is an interdisciplinary field that takes the entire human knowledge system as its object and employs object analysis and computational techniques to conduct comprehensive research on society's knowledge capabilities (including production, circulation, consumption, accumulation, and proliferation) and knowledge's social relationships (including organizational forms, collaborative networks, and social institutions) [2]. Influenced by both research perspectives and temporal factors, scholars from different disciplines and fields hold varying understandings of the connotation and extension of "knowledge measurement." In terms of research objects, existing knowledge measurement studies are concentrated in five domains: science of science, economics, management, computer science, and library and information science [3].

Scholars have long aspired to elevate knowledge measurement from indirect to direct approaches, proposing the smallest independent knowledge unit—the knowledge element (or knowledge unit). However, constrained by technological and methodological limitations of their time, they struggled to effectively separate knowledge elements from document units and information units. Consequently, although scholars theoretically articulated the concept of knowledge elements and highlighted their importance for knowledge measurement, they

could not achieve true knowledge content measurement. For a long time, knowledge element construction has constituted the fundamental bottleneck hindering in-depth knowledge measurement research in library and information science.

In recent years, the open access movement has made full-text data of literature resources readily available. Simultaneously, text structure identification technologies such as chapter segmentation and citation location extraction, along with text content mining technologies like topic extraction and citation strength calculation, have provided broader space and possibilities for analytical and measurement studies based on textual data from academic literature. In the past, scholars could only reveal ordinal relationships between knowledge through manual annotation, but with the development of knowledge graph technology, current research can not only automatically extract knowledge associations but also visually present the architecture, development 脉络, and frontier trends of domain knowledge using visualized graphs. These environmental changes and technological advancements have created new opportunities for knowledge measurement research in library and information science, rendering previous constraints obsolete and necessitating a rethinking of knowledge measurement approaches in the field.

1. Current Status of Knowledge Measurement Research in Library and Information Science

Knowledge is a complex concept, and its measurement presents a formidable challenge. Tracing the evolution of research objects, knowledge measurement research in library and information science has progressed through three stages: knowledge carrier measurement, knowledge attribute measurement, and knowledge content measurement. This dynamic evolution reflects scholars' arduous exploration and gradual deepening in conducting knowledge measurement research.

1.1 Knowledge Carrier Measurement Stage Taking knowledge carriers—document units—as the measurement unit, this stage benefited from increasingly sophisticated bibliometric indicators, methods, and evaluation systems. Scholars treated documents as knowledge and conducted quantitative studies on documents to indirectly measure knowledge's main characteristics. Representative and influential works during this period included F.J. Cole et al.'s [5] temporal distribution curves of publication volumes to compare research focuses and developmental changes in a discipline across different periods; P.L.K. Gross et al.'s [6] statistical analysis of journal paper references to identify core journals in a field; E. Garfield's [7] proposal of using citation indexes to retrieve scientific literature; P.O. Seglen's [8] discussion of limitations in applying impact factors to journal evaluation; H. Small's [9] use of co-citation concepts to measure the contribution of cited literature clusters to a field; S.C. Bradford's [10] discovery of the distribution pattern of papers in a field across journals; A.J. Lotka's [11] discovery of the quantitative relationship between authors and their produc-

tivity; J.E. Hirsch's [12] proposal of the h-index to quantify authors' research levels; and various laws of literature growth, aging, citation patterns, and Zipf's law.

Overall, the knowledge carrier measurement stage focused on the quantity and value characteristics of knowledge, measuring and evaluating materialized knowledge through its physical manifestations [13]. However, several problems existed: first, documents and knowledge are not completely equivalent—while documents contain knowledge, this approach cannot measure how much knowledge is contained within a document; second, it ignored the distinction between knowledge quality and knowledge value, equating knowledge quality solely with document value; third, it was relatively weak in measuring knowledge relationships.

1.2 Knowledge Attribute Measurement Stage Taking information units that reveal knowledge attributes as the measurement unit, this stage emerged alongside the development of information technology and new non-document resources for knowledge expression and transmission. Scholars used information to characterize knowledge, marking and expressing knowledge's internal and external features through information units and their combinations based on structured decomposition of document units. This stage produced numerous theoretical research outcomes on knowledge measurement. For instance, Wen Tingxiao et al. systematically reviewed research objects and content, basic units, disciplinary origins, and main methods of knowledge measurement [14], and elaborated on research dimensions and measurement units [15-16]; Yu Yisheng et al. [17] conducted in-depth discussions on the development history, measurement content, and main challenges of knowledge measurement research. These studies recognized the limitations of knowledge carrier measurement, pointed out the importance of knowledge content measurement, and proposed the idea of using knowledge elements as measurement units, but did not achieve knowledge content measurement implementation.

Overall, the knowledge attribute measurement stage served as a transitional or intermediate period. It neither overcame the inherent defects of knowledge carrier measurement nor achieved the ultimate goal of knowledge content measurement, suffering from several issues: first, information units are merely means to control and process knowledge carriers, lacking completely independent practical significance and typically cannot exist independently of knowledge carriers, still requiring knowledge carriers to locate knowledge content; second, this stage only explored knowledge clues, which could be considered as addressing knowledge relationships to some extent but could not precisely describe knowledge linkages.

1.3 Knowledge Content Measurement Stage Taking knowledge elements that describe knowledge content as the measurement unit, this stage has benefited from the widespread application of text mining technologies, particularly

knowledge identification and extraction technologies for literature resources. Three research areas are crucial for achieving knowledge content measurement. First, the in-depth development of knowledge element theory has enabled the academic community to develop a clearer understanding of knowledge elements, transforming them from a vague, virtual concept. Many studies have designed description rules and service models for knowledge elements, achieving knowledge element extraction and association to some extent [18-19]. Second, research on identification and segmentation of document structure and function has enabled locating knowledge content from literature resources, such as chapter segmentation [20], paragraph segmentation [21], and functional segmentation [22]. Third, research on identification and extraction of document content features has enabled separating knowledge content from literature resources, such as vocabulary identification [23], syntactic identification [24], and semantic identification [25].

Building on these foundations, with increasing data processing capabilities and growing availability of open literature resources, more research has begun using textual data from academic literature for measurement analysis, with citation content being widely employed. Citation content contains knowledge content extracted by citing authors from cited literature within citing documents. Through citation content analysis, the innovative relationship between citing and cited literature and the reasons for citation can be revealed. Therefore, conducting knowledge measurement research from the citation content analysis perspective can both penetrate knowledge elements within knowledge carriers and preserve the complex relationships between knowledge elements. Current citation content analysis research can be divided into four aspects: first, knowledge element and entity extraction, i.e., extracting the smallest representation units of knowledge and their associated entities from academic literature, using rule-based, statistical, and hybrid methods [24,26-27]; second, citation location and degree analysis, which extends traditional citation analysis by identifying the location and distribution of citation content in citing documents [28-29]; third, citation behavior and motivation analysis, which infers authors' true citation purposes based on description methods of citation content, incorporating sociological and psychological research methods [30-31]; fourth, citation sentiment and type identification, which uses semantic features of citation content to obtain citing authors' subjective sentiments toward cited literature and extends this to academic evaluation of cited literature [32-33].

Overall, existing research has made knowledge content measurement possible, and future development will inevitably focus on achieving accurate knowledge content measurement. However, currently only a few scholars recognize that existing research has created conditions for the emergence and development of a new knowledge measurement paradigm based on knowledge elements [34], some studies merely note the quiet rise of full-text-based bibliometric analysis [35-36], and others discuss "knowledge measurement" without achieving true knowledge content measurement. In fact, whether in the knowledge carrier measurement stage, knowledge attribute measurement stage, or knowledge content measure-

ment stage, scholars' original intention has always been to measure knowledge itself. Limited by cognitive levels and technological means, when precise direct measurement of knowledge itself was impossible, they could only conduct rough indirect measurement. Throughout this continuous exploration process, a series of extremely important measurement indicators and methods have been derived. Although these indicators and methods have not achieved direct measurement of knowledge itself, their research perspectives, design concepts, and application methods have laid an extremely important foundation for the in-depth development of knowledge measurement research.

2. Unique Characteristics of Knowledge Measurement Research in Library and Information Science

Knowledge measurement is a comprehensive, cross-disciplinary research challenge. Although scholars from different fields have conducted knowledge measurement research from various perspectives, their research purposes and significance differ, resulting in fragmented, scattered, and unsystematic outcomes. Moreover, some fields face insurmountable bottlenecks in knowledge measurement research. For example, in the science of science field, although scientific and technological achievements are recognized by authorized institutions, avoiding issues of research object scope, the transformation and evaluation of different types of achievements (basic research, applied research, and developmental work) greatly constrain knowledge measurement research. In management and economics fields, whether from knowledge management or knowledge economy perspectives, the greatest difficulty lies in measuring the flow of tacit knowledge. In computer science, existing research focuses on knowledge acquisition and application but lacks relatively mature knowledge measurement research paradigms. In contrast, library and information science possesses unique advantages for conducting knowledge measurement research.

2.1 Knowledge Measurement Research in Other Fields The proposal of the knowledge economy era and the rise of knowledge economics have led scholars to recognize the necessity of combining research methods and paradigms from scientometrics, bibliometrics, informetrics, and econometrics to conduct extensive interdisciplinary research on knowledge production and application, input and output, stock and flow, distribution and transfer, value and price from both macro and micro perspectives. Therefore, besides library and information science scholars, scholars from other fields have also formed relatively independent theoretical and methodological systems to address knowledge measurement and valuation issues according to their respective needs [15].

In the science of science field, scholars aim to measure, analyze, evaluate, and manage scientific and technological achievements, exploring the quantitative characteristics, developmental patterns, and structural relationships of knowledge within these achievements to reveal the laws of scientific and technological development and scientific research activities. In this process, scholars take

scientific and technological achievements as the main research object and extend to researchers, research institutions, regions, or countries related to these achievements.

In the management field, scholars examine knowledge as an input resource to investigate its output, efficiency, and effectiveness, making significant progress in knowledge asset evaluation. Particularly in knowledge classification, management scholars divide knowledge into codifiable explicit knowledge and non-codifiable tacit knowledge. The measurement of tacit knowledge and the transformation from tacit to explicit knowledge have become key issues in management knowledge measurement research.

In the economics field, scholars measure knowledge's impact on socio-economic development from the perspective of its economic value, focusing on applying knowledge economics theories and methods to measure knowledge contribution rates in the knowledge economy and knowledge industries. They treat knowledge as an independent production factor incorporated into enterprise or social production processes, examining this new economic growth model to measure knowledge input, stock, flow, output, etc., in economic entities.

In the computer science field, scholars have conducted extensive and in-depth research on explicit knowledge expression, understanding, processing, and storage, achieving major breakthroughs in knowledge mining, knowledge discovery, knowledge graphs, knowledge warehouses, and knowledge systems. They have particularly enabled knowledge extraction, association, and service for text, image, audio, video, and other resources to some extent, providing reliable technical means and methodological tools for knowledge measurement research on these resources.

2.2 Core Advantages of Knowledge Measurement Research in Library and Information Science Although knowledge measurement research in library and information science has experienced three different developmental stages, scholars have consistently focused on literature as the knowledge carrier. The only difference lies in the gradual evolution of knowledge measurement units. Whether using document units, information units, or knowledge elements as basic measurement units, they all represent valuable explicit knowledge recorded on carriers through certain methods and meaning expression systems. This precisely avoids tacit knowledge that is difficult to quantify directly.

Second, knowledge element research has been applied in practice. As early as the 1970s, experts in library and information science pointed out that the control unit of knowledge needed to advance from document units to knowledge elements, hoping to directly organize and analyze the logical content of documents to identify connection nodes in knowledge creation processes and deeply reveal the organic structure of knowledge [37]. Over decades, scholars have conducted extensive exploration and research from theoretical, methodological, technical model, and application platform perspectives, achieving remarkable

results in knowledge element theory and review, knowledge element representation and modeling, and knowledge element extraction and implementation. In recent years, with the development of cognitive theory and natural language processing capabilities, library and information science has basically achieved automatic extraction of full-text knowledge elements from literature resources, truly enabling organization and service targeting knowledge content.

Third, the field possesses relatively mature research paradigms. Library and information science has a long history of using metrological methods to conduct quantitative research on knowledge carriers, knowledge content, knowledge activities, and their impacts. After decades of development, scholars have gradually shifted their research perspective from knowledge carrier-based measurement to knowledge content-based measurement. Meanwhile, bibliometrics, informetrics, and webometrics—closely related metrological disciplines—have reached maturity in library and information science, providing solid foundations in techniques, tools, and methods for knowledge measurement research. Knowledge measurement research represents the extension, deepening, and development of these previous metrological studies.

2.3 Key Issues in Knowledge Measurement Research in Library and Information Science The most critical issue in early knowledge measurement research in library and information science was how to obtain the smallest, independent knowledge measurement unit from literature. Although scholars conceived the idea of using knowledge elements as measurement units long ago, unclear conceptual definitions of knowledge elements, particularly the difficulty of knowledge element construction, hindered the in-depth development of knowledge measurement research in library and information science for a long time. Existing research only focused on knowledge carriers and attribute features without achieving direct measurement of knowledge content.

To some extent, past knowledge measurement research in library and information science merely borrowed research methods and paradigms from bibliometrics, informetrics, and scientometrics. Beyond its forward-thinking research ideas and perspectives, its research content did not differ from existing metrological studies. Currently, the core issue remains knowledge element construction, with the prerequisite being clear identification of what knowledge elements are in literature and how to extract them. Additionally, the following key questions must be considered:

(1) **Knowledge element quantity measurement**: Since literature is a resource with historical significance or research value that necessarily contains a certain number of knowledge elements, how should the number of knowledge elements in a document be calculated? How are knowledge elements distributed within a document? How do knowledge element quantities differ between documents?

(2) **Knowledge element quality measurement**: As the smallest, indepen-

dent knowledge product, a knowledge element possesses characteristics and features that satisfy specified and potential needs. How should the actual quality of a knowledge element be measured? Do different knowledge elements within the same document have the same knowledge quality? What is the relationship between knowledge quality and knowledge value? In previous research, some scholars confused knowledge quality with knowledge impact, using citation counts of knowledge carriers in citation networks to measure knowledge quality. In fact, citation counts measure the usage frequency of knowledge carriers as knowledge products, representing a form of use value.

(3) **Knowledge element value measurement**: According to political economics definitions of use value and value, it is difficult to measure a knowledge element's value or accurately determine its use value, but we can use usage frequency to compare the use value of different knowledge elements. This raises important research questions: How can knowledge element usage frequency be obtained? Can knowledge elements only be used completely?

(4) **Knowledge element relationship measurement**: Complex logical connections exist between knowledge elements, necessitating differentiation of connection types in knowledge measurement. Within the same document, connection relationships between knowledge elements affect quantity measurement results; across different documents, they affect value measurement results. Additionally, calculation methods and scope of knowledge connections should differ between knowledge elements within the same document and those across different documents.

3. Main Ideas for Conducting Knowledge Measurement Research from the Citation Content Analysis Perspective

Beyond clear research objects, appropriate technical means, and mature research paradigms, citation networks provide a research platform supporting big data analysis for library and information science scholars. Previous knowledge measurement research using citation networks has been common because citation networks align with research objects and paradigms in library and information science. However, constrained by knowledge element construction difficulties, past research only macroscopically and superficially measured knowledge carriers or attributes without penetrating knowledge content. Therefore, this paper proposes a new idea for knowledge measurement research in library and information science from the citation content analysis perspective, using full-text mining technologies for literature resources, as shown in Figure 1 [Figure 1: see original paper].

Literature contains knowledge elements expressing viewpoints, definitions, theories, and methods in semi-structured or unstructured forms. First, knowledge description rules must be used to locate knowledge elements' specific positions in

documents. Then, based on a designed description framework, text content at corresponding positions undergoes knowledge extraction and storage. Finally, based on citation relationships between documents, knowledge elements at citation positions in citing documents are associated with knowledge elements in cited fragments of cited documents, constructing a knowledge network based on knowledge elements and their connections.

Currently, knowledge extraction follows two technical routes: statistics-based and rule-based. Since knowledge in literature typically consists of complex sentences, mathematical statistics and machine learning methods are often inadequate, making rule-based identification and pattern matching more suitable. Given current cognitive theory development and natural language processing capabilities, this paper designs knowledge element quantity measurement steps as shown in Figure 2 [Figure 2: see original paper].

First, knowledge element characterization and description must be addressed. To obtain high-quality knowledge element description rules from training texts that lay the foundation for automated knowledge element extraction, a semi-automated approach combining machine and human methods can be adopted. During this process, training corpora should be as comprehensive and representative as possible, with complete sentence segmentation, accurate syntactic structure and grammatical components, and comprehensive domain knowledge thesauri.

Second, knowledge element identification and extraction must be resolved. Through natural language processing including sentence segmentation, word segmentation, and keyword filtering of target documents, text fragments potentially containing knowledge elements are identified and extracted. Based on generated knowledge element description rules, text fragments' dependency syntax sequences are matched with description rule syntax sequences to discover knowledge elements hidden in semi-structured or unstructured text fragments.

Finally, knowledge elements' quantitative characteristics must be measured and analyzed from multiple dimensions to comprehensively reveal the total quantity, flow, growth, distribution, and structure of knowledge elements across different knowledge subjects.

3.1 Ideas for Measuring Knowledge Element Quantity in Documents

Extracting knowledge elements from documents is the foundation for calculating knowledge quantity. Based on knowledge networks, knowledge measurement research can be conducted from the citation content analysis perspective: (1) the quantity of knowledge elements in documents determines knowledge stock in knowledge nodes; (2) knowledge elements in cited fragments determine the knowledge content provided by cited literature; (3) knowledge elements at citation positions determine the knowledge content received by citing literature; (4) by comparing differences between knowledge elements in cited fragments and

citation positions, citation strength between knowledge nodes can be characterized, with citation strength determining knowledge flow between nodes; (5) by comprehensively analyzing citation position, sentiment, and motivation, citation types between knowledge nodes can be defined, with citation types determining specific connections implied between knowledge elements.

3.2 Ideas for Measuring Knowledge Element Quality in Documents
Following the classic assertion that quality lies in meeting and exceeding expectations, knowledge quality is actually the degree to which knowledge evaluators perceive knowledge characteristics as conforming to their expectations [37]. Unlike ordinary products, knowledge cannot be represented by direct quantitative parameters, but its quality can be indirectly reflected through quality characteristics or features. Therefore, knowledge quality evaluation should largely be subjective, deriving from differences in evaluation dimensions employed by different evaluators. As evaluation objects become increasingly abstract in product attributes from document units to information units to knowledge elements, quality connotations become more complex, requiring richer evaluation dimensions. Accordingly, this paper adopts a multi-dimensional evaluation standard at the theoretical level to comprehensively reveal a picture compatible with the complex concept of knowledge quality, as shown in Figure 3 [Figure 3: see original paper]. The advantage of multi-dimensional decomposition is that it transforms the abstract concept of knowledge quality into concrete terms while reducing misjudgment risks.

At the practical level, although different evaluation models may differ in evaluation dimensions, overall patterns should exist. From the information dimension, signals interpreting knowledge quality features can be input—such as author, institution, journal, and project—to measure knowledge element quality from knowledge attribute aspects. From the process dimension, texts revealing knowledge production processes are input to evaluate whether each production 环节 is reasonable and whether connections between 环节 are correct. From the ontology dimension, frameworks showing knowledge's intrinsic and contextual constructs are input to evaluate logical relationships between knowledge elements and existing knowledge systems. From the application dimension, empirical processes and reuse conditions are input to evaluate knowledge elements' practical effectiveness in solving real problems. On this basis, more specific evaluation indicators can be established under each dimension, and expert scoring for each indicator can be conducted using peer review methods to achieve quantitative calculation and normalization of knowledge element quality features.

Additionally, machine learning algorithms can enable rapid assessment of knowledge element quality. Text mining technologies can directly extract quality features x corresponding to each evaluation indicator from input data. For example, assuming a regression relationship between knowledge element quality features x and prediction results p, model training can first optimize actual parameters for each evaluation indicator in the assessment model, then select

optimization algorithms to continuously minimize error values between prediction results p and expert scores y, ultimately achieving convergence between peer review and machine learning results.

3.3 Ideas for Measuring Knowledge Element Value in Documents Knowledge value lies in its flow, depending not only on knowledge quality itself but more crucially on the depth and breadth of knowledge flow. Knowledge elements in literature do not exist in isolation. Typically, a document builds upon multiple documents through citation behaviors between knowledge elements. This objective fact aligns with both the inherent development 规律 of science and the inheritability and variability between knowledge elements in research activities. Generally, cited literature serves as the research foundation for citing literature, with cited fragments guiding citing literature. Therefore, knowledge elements contained in cited fragments can be considered as flowing to corresponding citation positions in citing literature. Accordingly, this paper designs a research 思路 for evaluating knowledge element value based on knowledge element flow conditions, as shown in Figure 4 [Figure 4: see original paper].

In the knowledge flow measurement module, based on citation relationships between citing and cited documents and using knowledge elements as measurement units, knowledge element flow in single triples is measured. First, citation positions must be identified in citing documents to extract citation content. Then, citation content is compared with cited fragments to extract knowledge elements from both citing and cited documents based on knowledge element description rules. Finally, similarity between knowledge elements is calculated to measure knowledge flow in a single citation process. When calculating similarity, subtopics, syntactic structures, and text content between knowledge elements must be considered, with weights designed through feature fusion to comprehensively calculate final values.

In the knowledge value measurement module, based on citation relationships provided by citation databases and using knowledge element flow conditions in citation networks as evaluation criteria, knowledge element value is assessed. First, flow relationships between knowledge elements are extracted from citation databases—i.e., knowledge elements flowing from cited to citing documents. Second, heterogeneous, weighted networks with knowledge elements as network nodes are constructed. Then, knowledge flow between knowledge elements is incorporated with external data revealing relationship types to comprehensively analyze knowledge elements' positions and roles in knowledge networks. Finally, content similarity algorithms and role importance algorithms are integrated to rank the value of interdependent and mutually influential knowledge elements.

3.4 Ideas for Measuring Relationship Types Between Knowledge Elements Various connections exist between knowledge elements, and revealing and utilizing these connections is the starting point for knowledge organization, management, and service. Relationship types between knowledge elements are

essentially the kinds of associations hidden between them that can be understood and ordered. For knowledge elements directly connected through citation relationships or knowledge attributes, their relationship types are obvious; however, relationship types between indirectly connected knowledge elements can also be mined through intermediary knowledge elements. Considering the origins of knowledge elements, relationship types must address both connections within the same document and across different documents. Accordingly, this paper designs a research 思路 for measuring relationships between knowledge elements as shown in Figure 5 [Figure 5: see original paper].

Within the same document, based on knowledge element classification, knowledge elements must first be divided into different types. According to knowledge description frameworks, they are structurally stored. Then, considering expert opinions and training results, reasonable knowledge relationship categories and classification standards are established to guide extraction of relationship features between knowledge elements. Finally, association vectors are used to store different association features between knowledge elements, with semantic similarity calculated separately for each association vector, ultimately integrating all association vector calculation results to comprehensively judge relationship types between knowledge elements.

Across different documents, relationship type classification should primarily rely on citation content analysis results. Citation relationships establish the most direct knowledge associations between knowledge elements, but these associations imply complex relationship types. Although current technology can correspond citation content in citing documents with cited fragments in cited documents, in practice it is difficult to unify citation behavior's initial motivation with final outcomes. Therefore, comprehensive analysis of corresponding knowledge elements in citation content and cited fragments is required, considering calculation results of citation position, sentiment, purpose, and strength to identify relationship types between knowledge elements from multiple angles, perspectives, and levels.

Overall, these two aspects have different emphases. For knowledge elements within the same document, since they share the same knowledge producer, cross-relationships, oppositional relationships, and parallel relationships should be primarily considered. For knowledge elements across different documents, due to knowledge "inheritance" and "variation," equivalence relationships, genus-species relationships, and oppositional relationships should be mainly considered. Additionally, the above research 思路 only involves binary relationships between two knowledge elements, yet association relationships between knowledge elements also exhibit transitivity. Considering ternary relationships from a structural perspective allows knowledge networks to be decomposed into analyzable micro-units and reveals richer structural features [39].

4. Innovative Applications of Knowledge Measurement Research from the Citation Content Analysis Perspective

Knowledge measurement is an important foundation for all knowledge activities, with broad applications in knowledge production, acquisition, organization, and presentation. The above considerations demonstrate that with current cognitive theory development and natural language processing capabilities, and based on existing research methods and theories of the "Five Metrics," knowledge measurement research from the citation content analysis perspective is fundamentally achievable. This new research perspective will bring even more 引人注目的 application scenarios, sufficient to promote the deepening and expansion of future knowledge measurement research. Examining domestic and international research trends, future knowledge measurement applications will concentrate in three main areas.

4.1 Application Directions in Knowledge Management Many information scientists have proposed that knowledge is the research object of library and information science, but this assertion did not receive sufficient academic attention for a long time [40]. Reviewing the research history of library and information science reveals that its workflow has always progressed along the information chain of "facts—data—information—knowledge—intelligence," with current research focusing on the 环节 of processing information into knowledge [41]. The research 思路 in this paper can transition knowledge management's control unit from document units and information units to knowledge elements, thereby elevating knowledge management from indirect to direct approaches and enabling effective identification, processing, and organization of knowledge content.

For knowledge management, achieving measurement research with knowledge elements as the basic unit will significantly adjust its current application scope: (1) it can serve knowledge quantity management by measuring the total and incremental knowledge of any knowledge owner, compensating for existing research that can only measure literature totals and growth using exponential growth models and information totals and growth using information entropy and Moore's law, but lacks quantitative methods for knowledge totals and growth beyond the difficult-to-calculate Brookes' knowledge equation; (2) it can serve knowledge quality management by evaluating academic knowledge quality and conducting knowledge quality detection and control, rather than evaluating research outcomes based on knowledge carriers like journal papers; (3) it can serve knowledge value management by evaluating knowledge input and output based on the smallest, independent knowledge units, calculating the use value, social impact, and contribution to scientific and technological progress of knowledge and related products; (4) it can serve knowledge relationship management by expanding single citation relationships to complex knowledge associations, revealing hidden, understandable, and usable internal logic between knowledge elements.

4.2 Application Directions in Scientific Evaluation Scientific evaluation in library and information science originated from bibliometrics and citation analysis that emerged in the mid-20th century. By employing mathematical and statistical analysis methods using various characteristic quantities of literature resources, it can describe, evaluate, and predict the current status and development trends of science and technology, and conduct quantitative assessments of entities involved in scientific activities. Throughout this long research history, many scholars have pointed out that existing theories, methods, and indicators, while playing an irreplaceable positive role in today's scientific evaluation, still have limitations and cannot be applied wholesale. They should be analyzed, used reasonably, effectively referenced, or even modified according to specific circumstances to make evaluation results more objective, reasonable, and scientific [42].

For scientific evaluation, measurement research with knowledge elements as the basic unit can be applied to all fields related to scientific activities. Academic literature contains rich knowledge elements. Using knowledge elements as evaluation objects and combining features such as citation position, strength, function, sentiment, and motivation enables truly journal-independent academic evaluation activities, fundamentally curbing undesirable tendencies in academic evaluation such as "paper-only" and "SCI-supremacy" that judge papers by their journals. Specific applications include: (1) integrating knowledge element measurement results into research outcome evaluation, including assessing the status and role of books, journals, papers, and patents; (2) integrating knowledge element measurement results into researcher evaluation, including qualitative and quantitative assessment of academic influence and research performance; (3) integrating knowledge element measurement results into research field evaluation, assessing a discipline or research field's development stage, status, level, prospects, and the correlation between knowledge structure and content; (4) integrating knowledge element measurement results into research institution evaluation, measuring the quantity, quality, effectiveness, and application of research outputs at various levels from cities and regions to countries, departments, and schools.

4.3 Application Directions in Knowledge Service After years of theoretical discussion, library and information science still struggles to conduct efficient and precise knowledge services, with the confrontation between "knowledge thirst" and "information overload" persisting. From the knowledge resource perspective, although knowledge is mostly distributed in literature resources, its current organization still primarily uses coarse-grained document units as basic units, unable to match users' multi-granular, hierarchical knowledge needs. Meanwhile, knowledge describing the same research object or problem is typically distributed across different disciplines with certain disciplinary barriers, preventing the revelation of associations between different knowledge elements [43]. These phenomena hinder the provision of superior knowledge service experiences for end users.

For knowledge service, measurement research with knowledge elements as the basic unit can clarify knowledge reserves and their system structures, selecting measurement results that match users' constraint conditions according to actual needs to achieve deeper specialized knowledge services: (1) it enables fine-grained knowledge service, transitioning from knowledge carrier-based (e.g., books, journals, patents) and information-based (e.g., topics, data, keywords) service methods to knowledge content-based service methods—for example, by mining, extracting, and analyzing knowledge elements in literature resources to provide users with minimal units of human knowledge such as specific examples, values, concepts, and laws; (2) it enables combinatorial knowledge service, transitioning from literature-based, discipline-based knowledge service methods to cross-disciplinary, knowledge content-based methods—for example, using knowledge relationship measurement results to construct objective knowledge systems across disciplines and fields, providing users with knowledge assembly solutions integrating multi-level concepts, solutions, and operational processes.

5. Summary and Outlook

Since the 1990s, to embrace the knowledge economy era, knowledge has been regarded as a strategic and economic resource in social competition, receiving unprecedented attention from all sectors of society. In recent years, although research in scientometrics, bibliometrics, and informetrics has involved measurement of knowledge carriers or attributes within their respective fields, differing research purposes and significance have prevented these studies from conducting systematic research at the height of knowledge science that penetrates knowledge content, making it difficult to convincingly reveal and clarify knowledge's important role in modern social development. This paper envisions a feasible path for knowledge content measurement in library and information science from the citation content analysis perspective, discusses foundational issues and solutions for knowledge content measurement from four aspects (quantity, quality, value, and relationship), and finally envisions application scenarios after achieving knowledge content measurement.

From the current perspective, knowledge measurement remains a research direction between multiple disciplines with interdisciplinary characteristics. This requires that when conducting knowledge measurement research, library and information science must not only borrow theories, methods, and technologies from other disciplines but also integrate them into the field to elevate them to a new level, deepening a more complete research system and paradigm that conforms to internal logic. Therefore, narrowing disciplinary gaps, removing barriers, and synthesizing and integrating ongoing knowledge measurement research from different levels and perspectives across various fields is particularly important.

For example, science of science's research performance evaluation and management methods, management's explicit and tacit knowledge management theories, economics' macro and micro economic value perspectives, computer sci-

ence's knowledge mining and storage technologies, and knowledge concepts from education and information science all have significant reference value for knowledge metrology research in library and information science.

More importantly, library and information science must generate new theories and methods that other disciplines do not possess or do not emphasize. Throughout history, new theories, methods, and technologies have profoundly impacted library and information science. However, at all times, the field should act according to its actual conditions and characteristics, doing something in some areas and refraining in others, rather than blindly pursuing comprehensiveness, otherwise it will lose its own characteristics. In summary, in this era of scientific research intersection and integration, in an environment with both opportunities and challenges, library and information science scholars should find their positioning in inheritance and innovation, 坚守 and expansion, and conduct knowledge measurement research with library and information science characteristics.

References

[1] Liu Zeyuan. Zhao Hongzhou and Chinese Scientometrics[J]. Studies in Science of Science, 1999, 17(4): 104-109. [2] Liu Zeyuan, Liu Fengzhao. Methodological Reflections on Knowledge Metrology Research[J]. Science of Science and Management of S&T, 2002, 22(8): 5-8. [3] Qiu Junping, Wen Tingxiao, Song Yanhui, et al. Knowledge Metrology[M]. Beijing: Science Press, 2014. [4] Ma Feicheng. Progress and Deepening of Information Science[J]. Journal of the China Society for Scientific and Technical Information, 1996, 15(5): 338-344. [5] Cole FJ, Hulme EW. The history of comparative anatomy—a statistical analysis of the literature[J]. Science progress, 1917 (11): 578-596. [6] Gross PLK, Gross EM. College libraries and chemical education[J]. Science, 1927, 66(1713): 385-389. [7] Garfield E. Citation indexes for science: a new dimension in documentation through association of ideas[J]. Science, 1955, 122(3159): 108-111. [8] Seglen PO. Why the impact factor of journals should not be used for evaluating research[J]. British medical journal, 1997, 314 (7079): 497-502. [9] Small H. Co-citation in the scientific literature: A new measure of the relationship between two documents[J]. Journal of the American Society for Information Science and Technology, 1973, 24 (4): 265-269. [10] Bradford SC. Sources of information on specific subjects[J]. Engineering, 1934, 137: 85-86. [11] Lotka AJ. The frequency distribution of scientific productivity[J]. Journal of the washington academy of sciences, 1926, 16 (12): 317-323. [12] Hirsch JE. An index to quantify an individual's scientific research output[J]. Proceedings of the National Academy of Sciences of the USA, 2005, 102(46): 16569-16572. [13] Wen Tingxiao. Research on Knowledge Measurement and Knowledge Evaluation[J]. Journal of the China Society for Scientific and Technical Information, 2007, 26(5): 670-676. [14] Wen Tingxiao, Liu Ying, Liang Xiujuan, et al. Review of Knowledge Measurement Research[J]. Library and Information Service, 2010, 26(1): 95-101. [15] Wen Tingxiao, Liu Xuan. On the Dimensions of

Knowledge Measurement Research[J]. Library and Information Science, 2013, 29(3): 103-109. [16] Wen Tingxiao. Comparative and Evaluative Research on Knowledge Measurement Units[J]. Information Studies: Theory & Application, 2007, 30(6): 731-736. [17] Yu Yisheng, Zhang Yang. Research on Knowledge Measurement and Evaluation[J]. Library and Information Service, 2008, 52(11): 18-21. [18] Wang Zhongyi, Huang Rong, Zheng Xin, et al. Research on Multi-granularity Integrated Knowledge Service in Digital Libraries[J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(2): 143-158. [19] Bi Chongwu, Wang Zhongyi, Song Hongwen. Research on Multi-granularity Integrated Knowledge Service in Digital Libraries Based on Knowledge Elements[J]. Library and Information Service, 2017, 61(4): 115-122. [20] Huang Yong, Lu Wei, Cheng Qikai. Structural Function Identification of Academic Text—Identification Based on Chapters[J]. Journal of the China Society for Scientific and Technical Information, 2016, 35(3): 293-300. [21] Huang Yong, Lu Wei, Cheng Qikai, et al. Structural Function Identification of Academic Text—Identification Based on Paragraphs[J]. Journal of the China Society for Scientific and Technical Information, 2016, 35(5): 530-538. [22] Lu Wei, Huang Yong, Cheng Qikai. Structural Function Identification of Academic Text—Functional Framework and Identification Based on Chapter Titles[J]. Journal of the China Society for Scientific and Technical Information, 2014, 33(9): 979-986. [23] Lu Wei, Li Pengcheng, Zhang Guobiao, et al. Academic Text Vocabulary Function Identification—Research on Automatic Keyword Classification Based on BERT Vector Representation[J]. Journal of the China Society for Scientific and Technical Information, 2020, 39(12): 1320-1329. [24] Hua Bolin. Research on Types and Description Rules of Method Knowledge Elements in Academic Papers[J]. Journal of Library Science in China, 2016, 42(1): 30-40. [25] Wen Hao. Research on Semantic Recognition and Classification Methods for Innovation Points in Scientific Abstracts[J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(3): 249-256. [26] Ying D, Min S, Han J, et al. Entity metrics: measuring the impact of entities[J]. Plos one, 2013, 8(8): e71416. [27] 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. [28] Hu ZG, Chen CM, Liu ZY. Where are citations located in the body of scientific articles? a study of the distributions of citation locations[J]. Journal of informetrics, 2013, 7(4): 887-896. [29] Hu ZG, Ling G, Sun TA, et al. Understanding multiply mentioned references[J]. Journal of informetrics, 2017, 11(4): 948-958. [30] Willett P. Readers' perceptions of authors' citation behaviour[J]. Journal of documentation, 2013, 69(1): 145-156. [31] Bornmann L, Daniel H. What do citation counts measure? a review of studies on citing behaviour[J]. Journal of documentation, 2008, 64(1): 45-80. [32] Catalini C, Lacetera N, Oettl A. The incidence and role of negative citations in science[J]. Proceedings of the National Academy of Sciences, 2015, 112(45): 13823-13826. [33] Sula CA, Miller M. Citations, contexts, and humanistic discourse: toward automatic extraction and classification[J]. Literary and linguistic computing, 2014, 29(3): 452-464. [34] Sun Zhen, Leng Fuhai. Analysis of a New Scientometric Paradigm Based on Knowledge Elements[J].

Journal of the China Society for Scientific and Technical Information, 2017, 36(6): 555-564. [35] Lu Chao, Zhang Chengzhi, Wang Yuzhuo, et al. Deepening of Semantic Feature Analysis—A Review of Full-text-based Bibliometric Analysis[J]. Journal of Library Science in China, 2021, 47(2): 110-131. [36] Hu Zhigang, Zhang Chengzhi. The Quiet Rise of Full-text-based Bibliometric Analysis[J]. Library Tribune, 2021, 41(3): 1-11. [37] Wen Youkui. Knowledge Element Mining[M]. Xi'an: Xidian University Press, 2004. [38] Zhou Bo, Zhang Fang. Knowledge Quality Evaluation Model in Knowledge Market Context[J]. Journal of Management, 2013, 10(10): 1535-1544. [39] Dong Ke, Liu Dehong, Jiang Hong. Structural Analysis of Citation Networks Based on Triple Relationships[J]. Information Studies: Theory & Application, 2010, 33(11): 50-53. [40] Qiu Junping, Duan Yufeng. On the Transformation of Library and Information Science by Knowledge Management[J]. Journal of Library Science in China, 2003, 29(2): 4-8. [41] Ma Feicheng, Li Zhiyuan. Development Prospects of Library and Information Science in China Under the Background of New Liberal Arts[J]. Journal of Library Science in China, 2020, 46(6): 4-15. [42] Su Xinning. Reflections on Issues in Bibliometrics and Scientific Evaluation[J]. Library & Information, 2013, 32(1): 79-83. [43] Wang Zhongyi, Xia Lixin, Li Yuhai. Construction of a Cross-disciplinary Multi-granularity Knowledge Representation Model for Digital Libraries Based on Knowledge Content[J]. Journal of Library Science in China, 2019, 45(6): 50-64.

Author Contributions: Bi Chongwu: Conceptualization, framework design, writing and revision; Ye Guanghui: Participated in framework design, provided ideas; Xia Lixin: Provided revision suggestions, approved the manuscript.

Rethinking Knowledge Measurement in Library and Information Science from the Perspective of Citation Content Analysis

Bi Chongwu[1], Ye Guanghui[2], Xia Lixin[2] [1] School of Information Management, Zhengzhou University, Zhengzhou 450001 [2] School of Information Management, Central China Normal University, Wuhan 430079

Abstract: [Purpose/Significance] The research of knowledge measurement in Library and Information Science cannot take the knowledge element as measurement unit, and always stays at the level of knowledge carrier measurement and knowledge attribute measurement. This paper puts forward a research idea of knowledge content measurement from the perspective of citation content analysis, which is a research direction worth exploring. [Method/Process] On the basis of combing the research status of knowledge measurement in different fields, this paper summarizes the core advantages and key problems of knowledge measurement research in Library and Information Science, puts forward the main ideas and implementation path of knowledge measurement research from the perspective of citation content analysis, and looks forward to the innovative application and development direction of this research idea. [Result/Conclusion] The research of knowledge measurement from the perspective of citation content analysis can break through the long-standing research bottleneck in Library and Information Science, truly penetrate the measurement object from knowl-

edge carrier to knowledge content, and realize the quantity, quality, value and relationship measurement of knowledge elements.

Keywords: knowledge measurement; citation content; knowledge element; knowledge evaluation

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

Source: ChinaXiv — Machine translation. Verify with original.