

An Analysis of Research Issues in Chinese Academic Papers: An Empirical Study of CSSCI-Indexed Publications in Library and Information Science, 2015-2020 (Postprint)

Authors: Suo Chuanjun, Li Muzi

Date: 2023-04-01T16:02:56+00:00

Abstract

[目的/意义] Academic research originates from questions; exploring the research questions in academic papers provides a reference for revealing the core content of academic papers and enabling question-based knowledge discovery in academic literature. [方法/过程] With papers published in CSSCI journals in the field of library and information science between 2015-2020 as the data source, 1,000 papers were randomly selected as samples. Employing content analysis, bibliographic information was statistically analyzed to examine the types and structures of research questions in academic papers, summarize the relationships between paper types and the types and quantities of research questions they contain, as well as the relationships and structural patterns among multiple research questions within a single academic paper. [结果/结论] The study reveals that different types of academic papers exhibit certain regularities in the types and quantities of research questions they contain; an academic paper typically contains more than one research question, and there is a correlation between the number of research questions and the number of question orientations; research questions within a single academic paper have relationships such as causal and progressive relationships, presenting structural patterns that are linear, tree-structured, or network-structured.

Full Text

Preamble

Analysis on the Research Questions of Academic Papers in China—Empirical Research Based on Papers Published in CSSCI in the Field of Library and Information Science from 2015 to 2020

Suō Chuánjūn, Lǐ Mùzǐ

School of Information Resource Management, Renmin University of China, Beijing 100872

Abstract: [Purpose/Significance] Academic papers originate from research questions. This study explores the research questions in academic papers to reveal their core content and provide a reference for knowledge discovery based on research questions. [Method/Process] Using papers published in CSSCI journals in the field of library and information science from 2015 to 2020 as data sources, 1,000 papers were randomly selected as samples. Content analysis was applied to bibliographic information to analyze the types and structures of research questions in academic papers, and to summarize the relationships between paper types and the types and quantities of research questions they contain, as well as the relationships and structural patterns among multiple research questions within a single paper. [Result/Conclusion] The study found that different types of academic papers exhibit certain regularities in the types and quantities of research questions they contain. A single academic paper typically includes more than one research question, and the number of research questions correlates with the number of question orientations. Research questions within a paper have causal, progressive, and other relationships, presenting linear, tree-like, and reticular structural patterns.

Keywords: Academic papers; Research questions; Types of academic papers; Types of research questions; Knowledge discovery

Classification Number: G250.2

DOI: 10.13266/j.issn.0252-3116.2021.19.011

Academic papers represent records of the process of discovering, analyzing, and solving problems by researchers [1]. Research questions are the prerequisite and foundation for academic papers, and they are also key for scientific users to acquire knowledge about the themes and content of academic literature. In scientific research, the discovery and acquisition of knowledge from academic papers constitute an important 环节 for researchers, who often focus on the problems and conclusions in literature when seeking information [2]. However, with the explosive growth of academic literature in recent years, information overload has led to knowledge disorientation. Existing retrieval methods lack direction and precision, requiring researchers to invest more time and effort to understand the research questions addressed in academic papers. A 2019 survey of researchers found that time spent on literature acquisition increased by 10%, while reading time increased by 11%. Therefore, how to quickly and accurately identify the research questions and knowledge in an academic paper has become an important issue in knowledge discovery.

In the context of big data, academic paper knowledge discovery presents both opportunities and challenges. Rich knowledge embedded in academic papers awaits discovery and excavation. Knowledge discovery should return to the

source of knowledge—research questions—integrating theories and research paradigms from philosophy of science, linguistics, and related fields. To promote understanding of academic paper content and improve the efficiency of knowledge acquisition, scholars have explored academic paper content from multiple perspectives, but none have addressed the relationship between academic papers and their research questions. Therefore, we randomly selected papers published in CSSCI journals in the field of library and information science from 2015 to 2020 as samples, using content analysis to examine research question types, quantities, forms, and relationships from the perspective of paper titles, providing a reference for knowledge discovery based on research questions.

1. Related Research on Research Questions in Academic Papers

Academic papers are one of the most important forms of academic achievements. Research questions in academic papers serve as the logical starting point for writing from a functional perspective, appear as interrogative sentences that authors aim to answer or explore from a formal perspective, and represent the thematic ideas contained in or expressed by the paper's topic from a semantic content perspective. Scholars from different disciplines have conducted research from various perspectives, primarily concentrated in library and information science, philosophy of science, and computer science.

1.1 Related Research in Library and Information Science

Research in this field mainly explores the essence of knowledge elements and research question knowledge elements, as well as question identification methods. In the late 1970s, American information scientist Professor Vladimir Slameca proposed that the unit of knowledge control would evolve from documents to “data elements” within documents, namely “knowledge elements” (knowledge element). Subsequently, B. C. Brookes suggested that the granularity of knowledge objects should be refined from documents to more microscopic knowledge elements [3]. Since then, the concept of “knowledge element” has been widely disseminated and attracted attention from domestic scholars. Recent research on knowledge elements has expanded from theoretical studies on concepts, classification, and representation to applied research on semantic models, and further to description rules and ontology construction for specific types of knowledge elements.

Wen Youkui argued that documents are composed of independent knowledge element links and studied knowledge element semantic link models, considering that descriptive knowledge elements include problem-description types, whose characteristic is proposing a new question and providing valuable answers [4]. Gao Guowei et al. proposed a knowledge element representation method of “concept, relationship, problem” and discussed the definition, expression, and

mathematical description of problems [5]. Suo Chuanjun studied the essence, structure, and description model of knowledge elements [6], and classified problem knowledge elements into three types based on problem-solving objectives: understanding-type, solving-type, and exploring-type problem knowledge elements, summarizing description rules for different types [7]. Qin Chunxiu defined problem knowledge elements as unresolved issues in scientific knowledge and practice proposed within specific knowledge contexts [8]. Li He and Du Xingye constructed an ontology of research questions and theories in academic papers [9].

These studies show that research questions in academic papers are generally considered a subclass of knowledge elements, possessing the attributes and structural features of knowledge elements, and constitute part of academic paper knowledge organization. Regarding knowledge element and research question identification methods, Chai Qingfeng et al. proposed a method combining manual and machine features for efficient and accurate extraction of knowledge elements from scientific literature [10]. Li Xin et al. identified research question words and method words in academic papers through lexical function identification [11]. Wen Hao et al. established an ontology model for creative knowledge points in academic abstracts, constructing a “question, method, result” triad pattern for creative knowledge points and implementing a text knowledge point deep mining method based on pattern recognition [12], and further proposed a cognitive analysis method for mining innovative points in academic abstracts [13].

1.2 Related Research in Philosophy of Science

Research in philosophy of science primarily focuses on theoretical studies of questions and scientific questions. Since ancient Greek times when debate flourished, philosophers have had profound understanding of the existence and significance of questions [14]. Today, the importance of questions for scientific research and development has reached consensus [15-17]; questions are not only the starting point of scientific research but also the growth point of disciplinary development [18]. Since the 20th century, scholars in philosophy, philosophy of science, and library and information science have conducted theoretical research on the definition, structure, types, and evolution of “questions” and “research questions” in scientific research.

Domestic scholars have also discussed scientific questions [19-22]. Compared with “questions,” “scientific questions” are those with researchers as the cognitive subject and scientific research as the scope [23]. Ma Laiping [23], Zhang Huaxia [24], Zhang Yi [25], Li Zhengkun [26], Qian Shaohua [27], and Shen Zhendong [28] have studied the types and structures of scientific questions from various perspectives. Other scholars have explored the logical mechanisms [29] and evolutionary mechanisms [30-32] of scientific question generation.

1.3 Related Research in Computer Science

In computer science and biomedicine, researchers have focused more on applied studies of “questions” and “question-answering” systems. For question decomposition and classification, S. M. Harabagiu et al. proposed a complex question-answering framework based on question decomposition, significantly improving answer relevance and comprehensiveness [33]. S. Harttumpf studied methods for decomposing complex questions and demonstrated related question decomposition classes to reduce question complexity and find more precise answers [34]. K. Roberts et al. conducted in-depth research on medical question classification, dividing common consultations into 13 question types and noting that question type classification aims to select optimal strategies for answers from publicly available health information resources [35]. Considering the diversity of question syntax and context, they also proposed a question decomposition annotation method for complex medical questions [36].

In summary, library and information science research describes research questions in documents from the perspective of knowledge elements but fails to delve into the relationship between documents and research questions. Philosophy of science research provides theoretical foundations for understanding the importance, concepts, structures, types, and evolution of “questions” and “research questions” in scientific research. Studies on question types, decomposition, and identification in professional fields like computer and biological sciences offer valuable references for classifying, identifying, and representing research questions in academic papers. Building on existing research, this study explores the relationship between academic papers and their research questions from the perspective of paper titles. The research questions examined here are scientific questions presented in academic papers that await investigation and resolution. This study aims to fill gaps and deepen understanding of research questions in academic papers for knowledge discovery.

2. Data Sources and Research Approach

2.1 Data Sources

Using CNKI as the primary data source, we collected and analyzed papers published in CSSCI journals in the field of library and information science from 2015 to 2020, obtaining 25,982 journal papers. We randomly selected 1,000 papers as research samples (ensuring coverage of 18 CSSCI journals in the field) and used Excel as a statistical tool to compile bibliographic information from the sample data.

2.2 Research Approach

Journal papers serve as pioneering research outputs in disciplinary fields, containing numerous domain questions with diverse content and stronger academic rigor, better reflecting research themes and directions in the field. Paper titles

are crucial components that accurately represent research objects and content [37]. As summarizing descriptions, titles express authors' writing intentions and core textual themes [38], serving as the most basic guide for reading academic papers and efficiently meeting readers' retrieval needs [39]. Studies have shown that title characteristics correlate with download/read counts [40] and citation frequency [41], and readers tend to prefer papers that present research questions or results in their titles [42]. Therefore, we employed content analysis to analyze research questions in academic papers from their titles.

The research approach includes four aspects: (1) identification of question factors contained in paper titles; (2) clarification of research questions included in academic papers; (3) analysis of the relationship between paper types and research question types; and (4) analysis of relationships among multiple research questions within a single paper. The research framework is shown in Figure 1 [Figure 1: see original paper].

3. Identification and Analysis of Research Questions in Academic Papers

3.1 Structure of Research Questions

Research questions have structure. A clear research question contains three elements: research object, question orientation, and research presupposition.

- (1) **Research object** is the basis of a research question, which depends on it. For example, in “What is the structure of knowledge elements?”, “knowledge element” is the research object.
- (2) **Question orientation** is the goal of question solving, representing an unknown part in knowledge that needs to be filled—a gap or logical fissure that generates a demand for researchers to investigate further and determine answers [43]. For example, in “What are the connotation, structure, and description model of...?”, “connotation,” “structure,” and “description model” are question orientations.
- (3) **Research presupposition** refers to the research background, methods, and answer domain for the research question. It constitutes part of the cognitive subject's background knowledge and serves as an important node connecting the question with the researcher's knowledge. For example, in “What is the application of... in digital humanities?”, “digital humanities” is the answer domain. Different research questions require different background knowledge and technical means, and researchers from different disciplines have different thinking patterns and paradigms regarding research questions. The answer domain is the scope determined in the question's discussion, assuming the solution lies within this range. When facing an unknown area they wish to understand, researchers always set an answer domain for the desired solution and screen each possible solution within it [44].

3.2 Identification of Question Factors and Question Analysis in Titles

Question factor identification involves recognizing structural elements of research questions. Since research questions in academic papers exist at the semantic level, we primarily relied on researchers' reading comprehension, supplemented by features listed in Table 1 such as part of speech and position, to manually read and annotate bibliographic information compiled in Excel. When identification was unclear from the title alone, we read the abstract for auxiliary judgment. Each researcher completed the annotation and then cross-validated with others; inconsistencies were resolved through secondary identification until consensus was reached to ensure accuracy and scientific rigor.

The specific information identified included three items: research object, question orientation, and research presupposition. Using the title "Research on the Structure and Process of Social Search Behavior: From the Perspective of Activity Theory" as an example:

- (1) **Research object identification:** Identify and annotate the research object of each paper as $M_b(X)$. The example contains one research object: $M_b(X) = \text{"social search behavior."}$
- (2) **Question orientation identification:** Through reading and understanding the title, determine the question orientation of the paper and annotate it as X . If multiple question orientations exist, annotate them separately and number them, counting the number of X as m ($m \geq 1$, m is a positive integer). The example clearly identifies two question orientations: $m = 2$, with $X_1 = \text{"structure"}$ and $X_2 = \text{"process."}$
- (3) **Research presupposition identification:** When research background or methods are explicitly presented in the title, annotate them as $M_c(X)$. In the example, the author explicitly shows the research perspective in the title: $M_c(X) = \text{"activity theory perspective."}$ Many papers' research perspectives or methods are not reflected in the title or are not explicitly stated, in which case no specific annotation is made.

Table 1 lists reference features for identification from titles, including part-of-speech features, position features, and linguistic features (characteristic words or sentence patterns). Research objects, presuppositions, and question orientations share similarities in part-of-speech features, all containing nouns and noun phrases with independent semantic information. Research objects and presuppositions typically appear at the beginning, middle, or end of titles, while question orientations never appear at the beginning since they are specific question directions following the research object. In terms of expression features, research objects and question orientations have no obvious characteristic words or collocations, while research presuppositions have relatively clear expression features.

3.3 Research Question Analysis and Description

Based on Section 3.2, we analyzed and described the identified research objects, question orientations, and research presuppositions together to clarify research questions in each paper. Research questions were uniformly expressed in interrogative form as $Q = (?X)M_b(X)M_c(X)$, where $M_b(X)$ is the research object, $M_c(X)$ is the research presupposition, X is the question orientation, and $(?)$ represents the interrogative feature of “What is...?” See the two examples in Figure 2 [Figure 2: see original paper].

Example 1: “Research on the Structure and Process of Social Search Behavior: From the Perspective of Activity Theory”

Social search behavior / structure / and / process / research: based on / activity theory perspective

$M_b(X) / X_1 / X_2 / M_c(X)$

Example 2: “Research on the Problems, Causes, and Countermeasures of Librarian Deprofessionalization”

Librarian deprofessionalization / problems, / causes / and / countermeasures / research

$M_b(X) / X_1 / X_2 / X_3$

Through statistics, among the 1,000 randomly selected samples, 915 were valid. For these valid samples, we counted the number of research questions per paper ($n \geq 1$, n is a positive integer), annotated and counted the types of research questions, and converted them into interrogative forms. The data analysis results are shown in Figure 3 [Figure 3: see original paper], and Table 2 provides three examples from the sample data.

Table 2 shows examples of research questions contained in paper titles, including research object $M_b(X)$, research presupposition $M_c(X)$, question orientation X (with quantity m), and research question $Q = (?X)M_b(X)M_c(X)$ (with quantity n).

4. Analysis of Academic Paper Types and Research Question Types

4.1 Academic Paper Types and Research Question Types

4.1.1 Types of Academic Papers Different disciplinary fields have their own theoretical systems, research paradigms, and characteristic research questions and methods. As primary forms of research output, academic papers manifest differently according to various research questions, methods, and approaches. Pang Fuxiang classified academic papers into argumentative, experimental report, discovery/invention, design/calculation, theoretical derivation, and review types [45]. Wang Fang [46] and Li Bowen [47] classified papers by research methods. Lin Gege et al. compared paper types in Nature and Web of Science databases, including commentary articles, news perspectives,

and reviews [48]. Building on these studies, we manually read 1,000 paper titles, annotated words, phrases, and sentence patterns that could identify paper types (e.g., “research progress,” “comparative analysis,” “empirical research”), and classified papers into seven categories: review papers, comparative papers, viewpoint papers, descriptive papers, principle exploration papers, empirical survey papers, and technical design papers, as shown in Table 3 .

Table 3 lists paper types and descriptions. Review papers take existing research on themes or phenomena as research objects $M_b(X)$, comprehensively analyzing and reviewing domestic and international research results, examining development history, current status, existing problems, and future prospects (question orientation X). Comparative papers take two or more phenomena or aspects as research objects $M_b(X)$, comparing similarities, differences, and relationships (question orientation X). Viewpoint papers present authors’ unique perspectives on research objects $M_b(X)$. Descriptive papers describe or interpret objective concepts, meanings, features, content, attributes, processes, and related information about research objects $M_b(X)$. Principle exploration papers explore internal principles, influencing factors, operating mechanisms, change causes, and evolutionary trajectories of research objects $M_b(X)$. Empirical survey papers verify causes, results, effects, and feasibility of phenomena or methods through surveys and empirical research. Technical design papers propose, construct, develop, improve methods, frameworks, models, and systems to solve practical needs.

4.1.2 Types of Research Questions in Academic Papers Academic paper writing is a knowledge creation process whose core 环节 is thinking. The thinking process first recognizes “what is” through differences in things, then observes “why” [49], and further considers “how to” and “how.” Based on existing question classification research and the thinking process in academic writing, we classified research question types into three categories according to question orientation: “what is” questions, “why” questions, and “how to” questions.

- (1) **“What is” questions** (“What is...?”) require judgment and analysis of research objects, with answers being factual statements. Researchers identify research objects and describe and explore their phenomena and essence, analyzing attributes, features, and related issues within research presuppositions to present explanations and analyses of existing things, phenomena, and theories.
- (2) **“Why” questions** (“Why...?”) primarily explore causal relationships and explanatory aspects between research objects and objective environments or internal factors, aiming to discover patterns and conditions behind phenomena, and reasons and influencing factors for things’ emergence and changes.
- (3) **“How to” questions** (“How to...?”) essentially seek solutions or application methods and practical results. Their resolution typically requires

seeking materials and designing instruments, techniques, and programming methods. Solution paths are not a finite set; even with the same goal, different solution paths can emerge, with only suitability and optimal solutions, no absolute measurement standards. Researchers need not only to “solve problems” but also to pursue “optimal solutions” with era development.

4.2 Analysis of Academic Papers and Their Contained Research Question Types

Different paper topics contain different research questions [50]. Through annotation and organization, the seven paper types contain 14 question orientations, with specific corresponding relationships between paper types and research question types.

- (1) Review papers include “research development history” and “research status” question orientations, both being “what is” questions. “Research status” usually includes “research development history” but emphasizes 梳理 of development processes with varying focuses depending on research objects.
- (2) Comparative papers include “relationship” and “comparison results” question orientations, focusing on summarizing and presenting comparison results of research objects or content, belonging to “what is” questions.
- (3) Viewpoint papers emphasize authors’ “viewpoints,” “existing problems,” and “value/significance” about research objects. These “viewpoints” represent authors’ new thinking and dialectical attitudes different from previous research.
- (4) Descriptive papers have relatively more question orientations with richer content and different levels: surface levels include concepts, attributes, and content of research objects; deeper levels involve change patterns and evolutionary trajectories, representing in-depth exploration of internal knowledge.
- (5) Principle exploration papers have exploratory question orientations investigating internal causes of research objects. As shown in Table 7, principle exploration papers address “What are the influencing factors, mechanisms, and principles of research objects?”, belonging to “why” questions, which only appear in principle exploration papers.
- (6) The question orientation “empirical and experimental results” mainly exists in empirical papers but also appears in technical improvement papers, serving different functions. In empirical papers, result analysis is fundamental; in technical improvement papers, it validates the answer to “What are the construction paths of frameworks, methods, systems, countermeasures, and models proposed by authors about research objects?”, demonstrating scientific validity.

Through statistics, the 915 valid samples contained 1,184 research questions, including 737 “what is” questions, only 11 “why” questions, and 436 “how to” questions. Table 7 shows the statistical distribution of samples containing n ($n \geq 1$, n is a positive integer) research questions. Each sample contained between 1-5 research questions, with papers containing one research question being most numerous (675 papers), followed by those with two questions (214 papers, accounting for a certain proportion). Only 26 samples contained more than two research questions, and only two samples contained 4-5 research questions.

5. Analysis of the Relationship Between Academic Papers and Research Questions

Based on Section 4, we conducted statistical analysis by paper type, including: (1) counting research question quantities and question factor quantities to summarize the distribution relationship between paper types and research question quantities; and (2) for papers containing $n \geq 2$ research questions, analyzing the logical relationships among the n research questions and summarizing the logical structures formed by research question relationships within a single paper.

5.1 Analysis of Academic Papers and Research Question Quantities

Table 6 lists the number of papers containing question factors and the statistical quantities of question factors. Among 915 valid samples, 477 could be annotated with research objects, presuppositions, and question orientations, while 438 could not clearly identify research presuppositions for two reasons: (1) in relatively macro and basic theoretical research, research presuppositions are often implicit in research objects or question orientations; (2) specific research presuppositions (research methods and answer domains) were not presented in titles. In Table 6, question factor quantities exceed paper quantities, indicating some papers contain multiple question factors. Statistics show 26 title samples contained two research objects (descriptive or comparative studies of multiple objects), 49 contained two research presuppositions (one from case explanations), and 240 contained more than one question orientation.

By combining question factors into research questions and counting them, the 915 samples contained 1,184 research questions. After classifying samples, we statistically analyzed research question quantities by paper type, with results shown in Figure 4 [Figure 4: see original paper]. The specific distribution is: 55 review papers containing 71 research questions; 26 comparative papers containing 31 questions; 168 viewpoint papers containing 220 questions; 153 descriptive papers containing 235 questions; 133 principle exploration papers containing 160 questions; 75 empirical survey papers containing 97 questions; and 305 technical improvement papers containing 370 questions. The average number of research questions per sample was 1.29, with descriptive papers containing the most (average 1.54), and other types fluctuating around the mean, indicating descriptive papers typically contain more research questions as collections of answers

to multiple question orientations about research objects.

5.2 Analysis of Relationships Among Research Questions Within a Single Paper

5.2.1 Internal Relationships Among Research Questions Section 5.1 shows that some papers contain more than two research questions, with a few samples identifying four or five questions, indicating that containing n research questions is normal. So what are the relationships among these n questions ($n > 1$, n is a positive integer)?

Based on rhetorical structure theory, we summarized relationships among n research questions in a paper as sequential, causal, and progressive relationships, with specific examples in Table 8 .

When a paper contains two research questions, their logical relationship is relatively simple, typically linear sequential, progressive, or causal relationships. Sequential relationships mainly exist in review and descriptive papers, reflecting the sequential connection of research ideas and questions. Progressive relationships appear in all paper types, representing deepening and extension of research questions, divided into deep progression (e.g., from “What are the concepts and content of the research object?” to “What are its characteristics?”) and expansive progression (e.g., extending to technical experiments or unknown operational fields). Comparative, principle exploration, empirical, and technical improvement papers often show causal relationships among n questions, validating causal effects between research processes and knowledge contained in papers.

5.2.2 Research Question Set Structure and Characteristics When a paper contains more than two research questions, the questions interconnect to form an organic whole presenting the paper’s 脉络, which we call a research question set. The question set has characteristics of integrity, synchronicity, and hierarchy, exhibiting linear, tree-like, and reticular structural patterns. Internal logical relationships include:

- (1) Research questions in the set influence each other sequentially, forming linear structures where directly connected questions are directly related and 跨越 questions are indirectly related. Linear structures are common in review, comparative, principle exploration, and empirical papers.
- (2) When every two research questions have significant relationships, forming linear relationships that interconnect, they create reticular structures common in viewpoint and descriptive papers.
- (3) When two questions have no obvious relationship but both relate significantly to another question, this forms a tree structure common in construction/improvement and descriptive papers.

Regardless of structure type, further research shows that research question types appear in certain sequences, and question sets have typological hierarchy and multi-dimensional content: (1) “What is” and “why” questions typically appear sequentially in progressive relationships, showing evolution from basic concepts to deeper causal mechanisms, from superficial to core research questions; (2) “What is” and “why” questions usually precede “how to” questions in causal relationships, indicating applied research questions require theoretical foundations; (3) Content-wise, a paper’s question set may contain multiple question types with multiple dimensions, directions, and perspectives.

6. Research Conclusions and Implications

6.1 Research Conclusions

Based on sample paper title information and content analysis of research question types, quantities, and relationships, we reached the following conclusions:

- (1) An academic paper typically contains more than one research question, and the number of research questions correlates with the number of question orientations, showing a linear relationship.
- (2) Research questions within a paper typically have linear, sequential, progressive, and causal relationships, forming question sets with linear, tree-like, and reticular structures. Research questions appear in certain sequences, and question sets have multi-typological and hierarchical characteristics.
- (3) Paper types relate to the types and quantities of research questions they contain. First, paper types correspond to research question types: review, comparative, descriptive, and viewpoint papers mainly contain “what is” questions; empirical and technical improvement papers mainly contain “how to” questions; principle exploration papers mainly contain “why” questions. Second, research question quantities have certain ranges, with sample statistics showing 1-5 questions per paper. Papers with one question were most numerous (675, 73.77%), followed by those with two questions (214, 23.39%). Only 26 papers (2.84%) contained three or more questions, and just two contained 4-5 questions. Third, descriptive papers contained the most research questions on average. Fourth, in the sample, “what is” questions accounted for 62.3%, “why” questions for 0.9%, and “how to” questions for 36.8%. Further analysis shows that recent library and information science papers focus on social hotspots (big data, cloud computing, AI) and application studies combining data/technology with library work, while research on underlying causes and mechanisms is relatively scarce.

6.2 Implications

- (1) The study found that 91.5% of paper titles effectively reveal research questions, with some clearly identifying question factors, while only 8.5% require abstracts and full text for clarification. Thus, obtaining research questions from titles is an important and effective approach for knowledge discovery.
- (2) All journal papers contain more than one research question. If research questions can be effectively identified and described, this can advance informetrics from lexical to semantic (question-based) analysis, enabling topic evolution analysis based on research questions. This can also enable classification and clustering based on research questions, facilitating organization and retrieval of academic papers by research questions, thereby improving literature discovery quality and efficiency for users.
- (3) The study found that 73.77% of papers contain only one research question. Effective identification and description would enable classification/clustering based on research questions, supporting question-based organization and retrieval of academic papers. For users, question-based retrieval could significantly improve literature discovery quality and efficiency, saving time and effort.

In summary, this study aims to identify research questions from paper titles and reveal relationships between papers and research question types and quantities, providing a new perspective for academic paper knowledge discovery. The approach of analyzing titles differs from previous knowledge discovery methods based on keywords, abstracts, or full text, offering valuable references for multi-perspective, in-depth exploration of research questions in academic papers. However, limitations include insufficiently rich disciplinary and linguistic samples, potentially limiting generalizability. Future research will expand multi-disciplinary and multi-language samples to test and deepen these conclusions.

References

- [1] Sheng Zhaohan. Problem-oriented: The driving force of management theory development [J]. *Journal of Management Sciences*, 2019, 22(5): 1-11.
- [2] Ribaupierre HD, Falquet G. Extracting discourse elements and annotating scientific documents using the SciAnnotDoc model: A use case in gender documents [J]. *International journal on digital libraries*, 2018, 19(2/3): 271-286.
- [3] Brookes BC. The foundations of information science (Part IV) [J]. *Journal of information science*, 1981, 3(1): 3-12.
- [4] Wen Youkui, Wen Hao, Xu Duanyi, et al. Text knowledge indexing based on knowledge elements [J]. *Journal of the China Society for Scientific and Technical Information*, 2006, 25(3): 282-288.

- [5] Gao Guowei, Wang Yajie, Li Yongxian. Research on knowledge element representation methods [J]. *Modern Information*, 2015, 35(3): 15-18, 33.
- [6] Suo Chuanjun, Gai Shuangshuang. Research on the connotation, structure, and description model of knowledge elements [J]. *Journal of Library Science in China*, 2018(4): 54-72.
- [7] Suo Chuanjun, Lai Haimei. Types and description rules of problem knowledge elements in academic papers [J]. *Journal of Library Science in China*, 2020, 47(2): 95-109.
- [8] Qin Chunxiu, Liu Jie, Liu Huailiang, et al. Research on a knowledge element-based content description framework for scientific texts [J]. *Library and Information Service*, 2017(10): 116-124.
- [9] Li He, Du Xingye. Intelligent evaluation of content innovation in academic papers based on knowledge elements [J]. *Library and Information Service*, 2020, 64(1): 93-104.
- [10] Chai Qingfeng, Shi Linyan, Mei Shan, et al. Knowledge element extraction from scientific literature based on fusion of manual and machine features [J/OL]. *Data Analysis and Knowledge Discovery* [2021-06-30]. <https://kns.cnki.net/kcms/detail/10.1478.g2.20210608.1550.002.html>.
- [11] Li Xin, Cheng Qikai, Liu Liubang. Design and implementation of a research literature analysis system based on lexical function identification [J]. *Library and Information Service*, 2017, 61(1): 109-116.
- [12] Wen Hao, Wen Youkui, Wang Min. Deep mining method for text knowledge points based on pattern recognition [J]. *Computer Science*, 2016, 43(3): 279-284.
- [13] Wen Hao, He Xiru. Cognitive analysis method for mining innovative points in academic abstracts [J]. *Journal of Intelligence*, 2021, 40(5): 489-499.
- [14] Wu Hongzhi. Aristotle's thought on question logic [J]. *Journal of Yan'an University (Social Science Edition)*, 1988(2): 35-38.
- [15] Zhang Tianrong. *What is science* [M]. Beijing: Tsinghua University Press, 2019.
- [16] Hilbert. *Mathematical problems* [J]. *Bulletin of the American Mathematical Society*, 1900(8): 251.
- [17] Barbara M. Wildemuth: *Applications of social research methods to questions in information and library science* [M]. 2nd ed. Santa Barbara, California: Libraries Unlimited, 2017.
- [18] Xiao Ximing. Library science research should be problem-oriented [J]. *Library*, 2005(1): 39-42.
- [19] Li Zuyang. Analysis of scientific questions [J]. *Studies in Dialectics of Nature*, 1996, 12(8): 10-13, 22.

- [20] Liu Guanjun. New exploration of the definition of scientific questions [J]. Theory Journal, 1999(4): 27-30.
- [21] Xia Congya, Liu Guohong. On the discovery and formulation of scientific questions [J]. Journal of Shanxi Normal University (Humanities and Social Sciences Edition), 2010, 55(2): 116-121.
- [22] He Ming. Problem consciousness and conscious problems—Characteristics, sources, and responses of questions in humanities and social sciences [J]. Academic Monthly, 2008, 40(10): 20-27.
- [23] Ma Laiping. Several types of scientific questions [J]. Shandong Social Sciences, 1992(1): 76-79.
- [24] Zhang Huaxia. On the logical structure of scientific questions [J]. Social Science Front, 1992(1): 54-61.
- [25] Zhang Yi. Dynamic research on scientific questions [J]. Philosophical Research, 1993(2): 27-33.
- [26] Li Zhengkun, Liu Wenjun. On the structure of scientific questions [J]. Journal of Central China Normal University (Humanities and Social Sciences Edition), 1998, 37(1): 51-54.
- [27] Qian Shaohua. Brief discussion on scientific questions [J]. Journal of Jiangsu University of Technology (Social Science Edition), 1999(2): 9-12.
- [28] Shen Zhendong. Logical system construction of scientific questions [J]. Journal of Jiangsu Second Normal College (Social Science), 2015, 31(1): 15-21.
- [29] Zhan Songsheng. On the logical mechanism of scientific question generation [J]. Modern Philosophy, 1992(3): 64-68.
- [30] Liu Min. Generation and evolutionary mechanism of scientific questions [J]. Journal of Northeastern University (Social Science Edition), 2015, 17(1): 8-13.
- [31] Liu Min. Super-cyclic development model of scientific questions from the perspective of generative theory [J]. Journal of Systems Science, 2015, 23(1): 24-27, 39.
- [32] Shen Zhendong. On the progress criteria of scientific questions [J]. Ningxia Social Sciences, 2012(4): 112-118.
- [33] Harabagiu SM, Lacatusu VF, Hickl A. Answering complex questions with random walk models [C]// SIGIR 2006: Proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval. New York: ACM, 2006.
- [34] Harttumpf S. Semantic decomposition for question answering [C]// ECAI 2008-18th European conference on artificial intelligence. Greece: IOS Press, 2008.

- [35] Roberts K, Masterton K, Fiszman M, et al. Annotating question types for consumer health questions [EB/OL]. [2021-02-22]. <https://www.semanticscholar.org/paper/Annotating-Question-Types-for-Consumer-Health-Roberts-Masterton/b4c4b9ce4b00ca2f05ac82fbdee98c4dd1e2a2f>.
- [36] Roberts K, Masterton K, Fiszman M, et al. Annotating question decomposition on complex medical questions [EB/OL]. [2021-02-22]. <https://www.webofscience.com/wos/alldb/full-record/WOS:000355611004035>.
- [37] Wei Ruibin. Research on the correlation between paper title characteristics and citations [J]. *Journal of Intelligence*, 2017, 36(11): 1148-1156.
- [38] Cheng Qikai, Li Pengcheng, Zhang Guobiao, et al. Academic text lexical function identification—Question method extraction based on title generation strategy and attention mechanism [J]. *Journal of Intelligence*, 2021, 40(1): 43-52.
- [39] Akman T. Selection of authors, titles and writing abstract [J]. *Turkish journal of urology*, 2013, 39(S1): 5-7.
- [40] Subotic S, Mukherjee B. Short and amusing: The relationship between title characteristics, downloads, and citations in psychology articles [J]. *Journal of information science*, 2014, 40(1): 115-124.
- [41] Letchford A, Moat HS, Preis T. The advantage of short paper titles [J]. *Royal Society open science*, 2015, 2(8): 1-6.
- [42] Jamali HR, Nikzad M. Article title type and its relation with the number of downloads and citations [J]. *Scientometrics*, 2011, 88(2): 653-661.
- [43] Zhang Huaxia. On the logical structure of scientific questions [J]. *Social Science Front*, 1992(1): 54-61.
- [44] Shen Zhendong. Logical system construction of scientific questions [J]. *Journal of Jiangsu Second Normal College (Social Science)*, 2015, 31(1): 15-21.
- [45] Pang Fuxiang, Jia Lihong, Zhang Hongxia. Classification and writing of academic papers [J]. *Journal of Taiyuan University of Technology (Social Science Edition)*, 2001, 19(1): 28-30.
- [46] Wang Fang, Shi Haiyan, Ji Xuemei. Application of theories in Chinese information science research: Content analysis based on *Journal of Intelligence* [J]. *Journal of Intelligence*, 2015, 34(6): 581-592.
- [47] Li Bowen, Zhang Chengzhi. Review of research methods in Chinese information science [J]. *Information Theory and Practice*, 2019(2): 1-6, 37.
- [48] Lin Gege, Hou Haiyan, Hu Zhigang. Classification and comparison of literature types in journal databases [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2019, 30(4): 418-425.
- [49] Zhang Chengwei, Liu Fanru, Hao Xutong. On the essence of knowledge and knowledge creation [J]. *Journal of Intelligence*, 2016, 35(4): 369-379.

[50] Zhu Xudong. On the connotation, types, and writing relevance of “questions” in dissertations [J]. Academic Degrees and Graduate Education, 2014(2): 18-22.

Author Contributions: Suo Chuanjun: Topic guidance, research design, paper revision and review; Li Muzi: Data processing and analysis, paper writing and revision.

Journal of Library and Information Service, Volume 65, Issue 19, October 2021

Call for Papers: 2021 Topics for Journal of Library and Information Service

1. Changes and impacts of academic information exchange models in the post-pandemic era
2. Library “14th Five-Year Plan” and 2035 long-term goals
3. Theories and methods for monitoring and identifying breakthroughs in key core technologies
4. Library and information work serving innovation-driven development strategies
5. National literature information resource security system integration development and service innovation
6. National literature resource security strategies under current international situations
7. Information resource blockade and counter-blockade for entity-listed institutions
8. Public information security from an intelligence perspective
9. Intelligent intelligence analysis technology and platform construction
10. Think tank construction and open data governance for major public health events
11. Application of new technologies and methods in government data opening
12. User cognition-oriented government open data management and services
13. Knowledge discovery theories and methods in government social media

14. Construction of library science basic theory in public cultural service systems
15. Research on public cultural digital resource service strategies
16. Research on university library public cultural system construction
17. Library cultural inheritance and dissemination services
18. Goals and key issues of high-quality library development
19. Library overall security and high-quality development research
20. Intelligence coordination mechanism design for emergency management
21. Health information behavior and personal health information management
22. Information organization and management in major emergency response events
23. Cultivation of public health information literacy for public health emergency management
24. Research on national intelligence work system innovation
25. Data management and utilization in different contexts
26. Open scientific data, data security, and personal information protection
27. Data identification, intelligence monitoring, and scientific early warning of public opinion
28. Intellectual property information open utilization mechanisms
29. Intellectual property information service capabilities and strategies
30. Public crisis governance policies and strategies
31. Long-term preservation of government digital resources
32. New generation metadata research
33. Standards and specifications for smart libraries
34. Smart library platform/third-generation library system platform con-

struction

35. Extended/augmented reality technology applications in digital libraries
36. Global Learning Tools Interoperability (LTI) open standards research
37. Digital inclusion and library information services
38. Reform and innovation of scientific research evaluation
39. Construction and application of public digital cultural resource knowledge graphs
40. Next-generation digital academic environment supported by cloud services
41. New Archives Law and archives governance research
42. Digital humanities and new liberal arts construction from LIS perspective
43. LIS discipline development under new liberal arts construction background
44. Positioning and value of LIS in digital humanities practice
45. Special collection technology applications in digital humanities
46. Library service innovation under new liberal arts and digital humanities background
47. Digital transformation research in LIS discipline
48. Current status and future of library, information, and archival science education
49. Re-examining research methods in library, information, and archival science
50. Construction of core competencies in library, information, and archival management

**Journal of Library and Information Service Editorial Office
December 12, 2020**

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

Source: ChinaXiv — Machine translation. Verify with original.