

数智时代的中国索引学：流变、趋向与展望^{*}

朱 禹，朱雨萌，叶继元

摘 要 数智技术正深刻改变索引学的研究范式与实务工作，实践快速发展与理论研究相对滞后形成鲜明反差。文章系统回溯中国索引学发展历程，通过对2014—2024年相关文献的计量与BERTopic主题建模分析，识别数智时代索引学研究的演进趋势与核心矛盾，为学科发展提供理论框架与实践路径。研究发现，中国索引学经历了萌芽奠基、恢复发展和转向深化三个历史阶段，在数智技术驱动下呈现出索引理论智能化、索引对象垂直化、索引组织学术化和索引产业科学化四大趋向；在此基础上，构建数智时代中国索引学研究整合框架，揭示学科发展面临机遇期与危机期、国际化与本土化、理论性与实践性、原创性与时代性等四对核心矛盾，提出危中寻机、植本开新、知行合一和守正创新四大发展路径，为索引学从治学之学向致用之学转型提供参考。

关键词 索引 索引学 数智时代 知识体系 学科建设

引用本文格式 朱禹，朱雨萌，叶继元.数智时代的中国索引学：流变、趋向与展望[J].图书馆论坛，2026，46（7）：25-37.

Chinese Index Studies in the Digital Intelligence Era: Evolution, Trends and Prospects

ZHU Yu, ZHU Yumeng & YE Jiyuan

Abstract Digital intelligence technologies are profoundly transforming the research paradigms and practical work of index studies, creating a stark contrast between rapidly advancing practice and relatively lagging theoretical research. This paper traces the developmental history of Chinese index studies and identifies evolutionary trends and core contradictions in the digital intelligence era through bibliometric analysis and BERTopic modeling of relevant literature from 2014 to 2024, thereby providing a theoretical framework and practical pathways for the discipline's development. The findings suggest that Chinese index studies have gone through three historical stages—sprouting and foundation-laying, recovery and development, and transformation and deepening—and, driven by digital intelligence technologies, demonstrate four major trends: intelligent indexing theory, domain-specific indexing objects, academically oriented indexing organizations, and scientifically oriented indexing industry. Furthermore, the paper establishes an integrative framework for Chinese index studies in the digital intelligence era, revealing four pairs of core contradictions facing the discipline: opportunity versus crisis, internationalization versus localization, theory versus practice, and originality versus contemporaneity. Accordingly, the paper proposes four developmental pathways—seizing opportunities amid crises, maintaining traditions while embracing innovation, integrating knowledge and action, and upholding fundamental principles while breaking new ground—to provide guidance for the transformation of index studies from a scholarship-focused discipline to a practice-oriented one.

Keywords index; index studies; digital intelligence era; knowledge system; disciplinary construction

^{*}本文系国家社会科学基金重大项目“学术‘全评价’视域下中国特色哲学社会科学评价体系建设研究”（项目编号：24&ZD323）、南京大学中国特色哲学社会科学自主知识体系建构“引领工程”重大研究专项“基于原创概念体系的中国图书情报学（信息资源管理学）自主知识体系与学术评价体系构建研究”（项目编号：202430059）和江苏省研究生科研与实践创新计划项目“生成式AI驱动的哲学社会科学‘全评价’模型构建与验证”（项目编号：KYCX25_0130）研究成果。

0 引言

索引学是一门研究索引理论、方法及其应用的学科。数智时代,以大数据和人工智能等为代表的新一代数智技术正深刻改变着与数据和技术有着天然联系的索引学。伴随着我国数字基础设施与数据要素的新基建,中国索引学实践迅速发生变化,索引技术从手工/机械式进化到自动/智能式,索引对象从文本式扩展到多模态,索引颗粒度从形式深入到内容^[1]。随着生成式人工智能快速赋能各行各业,高质量的大规模行业数据资源/知识资源成为推动新质生产力发展的重要基础要素。在此背景下,索引学作为以数据组织与知识揭示为核心的学科,其理论创新与实践发展对国家战略具有支撑作用。

然而,与快速发展的实践步伐和现实需求不符,数智时代索引学的理论研究相对滞后,我国索引学界尚未系统地开展对中国索引学学术体系、学科体系和话语体系的研究。进入数智时代,我国索引事业面临着理论研究式微、学科队伍松散和行业话语薄弱等问题。在数智技术浪潮的冲击下,如何在坚持中国学术本土化、自主性、原创性与学术国际化的前提与基础上,真正建立起符合中国国情的索引学知识体系至关重要。

现有关于索引学的研究综述对中国索引学之流变梳理不够深入,对数智时代索引学发展趋势的剖析也不够系统。系统梳理我国索引学的研究和实践历史,有助于发现主要问题、明确发展方向。因此,本文通过文献计量与BERTopic主题建模分析,系统考察2014—2024年中国索引学研究的演进脉络与发展趋向,旨在明确数智时代中国索引学自主知识体系的构建路径与发展定位。

1 文献综述

关于中国索引研究的综述性成果主要集中于索引编纂实践、技术演进及研究热点的计量分析层面。部分研究通过文献计量与可视化方法^[2],对索引研究的主题结构、技术热点与演进趋势进行描绘^[3-4];另有研究围绕数据库索引^[5]、无纸化

编纂^[6]、科学数据索引^[7]等具体实践,对21世纪以来我国索引工作的技术路径与应用成效进行回顾与总结^[8-9]。这些成果从不同侧面揭示了我国索引研究与实践的发展图景,为理解索引事业的技术进步与应用拓展提供了重要参考。然而,数智时代的中国索引学仍处于理论体系有待凝练、学科共识尚未形成的发展阶段,而当前研究多以“索引研究”或“索引工作”为分析对象,侧重对研究成果、技术手段或实践经验的整理与评述,尚未从整体上回应“索引学”作为一门学科的理论形态与知识结构问题。既有综述较少从历史连续性与知识形态演进的视角,系统梳理不同时期索引研究所关注的核心议题、研究重心及其内在逻辑;亦缺乏对索引学在数智时代所面临的理论转向、概念重构与学科定位等问题的综合性反思。有鉴于此,本文通过历史分期与数据驱动分析相结合的方式,考察中国索引学在不同时期的研究主题与知识形态特征,旨在为数智时代中国索引学自主知识体系的构建提供历史参照与实证依据。

2 研究方法

本文以中国知网、超星、万方和维普等学术数据库为数据源,检索时间为2024年11月。检索以“索引”为关键词,限定信息资源管理学科,同时采用人工录入方式补充主流数据库收录不全的《中国索引》辑刊论文。经数据清理与去重,最终获得327条有效论文。

考虑到研究重点在于揭示数智时代的具体研究趋向,本文进一步聚焦2014—2024年这一关键转型期的138篇研究性论文作为分析样本。2014年前后神经网络模型开始在自然语言处理中被广泛采用,为索引智能化提供了技术基础,因此将2014年作为数智时代索引研究的起点具有方法论意义。该时段年均发文量12.5篇,明显高于2014年前的3.2篇,且呈双峰分布(2015年47篇、2020年43篇),反映了数智技术对索引研究的阶段性推动。

本文采用BERTopic主题建模方法对上述

138篇研究性论文进行语义聚类分析,以识别数智时代索引研究的核心主题及其结构特征。BERTopic融合Transformer语义嵌入与密度聚类,能更有效捕捉中文学术文本中的深层语义关系,适用于样本规模中等、主题边界较为复杂的研究语料^[10]。具体实现上,选用多语言预训练模型paraphrase-multilingual-MiniLM-L12-v2对论文标题、摘要与关键词进行联合向量化表征,经UMAP降维后采用HDBSCAN密度聚类算法(min_cluster_size=5, min_samples=2)自动识别主题簇,最后通过c-TF-IDF算法提取各主题的高区分度词项。此外,为系统把握数智时代索引学研究趋向的历史根基,本文在数据驱动分析之前,首先对中国索引学的发展脉络进行历史梳理,以期为后续主题分析提供学科语境与参照框架。

3 中国索引学的历史流变

中国索引学的发展历程受学科内在演进逻辑与外部社会历史条件的共同塑造。从标志性制度事件来看,1949年中华人民共和国成立从根本上重塑了索引事业的制度环境与服务导向,1991年中国索引学会的成立则标志着索引学从分散的

编制实践走向自觉的学科与队伍建构,二者构成学科发展的两个关键转折点。与此相应,不同阶段的索引研究在核心议题、方法取向与知识形态上也呈现出实质性差异,信息技术变革与国家政策导向则为各阶段的演进提供了外部驱动。据此,本文将中国索引学的发展历程划分为萌芽与奠基(约1500—1948)、恢复与发展(1949—1990)和转向与深化(1991—)3个阶段。中国索引学的发展脉络与关键事件如图1所示。

3.1 萌芽与奠基(约 1500—1948)

索引学是历史悠久的学科，旧称“通检”“备检”“录”，近代也曾被称为“索引法”。汉语索引是英文“index”一词的日文译文之移用，我国台湾又据“index”一词音译为“引得”。我国严格意义上的索引产生于明清，清代学者章学诚不仅编制索引还提出索引的编纂理论^[11]，为现代索引的诞生作了实物准备和理论准备^[12]。17世纪起，西学东渐为中国索引学注入新的思想与工具，阳玛诺和裨治文等西方传教士在华编制的书后索引与期刊索引^[13]成为中国近现代索引学的启蒙基石^[14]。

新文化运动为索引学的发展提供了契机。1917年林语堂在《创设汉字索引制议》^[15]中首次

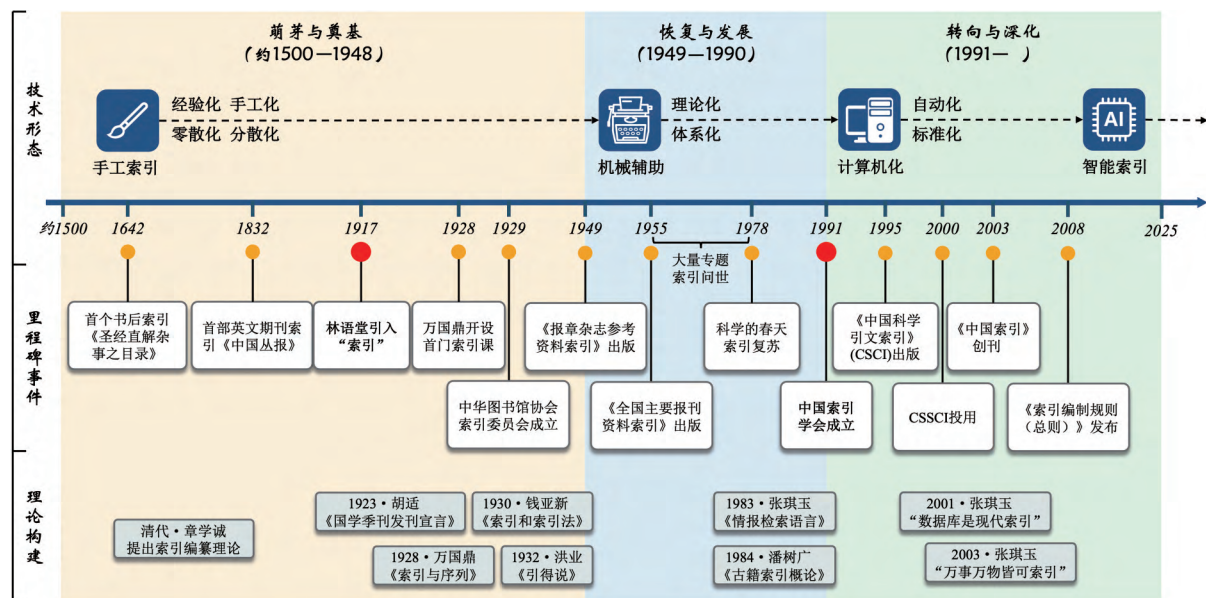


图1 中国索引学发展脉络

将日文“索引”一词引入中国并明确倡议为中国书籍编制索引。随后,他在《新青年》上进一步提出“汉字索引制”的具体设计,推动索引的价值逐步被学界认可。1928年秋,万国鼎在金陵大学开设“索引与序列”课程,是为中国索引学理论教育的肇始。胡适的《国学季刊发刊宣言》^[16]与万国鼎的《索引与序列》^[17]等文献奠定了索引运动和中国索引学初步体系化的思想基础。

1929年后,索引学进入机构化与理论化发展的新阶段^[18]。中华图书馆协会索引委员会^[19]、哈佛燕京学社引得编纂处和北平图书馆索引组等专门机构相继成立;钱亚新的《索引和索引法》和洪业的《引得说》等专著出版,标志索引学理论初步体系化;其他学者亦围绕中文书籍与期刊索引编纂问题展开了研究与学说创新。尽管抗日战争使我国索引学发展一度受挫,但索引及索引学的核心价值已深入学界,为后续发展奠定了坚实基础。

总体来看,这一时期中国索引研究以索引编制实践与方法引介为核心,关注点集中于索引的工具属性与学术辅助功能,尚未形成系统的理论反思与学科自觉。

3.2 恢复与发展(1949—1990)

中华人民共和国成立后,索引工作迎来恢复与发展的新阶段,不仅满足了人民群众的学习需求,也为科学研究和社会发展提供了重要的文献保障。

为适应中华人民共和国成立初期人民群众的学习需求,报刊索引率先得到发展。1949年10月,江苏省立教育学院编辑出版了中华人民共和国成立后最早的半月刊报刊索引——《报章杂志参考资料索引》。上海图书馆出版的《全国主要报刊资料索引》(后改名为《全国报刊索引》)持续至今,成为我国历时最长的连续性索引。

学者们也对专题性索引的编制表现出高度重视,尤其是重要思想家著作的索引。中央编译局和国内多所大学集中编制了马克思、恩格斯、列宁、斯大林和毛泽东等著作索引。1956年《人

民日报》社论提出“应该积极地有计划地进行书目和索引的编制工作”^[20],带动了全国索引事业的蓬勃发展^[21]。

1978年后,索引编制与研究全面复苏。1979年的《汉语主题词表》(试用本)为索引语言的发展奠定了理论基础;1983年,张琪玉的《情报检索语言》标志着中国索引语言研究走向成熟;《古籍索引概论》^[12]等著作推动了索引理论的深入发展。20世纪80年代后期,个人计算机逐渐被用于编制索引^[6],中国索引开始由手工方式向计算机辅助转型^[22],索引越来越多地以数据库的形式出现,开启数字化转型新时代。

这一阶段索引研究呈现出明显的国家任务与实务导向特征,索引作为知识基础设施的功能得到强化,但理论层面的系统建构仍相对薄弱。

3.3 转向与深化(1991—)

为促进索引理论研究和繁荣索引编辑出版,中国索引学会于1991年12月成立,标志着我国索引事业进入一个崭新的发展阶段,也标志着“索引学”作为一门学科的确立^[23]。

随着数字化、网络化和智能化技术的迅速发展,索引研究对象从传统印刷型逐渐扩展到数字型,推动了索引编制技术的自动化与智能化转型。1995年中国科学院研制的《中国科学引文索引》结束了中国没有本土科学引文索引的历史。随后,中国科学引文索引数据库(CSCD)、中文社会科学引文索引数据库(CSSCI)和中文图书引文索引(CBKCI)^[24]等索引数据库相继问世,为学术研究提供了重要的数据支撑。2003年专门研究索引学的期刊《中国索引》创刊,国家标准《索引编制规则(总则)》等陆续颁布,索引学研究和实务工作逐步走深走实。在此趋势下,索引/数据库也增长迅速。张琪玉教授及时提出“数据库是现代索引”^[25]与“万事万物皆可索引”^[26]的论断,为数智时代中国索引学自主知识体系建立奠定了良好的基础。

近年来索引研究的对象、技术路径与应用形态均在发生深刻变化,但相关研究仍以技术应用和实践经验总结为主,对索引学在数智时代的核

心概念、理论范畴与研究范式缺乏系统梳理。因此，在系统回顾历史脉络后，有必要借助文献计量与文本主题建模方法，对近年来中国索引学研究的主题结构与演进趋势进行数据驱动分析。

4 数智时代中国索引学研究趋向

BERTopic 文本聚类识别出 6 个一级主题，各主题及其核心关键词如图 2 所示。主题 1 “数智索引原理” 聚焦数智环境下的索引组织与检索技术；主题 2 “索引编制方法” 关注索引编制规范、流程与质量控制；主题 3 “古籍索引” 涉及中国古籍文献的索引编制与数字化；主题 4 “标引技术” 探讨汉字检字法与辞书编纂；主题 5 “索引事业进展” 涵盖索引学理论与期刊调

查；主题 6 “索引机构/组织” 涉及中国索引学会的组织建设与学术活动。

进一步采用层次聚类方法向上归纳(见图 3)。首先基于语义相似度，将 6 个主题聚合为 3 个关联主题组：主题 1 与主题 2、主题 3 与主题 4、主题 5 与主题 6，分别形成语义关联紧密的聚合单元。在此基础上，结合聚类结果的语义内涵，对 3 个聚合单元进行命名，最终形成“数智索引原理与方法”“古籍索引技术与实践”“索引机构与组织发展”三大主题类。基于主题聚类结果，本文归纳出数智时代中国索引学研究“理论智能化”“对象垂直化”“组织学术化”“产业科学化”等 4 种趋向，共同构成数智时代索引学研究的基本趋向。

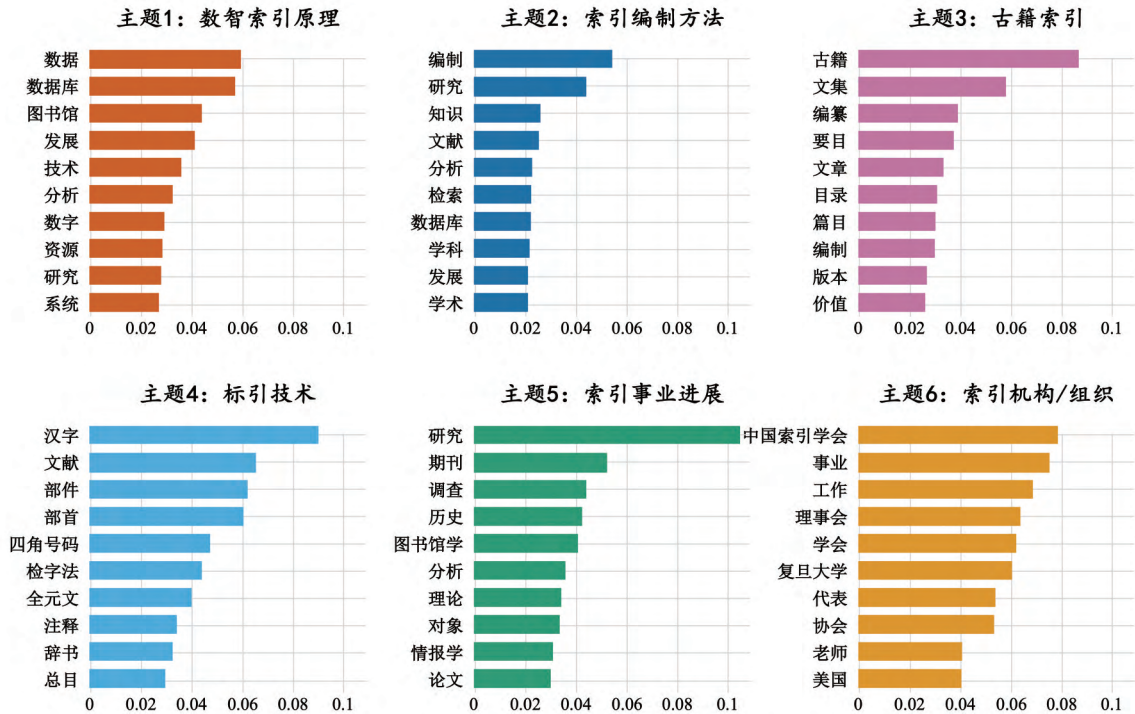


图2 主题聚类结果及其词频分布

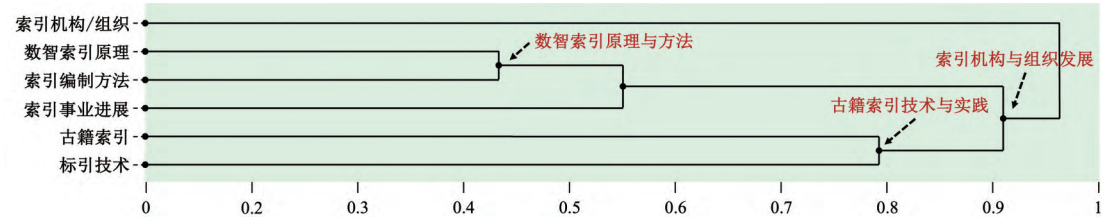


图3 层次聚类过程

4.1 索引理论智能化

索引理论智能化指人工智能和大数据等数智技术驱动的索引学理论革新与方法突破。中国索引从手工编纂、计算机辅助阶段迈入标准规范的自动索引编制时代^[27]，其技术突破体现在软件和支持技术两个层面。

在软件层面，“索引之星”“索引家”等中文索引软件相继研制成功^[28]，索引软件与辅助技术日臻完善^[29]。在支持技术层面，传统机器学习算法的引入推动了实质性进展。比如，张思龙等探索了基于知识图谱的智能索引技术^[30]，王弘毅分析了智能索引系统功能优化^[31]，这些研究标志着索引方法从字面匹配向语义关联演进^[32]，预示索引学与智能技术驱动的知识工程的深度融合。2023年后，生成式人工智能技术的兴起则为索引理论与实践开辟了新的空间。叶继元等探讨了智能检索下的数据库书目控制实施方案^[27]，以及以人机交互方式编制索引/数据库的原理与路径^①，朱禹等提出利用检索增强生成(Retrieval-augmented Generation, RAG)技术编制索引的方案^[33]，刘丹等探讨该技术在史料索引编制中的应用^[34]，这些探索展示了生成式人工智能在索引学领域的应用潜力。

然而，当前的智能化转型面临理论空白与实践困境的双重制约。在理论层面，尽管新技术为索引编制提供了迅捷的智能化工具，但索引学理论研究却未能同步跟进。现有研究虽提出智能标引、语义索引等新概念，但未形成系统理论框架，对数智时代索引的基本范畴(对象、功能、质量等)的重新界定尚未达成共识。在实践层面，专业索引的领域数据规模有限，质量标准与统计指标错位，领域知识理解不足。上述问题表明，技术创新与专业知识的融合路径仍待深入探索，索引学亟需构建适应数智时代的理论体系与方法论框架。

4.2 索引对象垂直化

索引对象垂直化指索引从通用型工具向特定领域、特定文献类型的专业化深度拓展，核心是通过聚焦细分场景实现知识服务的精准化与深度化。

4.2.1 古籍索引

古籍索引已成为数字人文基础设施的核心组成部分，古籍索引研究承载着中华优秀传统文化传承的重要使命。相关研究覆盖经、史、子、集等各类典籍^[35]，既包括《四库全书》《永乐大典》等大型文献集成，也涵盖地方志、家谱、金石碑刻等特色古籍，研究内容从传统编制技术延伸至数字化转化、知识组织及检索系统开发全链条。在数智技术支撑下，古籍索引正从静态检索工具向动态知识网络转型。王昊等提出中华古籍活化实践框架^[36]，通过智能标引、知识抽取和关联数据等技术构建古籍知识体系；范炜等运用自然语言处理和知识图谱等技术探索古籍智慧数据生成路径^[37]，彰显了索引学在文化遗产中的独特价值。

4.2.2 其他专题索引

随着通用检索需求已被搜索引擎较大满足，索引的核心价值逐渐聚焦于专业化、深度化的知识服务^[38]。具体而言，一是聚焦科学领域，除CSCD和CSSCI等外，研究者持续开发了诸多面向特定场域的科学数据索引/数据库^[7]。当前，本团队正利用生成式人工智能技术开发“中文学术评论索引数据库”，尝试将索引与学术规范、学术评价深度结合^[39]。二是覆盖特色文献类型，从论文集、思想家著作延伸至档案、手稿、口述史、日记等文献^[40]，如傅德华整理的方志敏烈士传记资料索引^[41]和王彦祥编撰的志书索引^[42]，为专题研究提供了不可替代的文献支撑。这类索引凭借领域知识深度、组织系统性及服务专业性，推动索引从“做大做全”向“做专做精”转型。

索引对象垂直化的深层逻辑在于信息环境分化对索引功能需求的重塑。然而这一趋向也面临现实制约：一是资源分散导致研究碎片化，不同领域、不同文献类型的索引成果缺乏有效整合，没有形成一体化系列；二是标准不统一阻碍互联互通，各垂直领域的索引编制规范难以兼容；三是可持续性不足，垂直化研究依赖稳定的资源投入与专业团队，部分小众领域面临支撑乏力问

题,如何在保持垂直深度的同时实现横向整合成为关键命题。

4.3 索引组织学术化

索引组织学术化指索引研究主体从实务机构向高校学术机构转移、索引学会从行业协会向学术共同体演进的转型过程,核心特征体现为学术话语权的结构性转移与学科化发展诉求。中国索引学会自1991年成立以来,通过《中国索引》构建“固守学术、开放包容”的学术发表平台,努力推动索引从“编制技艺”向“学术研究”转型,但主流图情期刊中的索引学研究发文有限,反映出索引学在学术场域的边缘困境,学科影响力亟待提升。

学术化转型的深层驱动是数智环境对传统索引工具的替代性冲击。随着数字化、智能化技术普及,传统实务机构的索引编制功能部分被技术工具替代,而高校凭借科研资源与人才优势,成为索引理论创新、技术研发的核心阵地。中国索引学会顺应这一形势,通过举办学术会议、开展国际合作^[43]及推动理论培训^[44]等方式,进一步强化了索引学的学术属性,致力将其构建为具有独立地位的二级学科。

然而,学术化转型过程中也暴露出三大矛盾:一是理论与实践脱节,高校的学术研究多聚焦理论与技术研究,研究成果难以有效转化为实务工具,导致对社会及行业的实际需求回应不足;二是普及性与专业性失衡,学术话语体系建设不足限制了索引概念及理念向大众的正确传播,导致社会上对索引的认知度较低或易与计算机科学所言“索引”混淆;三是学科认同薄弱,学科体系尚未完全建立,核心概念界定模糊,与图书馆学、情报学等相关二级学科的边界不清,导致研究力量分散,难以形成学术合力。与此同时,学术体系和话语体系与现实需求也有较大落差。如何在推进学术化的同时保持与实践的紧密联系、在提升专业性的同时扩大社会影响力、在学科建设中明确边界与特色,成为索引组织学术化进程中必须回应的核心命题。

4.4 索引产业科学化

索引产业科学化指索引编制从经验性操作向规范化、标准化与专业化转型,核心是通过建立行业规范与完善质量控制机制,来提升索引产业的整体发展水平。

这一转型在政策与标准层面已取得实质性进展。《关于进一步加强学术著作出版规范的通知》明确了书后索引的编制要求,为学术出版领域的索引应用提供了政策依据^[45];《索引编制规则(总则)》《地方志索引编制规则》《学位论文内容索引编制规则》等系列标准的颁布^[46],则规范了索引编制的流程、方法与质量要求,推动索引编制从个体经验向行业规范转变。此外,中国索引学会制定了《索引编制服务报价表》,形成了具有中国特色的索引费计算方法,为索引产业市场化运作提供了参考。

然而,政策与标准的完善尚未转化为产业的实质性突破。我国索引产业仍面临三重制约:一是索引应用覆盖率低,新书配置书后索引的比例不足10%,远低于西方学术图书80%—90%的水平^[47—48];二是专业人才稀缺,索引员队伍规模小、专业化水平参差不齐,缺乏系统的资格认证与培训体系^[49];三是市场机制不完善,索引费标准不统一、价值认同缺失,导致索引编制的社会效益与经济效益难以平衡^[48]。

尽管张琪玉很早就提出索引员资格认证和培训设想^[49],建议中国索引学会帮助索引员承接索引编纂项目,解决索引员的利益诉求问题,形成认证培训的良性循环;刘苏南也借助索引理论改进了黄页编制技术,推动索引走向大众化^[50],但索引编制的社会效益与经济效益双重价值尚未被充分认识。如何在市场化与公益性、成本与质量之间寻求平衡,推动索引产业科学化发展仍需持续探索。

4.5 数智时代中国索引学研究整合框架

为系统揭示四大趋向的形成机制与演化逻辑,本文以社会技术系统理论、“技术—组织—环境”理论、“物理—事理—人理”方法论为基

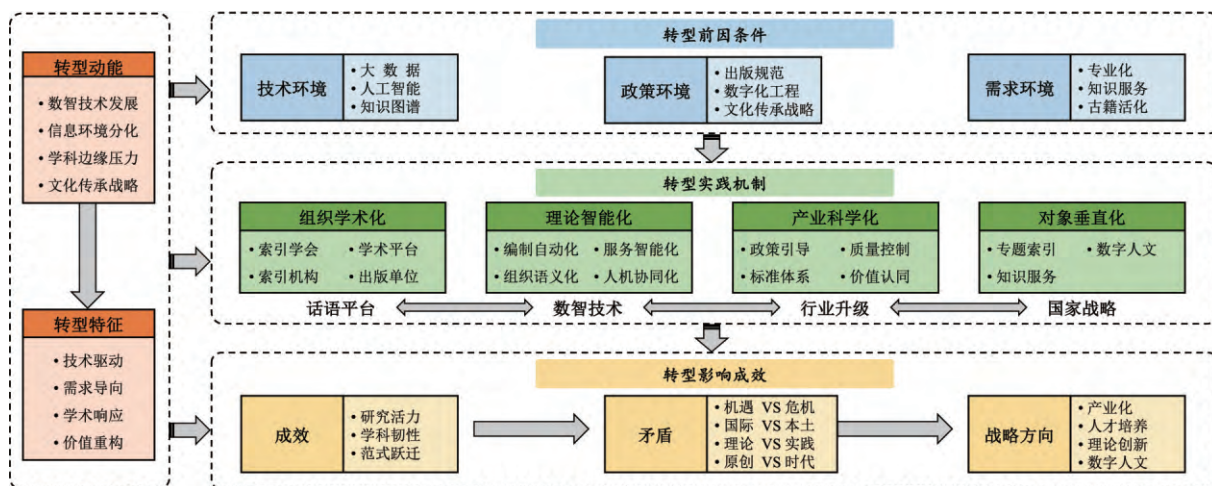


图4 中国索引学研究整合框架

础，构建“数智时代中国索引学研究趋向的整合框架”（见图4）。

(1)索引学数智化转型生发于复杂的社会技术环境。大数据、人工智能和知识图谱等数智技术革新索引编制模式和知识组织方法，推动索引从手工机械式向半自动、自动、智能式跨越。但转型仍面临学科体系未建立、核心概念模糊等学术挑战，以及索引价值认同缺失、专业队伍滞后等产业挑战。

(2)索引学数智化转型是多重因素交互作用的持续过程。环境层面，技术环境提供技术可能，政策环境(出版规范、数字化工程、文化传承战略)提供政策支撑，需求环境(专业化发展、深度知识服务、古籍活化利用)驱动价值重构。技术层面，索引编制技术经历“简单—复杂—简捷”演进路径，智能标引、语义索引和知识图谱等技术突破推动索引从字面匹配向语义关联演进。组织层面，索引机构、学术组织及出版单位协同推进标准规范建设与产业科学化进程。

(3)索引学数智化转型呈现4种趋向的协同演化机制。组织学术化为理论智能化提供话语平台，理论智能化为产业科学化注入理论动力，产业科学化对象垂直化提供实践场景，对象垂直化反哺组织学术化的学科建设。四种趋向的协同演化推动索引学从传统文献整理工具向数智知识服务范式跃迁。

(4)索引学数智化转型引发不同层面的成效与矛盾。“物理”层面体现为索引软件开发、数据库系统构建、知识图谱技术应用等技术工具与基础设施建设；“事理”体现为《索引编制规则(总则)》等系列标准的逐步建立和索引产业的科学化与规范化发展；“人理”层面体现为学术共同体建设与知识服务的专业化转向。转型成效体现为索引学研究的活力与韧性，但也暴露出四大深层矛盾。

5 数智时代中国索引学的挑战与展望

5.1 挑战

5.1.1 机遇期与危机期的矛盾

数智时代为中国索引学带来了前所未有的发展机遇。数智技术正在索引技术、对象与颗粒度3个维度推动深刻变革，为索引学开辟了广阔的应用场域。国家数据基础设施建设和文化数字化战略等政策导向，进一步赋予了索引学在数据要素增值与文化遗产传承中的关键角色，使其成为支撑新质生产力发展的重要基础。

然而，机遇背后潜藏着深刻的危机。传统索引学理论研究滞后于实践发展，学科体系尚未完全建立，核心概念界定模糊，索引学与图书馆学、情报学等相关二级学科的边界不清，导致学科认同感薄弱。学科队伍松散，青年学者流失严重，研究力量分散，难以形成系统性的理论创新

成果。此外,随着数据科学和人工智能等新兴学科的强势崛起,索引学面临被边缘化的风险,社会认知度偏低或被计算机类学科的话语体系所覆盖,多数人知晓SCI和CSSCI但对索引的来源及其核心价值认知不足,使得索引学在资源竞争、人才吸引等方面处于劣势。机遇与危机的并存,要求索引学界在拥抱技术变革与国家战略的同时,直面学科发展的内生短板,在破局中实现可持续发展。

5.1.2 国际化与本土化的矛盾

全球化背景下,中国索引学面临国际化与本土化的双重诉求。国际化为索引学提供了重要的发展动能,通过借鉴国际前沿的索引理论、技术方法和行业标准,能够快速提升我国索引学的研究水平和实践能力,推动索引工作与国际接轨,助力中国索引学走向世界舞台。与美国索引协会等国际组织的交流合作^[51],为我国索引学提供了跨界对话的平台,有助于拓宽研究视野并吸收先进经验。

与此同时,过度依赖国际理论体系,可能导致本土特色的丧失。中国索引学有着独特的历史积淀,丰富的文化典籍资源、独特的汉字体系以及国家治理中的特殊数据需求,都要求构建符合中国国情的索引学自主知识体系。当前,部分研究存在简单套用国外理论、忽视本土实践的问题,导致学术话语的主体性缺失。如何在吸收国际先进成果的同时,实现外来理论的本土化转化与本土经验的理论化提升,成为当前我国索引学发展中必须破解的核心矛盾。

5.1.3 理论性与实践性的矛盾

索引学作为一门兼具理论性与实践性的学科,其健康发展离不开二者的良性互动。

理论层面,数智时代要求索引学重构核心概念和知识体系,构建适应数智环境的理论框架。然而,当前我国索引学理论研究泛而不精,多停留在对技术应用的表层描述,对索引与数据、数据库、信息、知识和智能等概念的内在关联缺乏系统厘清,对索引本质、学科边界及发展规律的深度探究不足。部分研究过于依赖传统理论框

架,对新技术带来的范式变革回应滞后,难以形成具有学理性和指导性的原创理论。

实践层面,索引应用场景已从传统的文献索引延伸至科研数据、企业数据和民生数据索引等多个领域,但技术应用不充分、标准不统一与成果转化不畅等问题依然突出。部分索引编制仍依赖传统方法,对新技术的接受和应用能力不足;标引整理工作与信息组织、知识组织领域的知识融合不够,导致索引产品的实用性和适配性不强。

理论与实践需求的脱节,使得索引学难以形成理论指导实践、实践反哺理论的良性循环,制约了学科的整体发展。

5.1.4 原创性与时代性的矛盾

数智时代的技术变革和社会需求,要求索引学在保持学术传承的同时,实现原创性与时代性的有机统一。

原创性是学科发展的核心动力,索引学需要突破传统理论框架的束缚,基于数智技术的本土发展态势,提出新的概念、理论和方法,构建具有中国特色的自主知识体系。当前,我国索引学研究存在原创性不足问题,部分研究重复度高,缺乏对前沿问题的突破性探索,在国际索引学界的话语权较弱。

同时,时代性要求索引学紧跟技术发展和国家战略步伐,主动回应数智时代的新需求。生成式人工智能对大规模中文知识数据集的需求、国家文化数字化战略对文化遗产活化的要求、数据要素市场化配置对数据索引标准化的诉求,都为索引学提供了新的研究课题。但部分研究存在重传统轻创新、重形式轻内涵的倾向,对时代新命题的敏感度不足,未能及时将时代要求转化为学术创新动力。如何在坚守学科根脉、传承优秀学术传统的基础上,紧扣时代脉搏,增强原创意识,推出适应数智时代需求的创新性成果,实现学术传承与时代创新的辩证统一,是中国索引学面临的重要挑战。

5.2 展望

5.2.1 危中寻机:推动索引事业产业化发展

机遇与危机并存是数智时代中国索引学最突

出的特征。人工智能等技术为索引学开辟了广阔应用空间,但学科边缘化风险、社会认知度偏低、研究力量分散等危机同样严峻。化解该矛盾的关键在于将技术机遇转化为产业实效,通过产业化发展激活索引学的社会价值。

在中国特色社会主义市场经济背景下,应打破索引事业小众化、公益化的局限,构建专业化、职业化、规模化的产业生态。一方面,依托中国索引学会的各实践基地,培育一批专注于索引技术研发、数据标引服务与索引产品定制的高新技术企业,形成从技术研发到市场应用的完整产业链。例如,针对企业数据治理、社会民生数据整合、政务数据开放和科研数据管理等需求,开发定制化索引解决方案,将索引服务拓展至金融、政务及教育等多个领域^[52]。另一方面,结合灵活就业政策与市场需求,培育索引编制、数据标引和索引质量审核等新型职业,吸纳高校毕业生、专业技术人员及灵活就业群体参与,缓解就业压力。同时,推动索引应用的强制化与标准化,在学术出版和数据共享等信息服务领域建立索引编制规范,提高索引应用覆盖率,为产业发展创造稳定需求。

产业化发展的实质是以市场需求倒逼学科供给,使索引学在资源竞争与人才吸引中摆脱劣势地位,将机遇期的技术红利切实转化为学科可持续发展的内生动力。唯有让索引学从“治学之学”走向“致用之学”,方能在危机中把握机遇,实现学术价值与经济社会效益的统一。

5.2.2 植本开新:立足本土传统对接国际前沿

如何在国际化与本土化之间寻求平衡是贯穿中国索引学发展历程的长期命题。借鉴国际前沿索引理论、技术方法与行业标准是提升研究水平的重要途径,但过度依赖国际理论体系而忽视本土实践,则可能导致学术话语的主体性缺失。要化解这一矛盾,既不能闭门造车,也不能简单移植,而应认识到本土传统并非封闭的学术资源,恰是参与国际对话的独特优势,走以本土原创性成果对接国际前沿的道路。以本土之“本”开国际之“新”,方能在全球索引学版图中确立中国

索引学的独特地位。

理论层面,应立足中国索引学的历史传统与当代实践,提炼具有中国特色的核心概念与分析框架。中国索引学从明清古籍索引实践到近代索引运动,积累了丰富的本土经验;丰富的文化典籍资源、独特的汉字体系以及国家治理中的特殊数据需求,都构成了国际索引学理论所未能充分覆盖的独特问题域。应在系统整理这些本土资源的基础上,形成可与国际学界对话的原创性理论贡献,例如推动中国古籍索引标准的研制与国际化推广,将中国经验转化为国际学术话语。

在实践层面,应积极拓展与国际索引组织的交流合作机制^[53]。通过引进国外先进技术与教学经验,拓宽研究视野并吸收成熟经验;同时推动国内优秀学者参与国际索引学研究,在国际平台上展示中国索引学的研究成果与实践方案,培养具有国际视野的顶尖索引人才。

5.2.3 知行合一:贯通人才培养与理论实践融合

理论性与实践性的矛盾是当前中国索引学发展的核心瓶颈。理论研究泛而不精,实践转化不畅,二者难以形成良性循环。而人才是打通理论与实践的关键中介——缺乏兼具理论素养与实践能力的复合型人才,理论创新便难以落地,实践经验也难以升华为系统性理论。因此,完善人才培养体系是化解这一矛盾的基础性工程。

学历教育层面,可以推动高校在信息资源管理学科下设立索引学专业硕士方向或微专业,课程设置兼顾传统索引理论与数智技术应用,涵盖智能标引算法、多模态索引编制及数据要素索引等前沿内容,培养兼具理论深度与实践能力的复合型人才。同时,编制适配新时代需求的索引学教科书。借鉴杜定友《图书馆管理学》将索引知识纳入中学师范课程的理念,还可推动索引基础知识进高校、进职业院校,培育“索引思维”^[47]和“索引意识”,提升全社会对索引学的认知度。

实践培育层面,建立高校、科研机构与企业的协同育人机制,通过实习基地和项目制合作等方式,让学生参与真实索引项目,积累实战经验。完善人才评价体系,将索引产品研发、技术

创新、产业服务等成果纳入评价指标，激发科研人员与从业者的创新积极性。

从学历教育到实践培育，人才培养体系的完善旨在消除理论与实践之间的人为隔断，培养能够在二者之间自如切换的复合型人才。当理论研究者具备实践感知，实务工作者具备理论自觉，理论指导实践、实践反哺理论的良好循环才有可能真正建立，理论性与实践性脱节的困境也才能得到缓解。

5.2.4 守正创新：深化数智索引理论与方法创新

原创性与时代性的辩证统一是索引学持续发展的内在要求。原创性要求突破既有理论框架，形成具有中国特色的自主知识体系；时代性要求紧跟数智技术发展与国家战略步伐，主动回应新需求。当前部分研究存在“重传统轻创新、重形式轻内涵”的倾向，对时代新命题的敏感度不足，未能将技术变革及时转化为学术创新的动力。深化数智索引理论与方法创新，正是在守住学科根脉的基础上回应时代命题的核心路径。

理论层面，需在对索引本质属性的辨异基础上^[54]，系统厘清索引与数据库、数据、信息、知识和智能等核心概念的逻辑关联，建立具有学理性与包容性的数智索引理论框架。重点研究多模态数据索引、语义索引与数据要素索引等新领域的底层逻辑，形成智能标引原理、索引与数据产权衔接机制等原创性理论成果。在技术方法层面，推动索引编制从自动化向智能化升级，如积极利用大语言模型，优化中文智能标引算法，解决异体字、通假字和语义歧义等中文处理难题。同时，加强与国家数据局等机构的合作，推动数据索引与元数据标准化建设，编撰全国性数据库总目及其索引^[27]，参与人工智能训练数据集的标引与治理工作，为国家数据基础制度建设提供技术支撑。

守正方能固本，创新才能致远。在坚守索引学核心学术传统的基础上，紧扣数智时代的技术变革与国家战略需求，推出适应新时代的原创性理论与方法成果是实现原创性与时代性辩证统一的根本之道，也是强化索引学在信息资源管理学科核心地位的必由之路。

6 结语

数智时代的技术革新与国家战略需求为中国索引学带来了重构自主知识体系的历史契机。本文系统回溯中国索引学从萌芽奠基、恢复发展到转向深化的三阶段历史流变，梳理其在数智技术浪潮下的演进轨迹，通过对2014—2024年相关文献的计量与BERTopic主题分析，识别出数智时代索引理论智能化、对象垂直化、组织学术化和产业科学化等四大趋向，并在此基础上构建了数智时代中国索引学研究整合框架，揭示了数智时代中国索引学面临的机遇期与危机期、国际化与本土化、理论性与实践性、原创性与时代性等四对核心矛盾，据此提出危中寻机、植本开新、知行合一与守正创新四条发展路径。

本文的贡献主要体现在两方面。其一，突破了以往综述侧重“索引工作”的分析视角，从“索引学”作为一门学科的理论形态与知识结构出发，勾勒了中国索引学跨越五百余年的发展脉络，为理解索引学的学科演进提供了历史连续性视角。其二，聚焦数智时代索引学的趋向特征，以社会技术系统理论等为基础构建了数智时代中国索引学研究整合框架，将四大趋向纳入统一的分析逻辑，揭示了技术、组织与环境之间的协同演化机制，并从矛盾辩证视角提出了系统化的发展路径，为索引学从治学之学向致用之学转型提供了理论参照。

本文也存在一定局限。数据维度上，主要基于中文学术数据库进行文献检索，对国际索引学研究动态的关照相对有限；分析维度上，侧重学术文献层面的考察，对索引产业实践、政策执行效果及用户需求等方面的探讨尚不充分。未来研究可推进索引学核心概念与学科边界的系统厘定，构建具有学理性与包容性的索引学自主知识体系；开展中外索引学学科建制与知识体系的比较研究，在国际参照中提炼中国索引学的理论特色与话语优势；通过调查与案例分析等方法深化对索引产业需求与应用现状的实证考察，为上述体系建构提供经验支撑。

注释

- ①详见：叶继元，朱禹.人机交互编制索引/数据库的原理与模式[M]//中国索引（第十四辑）.北京：中央编译出版社，2026.

参考文献

- [1] 张心源，邱均平.国内外数据库索引编制研究的进展与趋势[J].图书馆杂志，2016，35（3）：60-67.
- [2] 张思龙，王兰成，高宾.基于知识图谱的国内近十年索引研究可视化分析[M]//中国索引（第五辑）.上海：复旦大学出版社，2018：152-164.
- [3] 王逸帆.21世纪以来中国索引理论与实践——基于Citespace的《中国索引》稿件题录信息分析[M]//中国索引（第十辑）.北京：中央编译出版社，2023：52-63.
- [4] 冯鲸洁.索引编制研究热点与主题演化分析[M]//中国索引（第十辑）.北京：中央编译出版社，2023：33-51.
- [5] 张洋，黄子莹，朱嘉麒.近30年国内外数据库索引研究主题对比分析[M]//中国索引（第十一辑）.北京：中央编译出版社，2023：12-26.
- [6] 王彦祥.中国索引已迈入无纸化快捷编撰时代——中国索引学会成立30年来索引编撰技术回顾与展望[M]//中国索引（第十辑）.北京：中央编译出版社，2023：3-20.
- [7] 尹璐，吴东岳.我国科学数据索引建设研究与实践[M]//中国索引（第十辑）.北京：中央编译出版社，2023：21-32.
- [8] 王彦祥.新世纪中国索引编纂与研究述评（上）——基于2000—2017年索引成果的定量研究[M]//中国索引（第五辑）.上海：复旦大学出版社，2018：37-60.
- [9] 王彦祥.新世纪中国索引编纂与研究述评（下）——基于2000—2017年索引成果的定性研究[M]//中国索引（第六辑）.上海：复旦大学出版社，2019：3-47.
- [10] 郑德俊，程为.基于三维主题特征测度的新兴主题识别研究[J].情报学报，2024，43（2）：167-180.
- [11] 章学诚.校雠通义通解[M].上海：上海古籍出版社，2009.
- [12] 潘树广.古籍索引概论[M].北京：书目文献出版社，1984.
- [13] 张西平.《中国丛报》篇名目录及分类索引[M].桂林：广西师范大学出版社，2008.
- [14] 平保兴.西方索引早期在中国的传播研究[C]//2009年中国索引学会年会暨学术研讨会论文集.兰州：兰州大学，2009：3-6.
- [15] 林语堂.创设汉字索引制议[J].科学，1917，3（10）：1128-1135.
- [16] 胡适.国学季刊发刊宣言[J].国学季刊，1923，1（1）：1-16.
- [17] 万国鼎.索引与序列[J].图书馆学季刊，1928（2）：3.
- [18] 熊静.索引运动与索引学说的建立[J].图书情报知识，2016（4）：27-36.
- [19] 王余光.索引运动的发生[J].出版发行研究，2003（6）：74-76.
- [20] 向科学进军中的图书馆工作[N].人民日报，1956-08-28（1）.
- [21] 潘树广.二十世纪的索引编纂与研究（代序）[M]//卢正言.中国索引综录.上海：上海辞书出版社，2000.
- [22] 印永清.中国索引发展史略[J].图书与情报，2007（1）：19-22.
- [23] 邱均平，楼雯.我国索引研究二十年回顾与展望——纪念中国索引学会成立20周年[C]//2011年中国索引学会年会暨成立二十周年庆典论文集.济南：山东师范大学图书馆，2011：3-25.
- [24] 叶继元.《中文图书引文索引·人文社会科学》示范数据库研制过程、意义及其启示[J].大学图书馆学报，2013，31（1）：48-53.
- [25] 张琪玉.现代的索引就是数据库[J].图书馆杂志，2001（12）：6-7.
- [26] 张琪玉.情报语言漫笔（L）[J].图书馆理论与实践，2003（6）：47-49.
- [27] 叶继元，陶蕊.智能检索下的数据库书目控制探讨[J].图书馆学研究，2024（4）：69-75.
- [28] 王雅戈，叶继元，黄建年，等.中文索引平台建设——以“索引家”开发为例[J].图书馆论坛，2019，39（11）：37-40.
- [29] 王兰成.基于Python设计的索引智能标引与检索研究——以《张琪玉索引学文集》和《中国索引（第一辑）》为例[M]//中国索引（第七辑）.上海：复旦大学出版社，2020：3-15.
- [30] 张思龙，蒋璵，王兰成.基于知识图谱的智能索引技术研究[M]//中国索引（第七辑）.上海：复旦大学出版社，2020：16-24.
- [31] 王弘毅.智能索引系统的功能与优化研究——基于《中国索引》数据库索引相关论文文本的分析[M]//中国索引（第七辑）.上海：复旦大学出版社，2020：25-33.

- [32] 张敏, 丁良萍, 刘欢. 面向科技文献的多维语义索引构建思路及实现[J]. 情报理论与实践, 2021, 44 (8): 139-145.
- [33] 朱禹, 叶继元. 生成式人工智能驱动的索引编制方法及其在学术规范和评价中的应用[J]. 图书馆杂志, 2024, 43 (10): 50-59.
- [34] 刘丹, 叶继元, 袁曦临, 等. 史料索引智能编制实验与效果测评研究[J]. 图书馆学研究, 2025 (11): 67-80, 124.
- [35] 段逸山. 中医典籍词语索引的编制及其作用[M]//中国索引 (第八辑). 上海: 复旦大学出版社, 2020: 3-14.
- [36] 王昊, 王蕾. 中华古籍活化的实践框架与发展路径[J]. 中国图书馆学报, 2024, 50 (4): 4-20.
- [37] 范炜, 曾蕾. AI新时代面向文化遗产活化利用的智慧数据生成路径探析[J]. 中国图书馆学报, 2024, 50 (2): 4-29.
- [38] 岳文玉, 曹树金, 赵怡琬. 基于知识图谱的中共党史知识索引设计与实现[J]. 图书馆论坛, 2025, 45 (3): 110-119.
- [39] 叶继元. 教学科研的得力工具——美国《科学评论索引》评述[J]. 中国科技期刊研究, 1996, 7 (2): 57-58.
- [40] 鲍国海. 《胡适日记全集》《顾颉刚日记》《顾颉刚年谱》三书索引评析[M]//中国索引 (第六辑). 上海: 复旦大学出版社, 2019: 180-195.
- [41] 傅德华. 方志敏烈士传记资料的整理与研究——以《20世纪中国人物传记资料索引》为中心[M]//中国索引 (第七辑). 上海: 复旦大学出版社, 2020: 145-156.
- [42] 王彦祥. 论志书索引对方志信息的有效揭示[M]//中国索引 (第八辑). 上海: 复旦大学出版社, 2020: 15-24.
- [43] 杨光辉, 温国强. 中国索引学会工作回顾与展望: 2012-2015——在第六届国际索引联盟峰会上的发言[J]. 中国索引, 2015, 13 (4): 2-4.
- [44] 郭丽芳, 温国强. 张琪玉教授对中国索引学会和中国索引事业的贡献[J]. 图书馆杂志, 2014, 33 (9): 18-24, 65.
- [45] 饶帮华, 何道刚, 谭熠. 浅谈学术著作出版技术规范——从引文、注释、参考文献、索引的规范谈起[J]. 科技与出版, 2014 (8): 72-75.
- [46] 孙涵涵, 王雅戈, 薛春香, 等. 国家标准《学位论文内容索引编制规则》制定: 缘起、作用与展望[J]. 图书馆论坛, 2023, 43 (7): 68-74.
- [47] 李彤. 关于编辑人员提升索引编制能力的若干思考[M]//中国索引 (第四辑). 上海: 复旦大学出版社, 2018: 67-76.
- [48] 王瑾, 陈春霞, 侯利娟. 中国学术著作书后索引的现状、成因与趋势——基于专家调查法的微调查[J]. 中国索引, 2015, 13 (2): 7-20.
- [49] 张丽, 王雨菲. 对张琪玉索引员资格培训和认证设想的进一步思考[M]//中国索引 (第九辑). 北京: 中央编译出版社, 2021: 66-76.
- [50] 刘苏南, 李华. 索引走向社会的倡导者——刘苏南先生谈黄页与中国索引学会的缘分[M]//中国索引 (第九辑). 北京: 中央编译出版社, 2021: 79-89.
- [51] 弗朗西丝·伦尼, 伏安娜. 美国索引协会: 过去、现在与未来——美国索引协会前主席 Frances·S·Lennie 女士在中国索引学会第四次全国会员代表大会暨学术研讨会上的演讲[J]. 中国索引, 2013, 11 (4): 58-61.
- [52] 叶继元. “万事万物皆可索引”: 可能性、可行性与实例分析[M]//中国索引 (第四辑). 上海: 复旦大学出版社, 2018: 3-10.
- [53] 叶继元. 索引研究的现状与走向——记2009年澳大利亚和新西兰索引学会国际研讨会[J]. 图书情报工作, 2010, 54 (8): 5-9.
- [54] 叶继元. 索引的本质属性及其在学术规范与评价中的作用[J]. 图书情报知识, 2023, 40 (6): 61-67.

作者简介 朱禹, 朱雨萌, 南京大学信息管理学院博士研究生; 叶继元 (通信作者, yejiyuan@nju.edu.cn), 南京大学信息管理学院特聘教授、博士生导师, 中国索引学会副理事长。

收稿日期 2026-01-13

(责任编辑: 刘洪)