多源异构数据视角下的学术评价:内涵、进展与展望*

邓三鸿[1] 杨 杰[1] 王 昊[1] 沈健威[2]

1.南京大学信息管理学院,南京 210023
2.江苏省质量和标准化研究院,南京 210023

摘 要:[目的/意义]综合多源异构数据的学术评价与科学计量方法逐渐发展成为领域热点,目前对于多源异构数据视角下的学术评价缺乏系统的理论探讨。[方法/过程]针对多源异构数据利用不充分、研究方法不妥当、研究目标不明确等问题,文章梳理多源异构数据视角下的学术评价发展状况、厘清领域内涵及外延、分析领域前沿和热点。[结果/结论]研究发现,在多源异构数据视角下,学术评价领域呈现出数据维度增加、数据粒度细化、评价指标复杂化等趋势。多源异构数据视角下的学术评价内涵界定、现状分析与未来展望,对于澄清领域认知、辨析相关概念、明确研究路径与方法论基础,以及构建中国特色的学术评价体系有着重大意义。

关键词:学术评价;科学计量;多源异构数据

中图分类号:G203 **文献识别码:**A **文章编号:**2096-7144(2023)04-0042-15

DOI:10.19809/j.cnki.kjqbyj.2023.04.004

1 引言

学术评价是科研活动的重要环节,在科学知识信息传播方面有着重大影响,同时,由于它能反映研究质量与研究水平,通常也与学术资源配置挂钩。学术评价在推动科技创新和社会进步方面发挥着重要作用,受到国家相关部门的高度重视,我国先后出台《关于深化项目评审、人才评价、机构评估的意见》[1]《关于优化科研管理提升科研绩效若干措施的通知》[2]《关于开展清理"唯论文、唯职称、唯学历、唯奖项"专项行动的通知》[3]以及《关于规范高等学校SCI论文相关指标使用树立正确评价导向的若干意见》[4]等政策,以期优化学术评价方法、合理分配学术资源,引导构建具有中国特色的全方面学术评价体系。

随着以数据为导向的科学研究第四范式的出现,学术评价领域迎来新的发展。在"开放获取"运动的背景下,越来越多的数据融入学术评价领域,这些数

收稿日期:2023-04-27 修回日期:2023-06-16
基金项目:国家社科基金项目"大数据环境下学术成果真实价值与影响的实时预测及长期评价研究"(编号:19BTQ062)
作者简介:邓三鸿(ORCID:0000-0002-6910-3935),男,1975年生,博士,教授,博士生导师,主要研究方向:知识图谱、科学计量、信息分析与科学评价,E-mail:sanhong@nju.edu.cn;杨杰,男,2000年生,博士研究生,主要研究方向:科学计量、学术评价、信息分析,E-mail:alexjieyang@outlook.com;王昊(ORCID:0000-0002-0131-0823),男,1981年生,博士,教授,博士生导师,主要研究方向:文本挖掘,知识图谱,E-mail:ywhaowang@nju.edu.cn;沈健威,男,1985年生,大学本科,高级工程师,主要研究方向:大数据应用与管理,E-mail:shjianw@163.com。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

据呈现出多源异构的特征。因此，开展多源异构数据视角下的学术评价进展梳理与理论探讨尤为重要。

数据的多源异构一般表现为数据来源、数据结构、存储模式、数据标准、平台及应用的多样性。随着大数据、人工智能、复杂网络和数据科学的发展，多源异构数据的融合与应用被广泛研究，学术评价领域也出现了基于多源异构数据视角的探索。多源异构数据视角下的学术评价拓展主要分为两大类型：基于多源数据的学术评价新指标，以及基于异构数据抽取的学术评价新方法。

目前，学术评价普遍存在对多源异构数据的利用不充分、研究方法不科学、研究目标不明确等问题。本文首先对学术评价的历史沿革及内涵特征进行分析探讨，总结多源异构数据视角下的学术评价新框架及研究热点。然后，从“基于多源数据的学术评价新指标”以及“基于异构数据抽取的学术评价新方法”两大维度出发，对多源异构数据视角下的学术评价发展进行系统梳理与总结。最后，对多源异构数据视角下的学术评价进行展望。

2　多源异构数据视角下的学术评价内涵及发展

2.1　学术评价的历史沿革及内涵特征

学术评价起源于同行评议和专家评价。同行评议的历史可以追溯到1665年英国皇家协会设立的*Philosophical Transactions*期刊。从1752年起，该协会开始邀请与期刊领域相关的学者，为文章提供建议和审查，这被认为是学术期刊的首个同行评议制度[5]。同行评议可以划分为“单盲”（Single-blind）评审制度和“双盲”（Double-blind）评审制度，“单盲”评审制度中作者不知道评审人的信息，“双盲”评审制度中作者和评审人互不知道对方的信息。随着学术规范的完善，同行评议逐渐发展成为学术论文发表的必要程序。开放同行评议（OPEN PEER REVIEW）是在互联网平台兴起的新定性学术评价方法，且更具效率和透明度[6]，但目前仍处于起步阶段，待进一步发展和研究。传统的学术评价也受到了一些批评，如PAULSON等[7]认为同行评议制度不能保证评审人对论文的主题完全了解，SILER等[8]认为专家评价和同行评议难以识别具有创新性的论文。

定量的学术评价起源于文献计量学，OTLET P是首位使用文献计量方法进行学术评价的学者[9]，他在1934年出版的文章*Traite de Documentation*中使用了该方法，但直到35年后，文献计量学才被PRITCHARD正式定义[10]。GARFIELD[11]提出影响因子（Impact Factor，IF），并创立科学引文索引（SCI），从此引文分析成为评价期刊质量的重要标准[12]。近年来，定量学术评价指标不断被提出，HIRSCH[13]于2005年提出了H-index，用于学者、期刊和地区等维度的评价。NOORDEN在2016年提出了CiteScore，拓展了引文时间窗口[14]，目前已成为Scopus对出版物评价的主要指标之一[15]。在机构预算紧张、竞争激烈的当下，基于引文数量的评价指标在大学排名、学术期刊评估、同行评议等方面均发挥着重要的作用[16-17]。然而，过度强调引文数量和论文数量也会导致学术不端等诸多问题。如影响因子的结果难以跨学科比较，且受自引影响较大，受到了广泛的批评[18]。H指数等评价指标也被认为导致了学术马太效应，不利于学术创新，受到了部分学者的反对和批评[19]。

学术评价是历史性的评价，常常发生在学术共同体之间，它也可被视作是学术群体之间的社会舆论[20]。学术评价的客体是作者、期刊、机构和地区等学术实体；主体一般是其他学术实体，通常为客体的学术共同体。学术评价的方法通常包括同行评议、专家评价等定性方法，基于文献计量指标的定量方法，或二者相结合的综合评价方法。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

苏新宁等[21]指出，学术评价可以分为3个层面：第一层是基于计量指标对学术实体的评价；第二层是探索学术研究规律、学科特征及关系、研究热点及领域；第三层在此基础上探索未来发展趋势，引领科学研究。

综上所述，本文将学术评价的内涵界定为：采用同行评议、专家评价以及文献计量等评价方法，对作者、期刊、机构和地区等学术实体在科学发展理论或实践影响程度等方面进行科学客观的度量，从而促进学术活动健康发展、推动学术继承与创新，最终对社会进步与科技创新产生贡献。

2.2 多源异构数据视角下的学术评价新框架

针对二元对立的定性和定量评价存在的“六化”问题，叶继元[22]提出了形式评价、内容评价和效用评价的“三位一体”框架，结合多源异构数据构建了专家评审与学术共同体相关数据库，全评价体系是定性学术评价与定量学术评价的有机结合，也是建设中国哲学社会科学特色评价的一个重大突破[23]。在异构、多源和海量的数据背景下，邱均平等[24]提出并拓展了“五计学”的评价框架，并将数据科学与计量学结合，提出了“数据计量学”的概念和内容体系[25]，融合了数据与计量两大范式，是多源异构数据视角下学术评价领域极具前景的发展方向。苏新宁等[21]7-9、邓三鸿等[26]、张宝隆等[27]和闫晓慧等[28]在既有评价范式和CSSCI的基础上，综合多源异构数据、网络影响力和社会影响力评价方法，探索出了人文社科多层次评价法、组合赋权法评价法、差异性视角下的期刊区分度以及研究创新度测度等，可用于论文、期刊、学科、高校等多层次的评价，从而形成中国人文社科领域的评价新框架。

除以上较为成熟的框架以外，张琳等[29]探索的交叉学科评价体系、赵蓉英等[30]探索的学术话语权评价框架、王贤文等[31]探索的动态复合评价模式、罗卓然等[32]探索的结合深度学习方法的创新评价方法，以及杨思洛等[33]探索的结合社交媒体数据多维测度指标等，都是基于多源异构数据视角下的学术评价新框架的研究和探讨。未来的学术评价应当综合这些框架体系，逐步构建形成健康完整的评价生态。

2.3 多源异构数据视角下的学术评价热点

在CNKI上检索含有“多源”“异构”等关键词，并包含“文献计量”或“学术评价”的文献，剔除停用词后基于VOSViewer-1.6.18进行关键词共现分析，结果如图1所示。多源异构视角下的学术评价以科学计量和信息计量为主体，在社会媒体数据、期刊开放获取、专利文献、政策文献、网络分析方法和人工智能技术的驱动下，呈现出多元化的发展趋势。一方面，基于多源数据产生了替代计量和网络计量等新的评价指标；另一方面，随着人工智能、深度学习和数据科学等方法的引入，可通过异构数据的抽取实现学术评价的客体维度补充。因此，本文将从两大角度出发，分别对基于多源数据的学术评价新指标，以及基于异构数据抽取的学术评价新方法的相关研究进行系统地梳理，以探究多源异构视角下学术评价的发展状况。

3 基于多源数据的学术评价新指标

多源数据反映了数据来源的多样性和丰富性。学术交流和知识共享的载体不仅局限于学术论文，还包括社交媒体数据、专利文献数据和政策文本数据等。在信息时代下，各种类型的文献呈现出爆炸式增长，基于多源数据网络分析的评价指标是一种面向未来的学科拓展思路。

3.1 基于社交媒体数据的评价指标

近年来，基于社交媒体数据的科学计量与学术评价迅速发展，并成为了研究热点，社交媒体数据计量与评价领域中的部分代表性研究，见表1。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

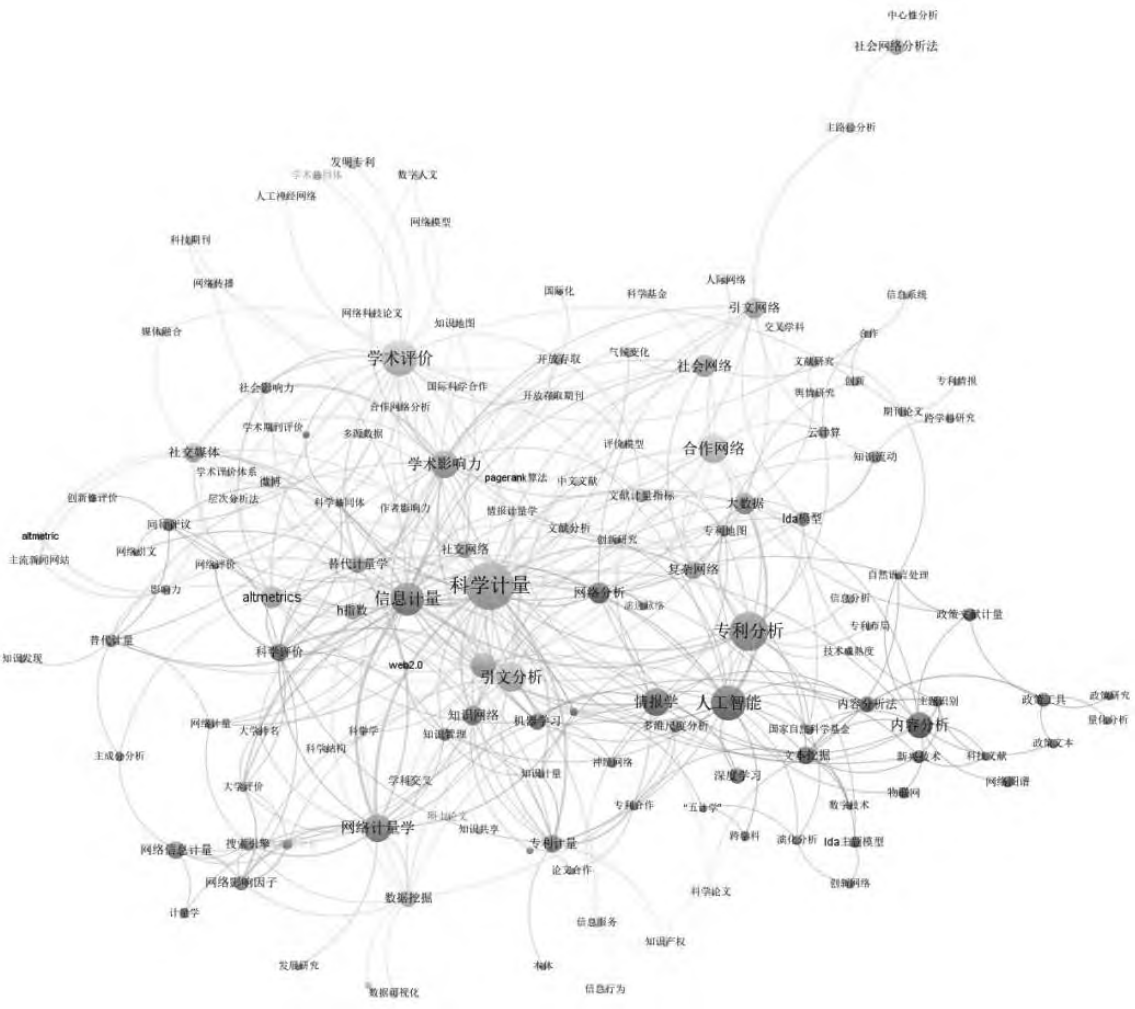


图1　多源异构数据视角下的学术评价热点

表1　社交媒体数据计量与评价领域的部分代表性研究

领域类别	相关领域	代表性学者	年份
替代计量学：Altmetrics的提出		PRIME J等	2010
Altmetrics的代表性数据源	社交网络、书评、博客、微博、维基百科和数据分享	GU F等	2011
	会议记录和日程数据	ROWLANDS I等	2011
	聚合新闻服务及用户评论数据	TENOPIR C等	2013
基于Altmetrics的拓展指标	Altmetric score	ADIE E等	2013
	期刊社会影响力因子	ALHOORI H等	2014
	社交媒体计量指数	THOMA B等	2014

2010年，PRIEM等[34]基于社交媒体数据提出了全新的信息计量指标——Altmetrics。社交媒体数据涵盖了社交网络、书评、博客、微博、维基百科和数据分享[35]，也有部分学者将会议记录和日程[36]、聚合新闻服务及用户评论[37]作为社交媒体数据。此后，基于Altmetrics的评价方法陆续被提出，如ADIE等[38]提出Altmetric score，ALHOORI等[39]提出期刊社会影响力因子，以及THOMA等[40]提出社交媒体计量指数。随着Altmetrics的补充与拓展，社交媒体数据如今已经成为了学术评价领域不可忽略的数据来源。BORNMANN[41]认为随着对社交媒体数据的不断强调和更广泛地使用，将会对文献计量学领域产生重大

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

变革。社交媒体可以提高学术交流的透明度以及加速学术的传播[42]，相关学者提出利用社交网络数据对学术社会影响力进行评价[43]。

Altmetrics在取得较大的成果和影响力的同时，也受到了一些批判。首先，Altmetrics在评价一致性、评价效率、评价专业性、可信程度等方面尚有改进空间[44]。其次，Altmetrics在与同行评议的结合方面也有待提高[45]。同行评议作为学术评价的重要环节之一，可以提供更加专业、全面和深入的评价。因此，未来的研究可以探讨如何将Altmetrics与同行评议相结合，以更好地评价学术成果。

为此，我们认为融合社交媒体数据的学术评价应注意以下几个问题：一是，为提高评价一致性，可以尝试建立统一的数据收集和处理标准，以确保不同来源的数据质量和可靠性。此外，通过加权或归一化方法综合多个来源的数据，也可以降低评价结果的不一致性。二是，为提高评价效率，可以借助现有的计算机技术和算法，如分布式计算、机器学习等，对大量的社交媒体数据进行高效处理和分析。此外，还可以利用自然语言处理技术，对社交媒体中的文本内容进行深入挖掘，以提取更有价值的信息。三是，为提高评价专业性，可以考虑在评价过程中引入领域专家的意见，以消除非专业性信息的干扰。同时，可以研究如何将领域专家的评价与社交媒体数据相结合，以实现更全面、准确的学术评价。四是，为提高可信程度，可以设计和实施相应的数据验证机制，以检测和剔除潜在的恶意操纵或虚假信息。此外，通过对比分析不同来源的数据，也可以提高评价结果的可信度。五是，为促进Altmetrics与同行评议的结合，未来可以探索建立一个综合评价体系，将社交媒体数据、同行评议以及其他相关数据进行整合，以实现对学术成果的全面评价。此外，可以研究如何利用社交媒体数据辅助同行评议，例如挖掘潜在的评审专家、发现学术热点等。

3.2 基于多源数据网络分析的评价指标

Web of Science、Google Scholar、Scopus等数据库为学者提供了广泛的文献数据，从多源数据中抽取节点，以引用关系作为链接，形成了巨大的引文网络。网络分析方法可针对学术期刊文献、会议文献、专利文献、政策文献以及社交媒体数据等多源数据进行深入的评价探索。基于多源数据网络分析的学术评价领域代表性研究，见表2。

表2 基于多源数据网络分析的学术评价领域代表性研究

领域类别	相关领域	学者	年份
提出网络计量学：Webmetrics		ALMIND T C等	1997
引文网络的社群研究	提出复杂网络的社群	NEWMAN M E J	2003
	总结社群在引文分析中的作用	FONTOURA C L	2011
多源数据在引文网络中的应用		VAN-ECK N J等	2014
引文网络中的评价指标	特征因子（Eigenfactor）	BERGSTROM C T等	2008
	学术迹（Academic Trace）	YE F Y等	2014
	RCR（Relative Citation Ratio）	HUTCHINS B I等	2016
基于引文网络的创新评价指标	颠覆性指数（Disruption Index）	FUNK R J等	2017
	文献网络中颠覆性指数	WU L等	2019
	网络多维度评价指标	BU Y等	2021
	改进颠覆性指数	LEYDESDORFF L等	2021
	X指数	LATHABAI H H等	2021

ALMIND等[46]提出了网络计量学（Webmetrics），引发了基于引文网络评价的热潮，引文网络类似于复杂网络，具有无标度、小世界等特征。NEWMAN等[47]提出了复杂网络中的社群，FONTOURA[48]总结了

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

其在引文分析中的应用。随着多源数据的融入，引文网络中的合作关系、共被引、关键词共现等现象得到广泛研究[49]。BERGSTROM 等[50]提出了特征因子（Eigenfactor），它是网络中一种类似 PageRank 算法的评价指标，通过复杂的迭代算法计算引文网络中论文的影响力。YE 等[51]基于学术矩阵提出了学术迹（Academic Trace），用于评价学者、机构或地区的科研水平。HUTCHINS 等[52]基于引文网络构建了 RCR（Relative Citation Ratio），可以在引文网络中识别出共被引聚类簇，以评价专利或论文的影响力。

在多源数据的视角下，网络计量新指标通常将拓展数据来源和评价方法细化结合。FUNK 等[53]在专利文献中提出了颠覆性指数（Disruption Index），基于多层次的引用网络分布衡量某一专利是发展还是颠覆了现有的技术趋势。WU 等[54]将其拓展到科学文献领域，并提出文献网络中颠覆性指数。BU 等[55]将颠覆性指数分解为多个评价维度：引文数量、引文深度、引文广度、引文独立性。LEYDESDORFF 等[56]对颠覆性指数进行了实证和改进，充分证明了该指数的有效性。LATHABAI[57]基于 NLP 技术和网络分析方法提出了拓展维度的 X 指数，用以评价机构的学术水平，它包含 2 个评价部分：核心能力和主题研究水平。核心能力反映其在某学科的研究广度，主题研究水平反映其在特定主题的研究深度。

基于多源数据网络分析的评价指标为学术评价提供了新的视角和方法。然而在实际应用中，这些评价指标仍存在一些挑战和不足之处。

首先，多源数据网络分析的评价指标面临数据质量和可靠性的问题，不同数据源之间可能存在数据重复、缺失或不一致的情况，这将影响评价指标的准确性。

其次，多源数据网络分析的评价指标在有效性和稳定性方面存在挑战。由于多源数据的复杂性和动态性，评价指标可能受到不同因素的影响，不可避免地导致评价结果的波动。

为此，基于多源数据网络分析的学术评价可以从以下几个方面来进行完善。

在数据处理阶段，应该关注数据质量和可靠性问题，通过数据清洗、去重、缺失值处理等方法，提高数据质量。此外，可以尝试构建统一的数据模型，以整合不同来源的数据，提高评价的一致性。

在评价指标的设计和改进过程中，应关注指标的有效性和稳定性。例如，可以尝试引入权重因子，以平衡不同数据源的影响；或者采用融合算法，结合多个评价指标，提高评价结果的稳定性。

同时，未来的评价模式应结合同行评议和多源数据的量化评价，以实现更全面、准确的学术评价。同行评议可以提供更专业、深入的评价意见，而多源数据的量化评价可以补充更多维度的信息。因此，可以尝试探索如何将同行评议与多源数据网络分析相结合，以构建完整的评价生态。

4　基于异构数据抽取的学术评价新方法

信息时代带来的数据不规范性进一步加剧，为学术评价带来了机遇与挑战。基于异构数据抽取的学术评价拓展一般是将既有评价指标作为预测变量，从异构数据中挖掘出文献特征的外部特征和内部特征。传统的学术评价一般仅基于文献外部特征，文献的内部特征通常具有更细的挖掘粒度。随着“开放获取”运动的兴起，全文本数据较以往更易获取，不少学者开始研究文献的内部特征，对异构数据的深层内容以及深层动机进行抽取，以拓展评价维度。

4.1　异构数据下的深层内容抽取

词、公式、图表以及结构目录等是文献内容的重要组成部分，但由于这些深层内容难以直接获取，且相关数据源难以获取或长期处于缺失状态，传统的

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

学术评价指标鲜有涉及。在“开放获取”运动的背景下，大量可用于学术评价的异构数据得以获取，基于异构数据的深层内容抽取也在人工智能的助力下逐渐发展起来，相关代表性研究，见表3。

表3 异构数据下的深层内容抽取代表性研究

类别	方法或模型	对象	学者	年份
词抽取	基于字典或规则	期刊论文全文本	SODERLAND S	1999
	Google BERT模型	期刊论文全文本	陆伟等	2020
	Encoder-Decoder机制	期刊论文全文本	YANG L等	2019
	分类和排序模式	期刊论文全文本	姜艺等	2021
语义抽取/实体识别	语义建模与计算（系统）	会议或期刊论文标题	KONDO T等	2009
	核支持向量机SVM	专利和论文全文本	NANBA H等	2010
结构/目录/元数据抽取	BiLSTM-CRF和BiGRU-HAN模型	专利文献全文本	CHEN L等	2020
	无监督算法	会议和社交媒体文章	DOUCET A等	2017
	RCNN方法	不同科学出版物的文献	SAFDER I等	2018
图表抽取	Seed-Guided Topic（STM）模型	不同科学出版物的文献	LI C等	2016
	远程监督神经网络模型	PDF文档	SIEGEL N等	2018
	知识抽取系统	不同科学出版物的文献	AL-ZAIDY R A等	2017

早期SODERLAND[58]基于字典或规则的方法对期刊论文进行了小规模的词抽取，但随着文献数量的大幅增长，传统的词抽取方法难以满足科研需求。近年来，深度学习尤其是预训练模型的迅速发展，使得大规模实现词汇粒度的文献内容抽取成为可能，已有学者基于Google BERT模型[59]、Encoder-Decoder机制[60]和分类和排序模式[61]对期刊论文的文本词汇的自动抽取进行了研究。随着专利文献、会议论文、图书和社交媒体数据为主的异构数据融入学术评价的范畴，内容抽取出现了新进展。

KONDO等[62]基于语义建模与计算方法构建了词识别系统，对不同类型学术文献的词语义进行了抽取和分类。NANBA等[63]基于SVM技术对专利文本以及学术文本内容语义进行了识别。CHEN等[64]利用BiLSTM-CRF和BiGRU-HAN模型对专利文献的信息进行实体识别和语义关系提取。DOUCET等[65]采用无监督算法对会议和社交媒体文章的结构和目录进行了抽取。SAFDER等[66]基于RCNN方法从不同科学出版物中提取特定元数据。LI等[67]利用机器学习方法对各种文献的图表进行了抽取，SIEGEL等[68]使用远程监督神经网络模型从PDF文档中提取图表。AL-ZAIDY等[69]基于知识抽取系统提取柱状图等特定类型的图表数据。

深层内容抽取在学术评价中具有重要意义，它为评价指标提供了丰富的信息来源。然而，在实际应用中，异构数据下的深层内容抽取仍面临一些挑战。首先，异构数据的多样性和复杂性增加了深层内容抽取的难度。不同类型的数据具有不同的结构和表达方式，需要针对性地设计抽取方法和技术。此外，数据质量和可靠性问题也会影响深层内容抽取的准确性和有效性。其次，深层内容抽取的算法和技术仍有待进一步改进。虽然已有许多研究采用了深度学习、预训练模型等先进技术，但在某些复杂场景下，深层内容抽取的准确性和鲁棒性仍有待提高。

要解决这些挑战，促进异构数据下的深层内容抽取研究，需要重视以下几个问题。

首先，要重视异构数据的多样性和复杂性问题，并尝试提出通用性更强、适应性更好的抽取方法。例如，可以考虑采用迁移学习、元学习等技术，以提高抽取方法在不同数据源和场景下的适用性。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

其次，在深层内容抽取的相关算法和技术研究中，应适当关注准确性和鲁棒性问题。可以尝试采用多模型融合、强化学习等技术，以提高抽取结果的准确性。同时，可以研究对抗性训练、模型压缩等方法，以提高模型的鲁棒性。

同时，为了提高深层内容抽取的实用性和应用价值，应关注其在学术评价中的具体应用场景。可以尝试探索如何将深层内容抽取的结果与其他评价指标相结合，以实现更全面、准确的学术评价。

4.2　异构数据下的深层动机抽取

学术论文的引用是定量学术评价的主要方法介质之一，传统学术评价指标一般赋予所有引用相同的权重，但引用往往出于不同动机，GARFIELD[70]将引用动机划分为15个类别，BROOKS[71]总结出7大引文目的。在多源异构数据开放获取的背景下，引用目的、极性与语义提取为学术评价提供了新的拓展方向。异构数据下的深层动机抽取的部分代表性研究，见表4。

表4 异构数据下的深层动机抽取部分代表性研究

类别	方法或模型	学者	年份
提出引文论证区域的概念	—	TEUFEL S等	1999
引文极性和引文目的识别	Word2Vec模型和CNN模型	HUANG H等	2021
引文分布情况识别	NLP技术	DING Y等	2013
自动引文权重计算框架	机器学习和回归方法	WAN X等	2014
引用动机识别和分类	机器学习方法	EBRAHIMY S等	2014
引文内容标注框架	—	陆伟等	2014
引文逻辑层次分析	深度学习方法	AN J等	2017

TEUFEL[72]提出引文的论证区域概念，为引文内容的自动识别做了铺垫。HUANG等[73]将Word2Vec模型和CNN模型结合，以识别引文极性和引文目的，从而更真实地评价高被引论文的影响力。DING等[74]基于NLP技术分析了文献文本中的引文分布情况。WAN等[75]构建了自动引文权重计算的框架，用机器学习和回归的方法赋予不同引文相应的权重。EBRAHIMY等[76]在规范性、信息性和识别性3个层次上，基于机器学习方法建立了分析引用一致性的工具，能够对作者的引用动机进行识别和分类。陆伟等[77]对LDA主题模型的相关文献的引文内容标注框架进行了探索。AN等[78]利用深度学习方法，对引文位置、前后文内容和逻辑层次进行分析，并基于引文逻辑层次分析出了高被引作者的不同学术特征。

目前对于异构数据的深层内容和深层动机抽取仍然处于探索阶段，研究以小规模实证为主，尚未完全融入成熟的评价体系中。具体面临的挑战包括：①当前的引文动机抽取方法仍需要在算法和技术上进行改进，可以尝试采用更先进的深度学习模型，以提高引文动机抽取的准确性和鲁棒性。②在异构数据下进行深层动机抽取时，如何有效地融合不同类型的数据源是一个重要挑战，可以探讨更有效的数据融合方法，以提高引文动机抽取的综合性能。③当前的深层动机抽取研究主要集中在英文文献，可以拓展到中文领域和跨领域应用，以实现更广泛的学术评价。④为了更好地评估深层动机抽取的有效性，有必要建立一套完整的评价指标和标准，这将有助于推动该领域的发展和应用。⑤深层动机抽取在实际应用场景中的应用还有待加强，可以探讨如何将深层动机抽取的结果与其他评价指标相结合，以实现更全面、准确的学术评价。未来随着人工智能和语义网络进一步发展，有望建立完整、高效、科学、系统的提取模式，并成为学术评价生态的重要组成部分。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

5 总结与展望

在多源异构数据视角下,学术评价产生了许多新趋势:一是数据维度的增加,以社交媒体数据和网络数据为主的多源数据逐渐融入学术评价领域,与数据科学、人工智能、云计算和语义计算等方法相结合,引发了网络计量学、替代计量学和数据计量学的迅速发展,拓展了评价的深度和广度。二是数据和信息来源的粒度细化,一方面,异构数据本身具备更复杂的数据层次,从而能够挖掘更细的粒度;另一方面,在深度学习的助力下,全文本数据、词、公式、图表、目录和引用动机等深层数据被提取,并广泛应用于学术评价。三是评价体系的复杂化,在多源异构数据背景下,学术评价的方法、主体和客体都得到了拓展,单一维度的评价已然成为过去式,新的学术评价体系一般基于细粒度数据或深层引证网络,考虑多样化和创新性的方法,将定量评价方法与定性评价方法有机结合,逐渐发展为完整的评价生态。

总体看来,尽管在多源异构数据的支撑下,学术评价相关研究取得了明显进展,但目前依然存在过度量化、方法运用不当以及唯引文量等诸多问题。未来基于多源异构数据的学术评价,应该从评价侧重点、评价方法拓展、评价生态构建和保持中国特色4个方面进行探索。

5.1 注重创新和未来发展

论文的创新价值评价是当下的研究热点,但目前仍然缺乏一种有效的、认可度高的手段评价论文的创新性。结合多源异构数据,综合评价论文的创新价值,仍是本领域亟需解决的一大难题,也是解决“破五唯”问题的关键所在。学术评价的最终目的是促进科学进步和社会发展,因此,学术评价要兼顾未来科学发展方向。根据科学发展的一般规律进行推导,学术评价的结果可以用于预测未来的方向和趋势。未来在人工智能、语义网络和多源异构数据融合的助力下,学术评价很可能将产生新的范式,即通过学术评价预测未来学科的发展趋势。

5.2 合理运用大数据与深度学习

大数据时代下,基于社交网络的学术评价、基于网络分析的学术评价以及基于全文本分析的学术评价逐渐兴起。在面对海量数据进行学术评价研究的同时,也要注意到大数据的缺点,如数据欺骗性、资源利用率低等。适当且合理地运用大数据的优势,将会给学术评价的发展带来巨大的动力。深度学习的快速发展为学术评价带来了新的驱动力,但也有部分学者盲目将其应用于学术评价各个领域,难以产出具备现实意义的成果。目前深度学习在学术评价的应用集中在文本内容的抽取、识别、特征分析以及引用极性和语义分析等方面,且以小规模实证为主。未来应当结合数据挖掘、语义网络、云计算以及认知计算方法,追求更加高效、科学、系统的方式,并从不同维度探索深度学习在学术评价领域的新应用前景。

5.3 建设多元评价生态体系

面向未来的学术评价应该是多元的,将定量评价和定性评价相结合,并逐渐构建一个完整的评价系统和模式。如ZHAO等[79]提出的“数智人”评价体系,以多源异构数据客观表征,以人工智能技术全面认知,以领域专家经验智慧深入揭示,融合数据、智能、专家和决策,是对未来智能化学术评价成熟模式的一种有效探索。面向未来的评价生态应当将专家思维和数据逻辑相结合,综合利用数据科学、深度学习、语义计算和云计算等技术,深度融合定性评价和定量评价等方法,科学探讨不同学科评价的方式与权重,综合提升学术评价的效率和效果。

5.4 坚持中国特色的学术评价新模式

未来学术评价体系的构建要结合中国科学研究

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

的具体国情，打造出具有中国特色的全面科学评价体系。具体而言，未来的学术评价要紧密结合国家的发展战略和社会需求，利用学术评价为国家和社会的重大问题提供解决方案。同时，基于学术评价的过程和结果，发现新的社会需求并转化政策导向，从而引导科研人员从事相关研究。最终，学术评价将发展成为国家战略发展和社会科学进步的新驱动力。

总之，多源异构数据视角下的学术评价未来发展将会呈现多元化、深度化和智能化的趋势，需要注重创新、合理运用技术、建设多元评价生态体系、坚持中国特色的学术评价新模式，同时关注伦理道德问题、培育跨学科交叉研究以及加强国际合作与交流。这将有助于实现学术评价的长足进步，并推动科学研究和社会发展。

参考文献：

[1] 中共中央办公厅，国务院办公厅.关于深化项目评审、人才评价、机构评估改革的意见[EB/OL].(2018-07-03)[2019-04-18].https://www.gov.cn/zhengce/2018-07/03/content_5303251.htm.

[2] 国务院.关于优化科研管理提升科研绩效若干措施的通知[EB/OL].(2018-07-18)[2019-04-18].https://www.gov.cn/zhengce/content/2018-07/24/content_5308787.htm.

[3] 科技部，教育部，人力资源社会保障部，中科院，工程院.关于开展清理“唯论文、唯职称、唯学历、唯奖项”专项行动的通知[EB/OL].(2018-10-15)[2019-04-15].https://www.gov.cn/zhengce/zhengceku/2018-12/31/content_5446309.htm.

[4] 中华人民共和国中央人民政府.教育部 科技部印发《关于规范高等学校SCI论文相关指标使用 树立正确评价导向的若干意见》的通知[EB/OL].(2020-02-18)[2020-08-10].https://www.gov.cn/zhengce/zhengcekn/2020-03/03/content_5486229.htm.

[5] SPIER R.The history of the peer-review process[J].Trends in Biotechnology,2002,20(08):357-358.

[6] BESANCON L, PEIFFER-SMADJA N, SEGALAS C, et al. Open science saves lives: Lessons from the COVID-19 pandemic[J].Medical Research Methodology,2020,21(01):117.

[7] PAULSON E J,ALEXANDER J,ARMSTRONG S.Peer review re-viewed:Investigating the juxtaposition of composition students′ eye movements and peer-review processes[J].Research in the Teaching of English,2007,41(03):304-335.

[8] SILER K, LEE K, BERO L. Measuring the effectiveness of scientific gatekeeping[J]. Proceedings of the National Academy of Sciences,2015,112(02):360-365.

[9] WILSON C S.Informetrics[J].Annual Review Of Information Science and Technology,1999,34:107-247.

[10] PRITCHARD A.Statistical bibliography or bibliometrics?[J].Journal of Documentation,1969,25:348-349.

[11] GARFILED E.Citation analysis as a tool in journal evaluation[J].Science,1972,178(4060):471-479.

[12] GARFIELD E. How can impact factors be improved? [J]. British Medical Journal, 1996, 313 (7054): 411-413.

[13] HIRSCH J E. An index to quantify an individual′s scientific research output[J]. Proceedings of the

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

National Academy of Sciences of the United States of America,2005,102(46):16569-16572.

[14] VAN N R.Controversial impact factor gets a heavyweight rival[J].Nature,2016,540(7633):325-326.

[15] ATAYERO A,POPOOLA S I.Citation analytics:data exploration and comparative analyses of CiteScores of open access and subscription-based publications indexed in Scopus (2014-2016)[J].Data in Brief,2018,19:198-213.

[16] BERGER J M,BAKER C M.Bibliometrics:An Overview[J].RGUHS Journal of Pharmaceutical Sciences,2014,4(03):81-92.

[17] GUNASHEKAR S,WOODING S,GUTHRIE S.How do NIHR peer review panels use bibliometric information to support their decisions?[J].Scientometrics:An International Journal for All Quantitative Aspects of the Science of Science Policy,2017,112(03):1813-1835.

[18] QUINDOS G.Confusing the confused:thoughts on impact factor,h(irsch) index,Q value,and other cofactors that influence the researcher's happiness[J].Revista Iberoamericana De Micologia,2009,26(02):97-102.

[19] CONROY G.What's wrong with the H-index,according to its inventor[EB/OL].Nature Index,2020,24.

[20] 王浩斌.学术共同体、学术期刊与学术评价之内在逻辑解读[J].中国社会科学评价,2015(03):69-81,126-127.

[21] 苏新宁,王东波.学术评价相关问题与思考[J].信息资源管理学报,2018,8(03):4-11.

[22] 叶继元.人文社会科学评价体系探讨[J].南京大学学报(哲学·人文科学·社会科学),2010,47(01):97-110,160.

[23] 叶继元.学术"全评价"体系与中国特色哲学社会科学学术评价体系的构建与完善[J].情报资料工作,2021,42(03):15-22.

[24] 邱均平,容瑀航,汤建民.我国"五计学"的交叉融合发展:进程与趋势[J].图书馆理论与实践,2020(06):1-6,11.

[25] 邱均平,余波,杨思洛.大数据背景下一门交叉学科的兴起:论数据计量学的构建[J].中国图书馆学报,2021,47(05):48-58.

[26] 邓三鸿,陈超群,王昊,等.基于组合赋权法的高校网络影响力综合评价:以"双一流"高校为例[J].西南民族大学学报(人文社科版),2019,40(09):227-235.

[27] 张宝隆,王昊,邓三鸿,等.基于差异性视角的期刊区分性测度及分析研究[J].情报学报,2020,39(03):284-296.

[28] 闫晓慧,邓三鸿,张艺炜,等.基于BJRC模型的学术论文创新度测量研究[J].信息资源管理学报,2022,12(04):46-55.

[29] 张琳,孙蓓蓓,黄颖.交叉科学研究:内涵、测度与影响[J].科研管理,2020,41(07):279-288.

[30] 赵蓉英,刘卓著,张兆阳,等.论学术话语权及其评价[J].图书情报工作,2022,66(11):14-23.

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

[31] 王贤文,方志超,王虹茵.连续、动态和复合的单篇论文评价体系构建研究[J].科学学与科学技术管理,2015,36(08):37-48.

[32] 罗卓然,陆伟,蔡乐,等.学术文本词汇功能识别:在论文新颖性度量上的应用[J].情报学报,2022,41(07):720-732.

[33] 杨思洛,郑梦雪.完善人文社会科学成果国际影响力评价[N].中国社会科学报,2022-05-31(008).

[34] PRIEM J,TARABORELLI D,GROTH P,et al.Altmetrics:A manifesto[EB/OL].(2010-10-26)[2019-04-15].http://altmetrics.org/manifesto.

[35] GU F,WIDEN-WULFF G.Scholarly communication and possible changes in the context of social media:A Finnish case study[J].The Electronic Library,2011,29(06):762-776.

[36] ROWLANDS I,NICHOLAS D,RUSSELL B,et al.Social media use in the research workflow[J].Learned Publishing,2011,24(03):183-195.

[37] TENOPIR C,VOLENTINE R,KING D W.Social media and scholarly reading[J].Online Information Review,2013,37(02):193-216.

[38] ADIE E,ROE W.Altmetric:enriching scholarly content with article-level discussion and metrics[J].Learned Publishing,2013,26(01):11-17.

[39] ALHOORI H,FURUTA R.Do altmetrics follow the crowd or does the crowd follow altmetrics?[C]//2014 IEEE/ACM Joint Conference on Digital Libraries,London:2015.

[40] THOMA B,ROLSTON D,LIN M.Global Emergency Medicine Journal Club:Social Media Responses to the March 2014 Annals of Emergency Medicine Journal Club on Targeted Temperature Management[J].Annals of Emergency Medicine,2014,64(02):207-212.

[41] BORNMANN L.Scientific Revolution in Scientometrics:the Broadening of Impact from Citation to Societal[J].Theories of Informetrics and Scholarly Communication,2016:347-359.

[42] VAN-NOORDEN R.Online collaboration:Scientists and the social network[J].Nature,2014,512(7513):126-129.

[43] 苏娜.科学研究的社会影响力评价:研究与实践进展[J].情报学报,2020,39(10):1114-1119.

[44] GRANDE D,GOLLUST S E,PANY M,et al.Translating Research For Health Policy:Researchers' Perceptions And Use Of Social Media[J].Health Affairs,2014,33(07):1278-1285.

[45] HARLEY D,ACORD S K,EARL-NOVELL S,et al.Assessing the future landscape of scholarly communication:an exploration of faculty values and needs in seven disciplines[C]//The Center for Studies in Higher Education,Univ of California Press,Berkeley:2010.

[46] ALMIND T C,INGWERSEN P.Informetric analyses on the world wide web:methodological approaches to 'webometrics'[J].Journal of Documentation,1997,53:404-426.

[47] NEWMAN M E J.The structure and function of complex networks[J].Siam Review,2003,45(02):

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

167-256.

[48] FONTOURA C L, OLIVEIRA O N, TRAVIESO G.et al.Analyzing and modeling real-world phenomena with complex networks: A survey of applications[J].Advances in Physics, 2011, 60: 329-412.

[49] VAN-ECK N J, WALTMAN L.Visualizing Bibliometric Networks[J].Measuring scholarly impact: Methods and practice, 2014: 285-320.

[50] BERGSTROM C T, WEST J D, WISEMAN M A.The Eigenfactor metrics[J].Neurosci, 2008, 28(45): 11433-11434.

[51] YE F Y, LEYDESDORFF L.The "Academic Trace" of the Performance Matrix: A Mathematical Synthesis of the h-Index and the Integrated Impact Indicator (I3) [J].Journal of the Association for Information Science and Technology, 2014, 65(04): 742-750.

[52] HUTCHINS B I, YUAN X, ANDERSION J M, et al.Relative citation ratio (RCR): A new metric that uses citation rates to measure influence at the article level[J].PLos Biology, 2016, 14(09): e1002541.

[53] FUNK R J, OWEN S J.A Dynamic Network Measure of Technological Change[J].Management science: Journal of the Institute of Management Sciences, 2017, 63(03): 791-817.

[54] WU L, WANG D, EVANS J A.Large teams develop and small teams disrupt science and technology[J].Nature, 2019, 566(7744): 378.

[55] BU Y, WALTMAN L, HUANG Y.A multi-dimensional framework for characterizing the citation impact of scientific publications[J].Quantitative Science Studies, 2021, 2(01): 155-183.

[56] LEYDESDORFF L, BORNMANN L.Disruption indices and their calculation using web-of-science data: Indicators of historical developments or evolutionary dynamics? [J/OL].Journal of Informetrics, 2021, 15(04): Article 101219.

[57] LATHABAI H H, NANDY A, SINGH V K.X-index: Identifying core competency and thematic research strengths of institutions using an NLP and network based ranking framework[J].Scientometrics, 2021, 126(12): 9557-9583

[58] SODERLAND S. Learning information extraction rules for semi-structured and free text[J]. Machine Language, 1999, 34: 233-272.

[59] 陆伟, 李鹏程, 张国标, 等.学术文本词汇功能识别: 基于BERT向量化表示的关键词自动分类研究[J].情报学报, 2020, 39(12): 1320-1329.

[60] YANG L, ZHANG Z, CAI X, et al. Attention-Based Personalized Encoder-Decoder Model for Local Citation Recommendation[J].COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE, 2019: 1-7.

[61] 姜艺, 黄永, 夏义堃, 等.学术文本词汇功能识别: 在关键词自动抽取中的应用[J].情报学报, 2021, 40(02): 152-162.

[62] KONDO T, NANBA H.Technical trend analysis by analyzing research papers' titles[C]//Proceedings of

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

the Language and Technology Conference on Human Language Technology, Challenges for Computer Science and Linguistics.Heidelberg:Springer,2009:512-521.

[63] NANBA H, KONDO T, TAKEZAWA T. Automatic creation of a technical trend map from research papers and patents[C]//Proceedings of the 3rd International Workshop on Patent Information Retrieval. New York:ACM Press,2010.

[64] CHEN L, XU S, ZHU L, et al.A deep learning based method for extracting semantic information from patent documents[J].Scientometrics,2020,125(01):289-312.

[65] DOUCET A, COUSTATY M.Enhancing table of contents extraction by system aggregation[C]//2017 14th IAPR international conference on document analysis and recognition(ICDAR).IEEE,Kyoto:2017.

[66] SAFDER I, HASSAN S U, ALJOHANI N R. AI cognition in searching for relevant knowledge from scholarly big data,using a multi-layer perceptron and recurrent convolutional neural network model[C]// Companion of the web conference 2018.Lyon:2018.

[67] LI C, XING J, SUN A, et al. Effective document labeling with very few seed words: A topic model approach[C]//Proceedings of the 25th ACM international on conference on information and knowledge management, Indianapolis:2016.

[68] SIEGEL N, LOURIE N. Extracting scientific figures with distantly supervised neural networks[C]// Proceedings of the 18th ACM/IEEE on joint conference on digital libraries,New York:2018.

[69] AL-ZAIDY R A, GILES C L. Automatic knowledge base construction from scholarly documents[C]// Proceedings of the 2017 ACM symposium on document engineering,Beijing:2017.

[70] GARFIELD E. Cam citation indexing be automated[C]//Statistical association methods for mechanized documentation,Symposium proceedings.1964:189-192.

[71] BROOKS T A. Evidence of complex citer motivations[J].Journal of the American society for Information Science,1986,37(01):34-36.

[72] TEUFEL S, CARLETTA J, MOENS M. An annotation scheme for discourse-level argumentation in research articles[C]//Proceedings of the Ninth Conference on European Chapter of the Association for Computational Linguistics,Bergen:1999.

[73] HUANG H, ZHU D, WANG X. Evaluating scientific impact of publications: combining citation polarity and purpose[EB/OL].Scientometrics,2021.

[74] DING Y, LIU X, GUO C, et al.The distribution of references across texts: some implications for citation analysis[J].Journal of Informetrics,2013,7(03):583-592.

[75] WAN X, LIU F. Are all literature citations equally important? Automatic citation strength estimation and its applications[J].Journal of the Association for Information Science and Technology,2014,65(09):1929-1938.

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

[76] EBRAHIMY S, OSAREH F.Design, validation, and reliability determination a citing conformity instrument at three levels: normative, informational, and identification[J].Scientometrics, 2014, 99(02): 581-597.

[77] 陆伟, 孟睿, 刘兴帮.面向引用关系的引文内容标注框架研究[J].中国图书馆学报, 2014(06): 93-104.

[78] AN J, KIM N, KAN M Y, et al.Exploring characteristics of highly cited authors according to citation location and content[J].Journal of the Association for Information Science and Technology, 2017, 68(08): 1975-1988.

[79] ZHAO S X, QIAO L, QIAN X H D I.Man: An analytical framework for solving complex Issues[J]. Science, 2021, 374(Oct.1 App.TN.6563): 39-41.

Academic Evaluation from the Perspective of Multi-source Heterogeneous Data:Connotation, Status and Prospect

DENG Sanhong[1], YANG Jie[1], WANG Hao[1], SHEN Jianwei[2]

1.School of information management, Nanjing University, Nanjing 210023

2.Jiangsu Institute of Quality and Standardization, Nanjing 210023

Abstract: [Purpose/significance]Academic evaluation and scientometric methods of synthesizing multisource heterogeneous data have gradually become hot spots in the field. At present, there is no systematic theoretical discussion on academic evaluation from the perspective of multisource heterogeneous data.[Method/process]In view of insufficient utilization of multi-source heterogeneous data, inappropriate research methods, and unclear research objectives, this paper combs the development of academic evaluation from the perspective of multi-source heterogeneous data, and clarifies the connotation and extension of the field, and analyzes the frontiers and hot spots in the field.[Result/conclusion]The results state clearly that from the perspective of multi-source heterogeneous data, academic evaluation manifests the trend of increasing data dimension, data granularity refinement, and evaluation index complexity. The definition of academic evaluation connotation, current situation analysis and future prospects from the perspective of multi-source heterogeneous data are of great significance to clarify domain recognition, identify related concepts, clarify research paths and methodological basis, and build an academic evaluation system with Chinese characteristics.

Keywords: academic evaluation; scientometrics; multi-source and heterogeneous data

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net