

• 信息计量与科学评价 •

主题新颖性程度—融合程度—影响力三维 视角下的学术论文创新性评价研究

魏绪秋 常霞 姜召昊 申力旭

(山东理工大学信息管理学院, 山东 淄博 255000)

摘要: [目的/意义] 如何科学、合理地评价学术论文创新水平和科学价值成为研究的关注点。本研究构建了融合主题新颖性程度、融合程度、影响力三维的学术论文创新性评价方式, 将为学术论文创新性评价相关研究提供借鉴。[方法/过程] 本研究基于学术论文研究主题内容及价值, 从学术论文研究主题新颖性程度、主题融合程度和主题影响力 3 个维度, 采用 TOPSIS 法, 构建主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价模型, 并以 2021 年度人文社会科学中区块链为研究主题的学术论文为研究对象, 进行实证研究。[结果/结论] 研究发现, 当前我国人文社会科学区块链研究维持着较高的增长态势, 且其关注度在稳定的同时略有变化; 所构建的主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价模型可作为研究活动和绩效的衡量标准, 丰富了学术论文创新性评价方式。

关键词: 学术论文; 创新性评价; 主题新颖性程度; 主题融合程度; 主题影响力; TOPSIS 法; 区块链

DOI: 10.3969/j.issn.1008-0821.2024.03.012

(中图分类号) G250.252 (文献标识码) A (文章编号) 1008-0821 (2024) 03-0131-09

Study on the Evaluation of Academic Papers' Innovativeness from the Three-Dimensional Perspective of the Degree of Novelty-Integration-Influence of the Topic

Wei Xuqiu Chang Xia Jiang Zhao hao Shen Lixu

(School of Information Management, Shandong University of Technology, Zibo 255000, China)

Abstract [Purpose/Significance] How to scientifically and rationally evaluate the innovation level and scientific value of academic papers has become a research concern. This study constructs a three-dimensional evaluation method of academic papers' innovativeness integrating the degree of novelty, integration and influence of the topic, which would provide reference for the research related to the evaluation of academic papers' innovativeness. [Method/Process] Based on the content and value of academic paper research themes, this study used the TOPSIS method to construct a three-dimensional evaluation model of academic papers' innovativeness integrating the degree of novelty, integration and influence of the topic. This research conducted empirical research on academic papers with the theme of blockchain in the humanities and social sciences in 2021 in China. [Result/Conclusion] The study finds that: the current blockchain research in the humanities and social sciences maintains a high growth trend, and its attention is stable while slightly changing in China; the construction of the three-dimensional evaluation model of academic papers' innovativeness integrating the degree of novelty, integra-

收稿日期: 2023-07-04

基金项目: 国家社会科学基金青年项目“基于文本挖掘的学术论文创新性识别研究”(项目编号: 21CTQ022)。

作者简介: 魏绪秋 (1990-), 男, 副教授, 博士, 硕士生导师, 研究方向: 信息计量、科学评价与数据挖掘。常霞 (1999-), 女, 硕士研究生, 研究方向: 信息计量、科学评价与数据挖掘研究。姜召昊 (1999-), 男, 硕士研究生, 研究方向: 信息计量、科学评价与数据挖掘研究。申力旭 (1998-), 女, 硕士研究生, 研究方向: 信息计量、科学评价与数据挖掘研究。

tion and influence of the topic can be used as a measure of research activities and performance, enriches the way of evaluating the innovativeness of academic papers, can be used as a measure of research activities and performance, enriches the innovative evaluation methods of academic papers.

Key words: academic paper; innovativeness evaluation; degree of topic novelty; degree of topic integration; degree of topic value; TOPSIS method; block chain

中国共产党第十八届中央委员会第五次全体会议明确了“创新、协调、绿色、开放、共享”的新发展理念。而在其中,“创新”一词排在第一位。可见,创新引领发展,科技赢得未来。学术论文作为科研活动的重要研究成果,通常反映了研究人员在科研活动中所取得的新进展、新知识、新价值与新应用。2020年2月18日,教育部、科技部研究制定的《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》中明确指出:坚决摒弃“以刊评文”,评价重点是论文的创新水平和科学价值。因此,如何科学、合理地评价学术论文创新水平和科学价值引发科研人员和科研单位的热议,并成为科研人员和科研单位关注、研究的重点。

当前,科技创新成为国际战略博弈的主要战场,围绕科技制高点的竞争空前激烈,各国(地区)对创新性知识的需求急剧增加。研究人员需要从海量的学术论文中发现创新性较高的论文,挖掘出创新性知识,以满足对创新性知识的需求。然而,学术论文创新性的测度逐渐趋于复杂化和技术化,如贺婉莹^[1]认为运用机器学习的方法,通过对多维数据学习和理解,以一种自适应的方式评价学术论文的创新力有望解决定性评价和定量评价中的缺陷;沈超^[2]认为可以借助文本挖掘技术,从审稿人的评论文本中,挖掘有关创新力的内容,运用机器学习模型和“投票原则”,构造一个基于评论文本创新力情感分析的学术论文创新力评价指标,综合评价学术论文的创新力。

学术论文创新性评价具有复杂性,受多重因素的影响,例如,评价角度的多样性、评价层次的差异性、评价方法的灵活性以及学术论文自身的多样性与复杂性等。汪雪锋等^[3]认为,基于知识元组合视角的学术论文创新质量评价,将创新视为对已有知识元素的新的或者非典型的结合,知识元素包括关键词、主题词等。黄晨等^[4]认为,以关键词形式出现的动态新概念可能携带创新基因。钱玲飞等^[5]

认为,学术论文的创新具有继承性和发展性。闫晓慧等^[6]认为,学术论文创新度越高,受到的关注度也越高,得到的引用次数也就越多。因此,本研究一方面基于主题内容方面,主要考虑学术论文研究主题的新颖性、主题继承与发展的分裂与融合程度;另一方面兼顾主题学术影响力方面,主要考虑学术论文主题学术影响程度,有机融合学术论文主题新颖性程度、主题融合程度、主题影响力,构建多维视角下的学术论文创新性评价模型,来评价学术论文创新性。

1 相关研究

学术论文的创新性是指研究工作在学术领域中的独创性和独特性,钟细军^[7]认为,科技学术论文的创新具有3个重要属性:学科领域性和专业方向性;真实性和可重复实现性;比对性或继承性。陈建青^[8]认为,可以从文章内容、研究方法、研究视角等多个方面来阐述创新的内涵与外延。魏绪秋等^[9]认为,学术论文的创新性包括观点创新、理论创新、方法与技术创新、结果与结论创新和应用创新5类。当前,学者们尝试了多种方式来评价学术论文创新性:一是运用同行评议方法评价学术论文创新性;二是基于引用、社交媒体数据评价学术论文创新性;三是从学术论文内容本身出发评价学术论文创新性。

1.1 同行评议方法在学术论文创新性中的应用

同行评议方法是当前学术奖励制度和学术文献出版过程中通用的主观评价方法,基本原理是:在某特定领域中,专家对该领域或相近领域其他科研人员的科技著作的专业表现、创新性以及质量进行评价的方法^[10]。如朱大明^[11]认为,初审编辑可以从学术论文标题、引言、正文、结论、参考文献、来源等部分鉴审学术论文创新性;曾群等^[12]认为,学术论文审稿人对科技论文的学术评价内容就是对学术创新性进行的评价。

1.2 引用、社交媒体数据在学术论文创新性中的应用

Garfield E^[13]认为,引文索引为阐明新颖性等

概念提供了新基础。因此, 众多研究人员用引用数据量化学术论文创新性, 如 Uzzi B 等^[14]认为, 参考文献可以追溯论文研究的知识组合过程, 期刊被视为知识块的主体, 通过引用的期刊对的常规性和新颖性评估论文的创新性; 姚飞等^[15]基于引文分析方法研究了《重庆大学学报(自然科学版)》的学术质量, 发现普赖斯指数高于平均值, 论文新颖度高; Chen C 等^[16]提出了一种基于引文网络的方法, 即引文权威扩散(Citation Authority Diffusion, CAD), 快速挖掘某个主题的有限关键论文, 衡量文献新颖性; Matsumoto K 等^[17]运用新颖性指标来量化自然科学各领域重点论文和先前存在的同一领域论文间的引用相似度, 识别属于同一领域重点论文的新方法。随着信息技术的发展, 社交媒体研发、普及及应用, 在线实时分享、跟踪与评价科学数据成为可能。社交媒体数据具有实时性特点, 被研究人员用来快速量化学术论文价值, 在一定程度上间接反映了学术论文创新性, 如 Hammarfelt B^[18]认为, Altmetrics 可以发展成为评估人文学科研究的宝贵工具; 赵蓉英等^[19]认为, 基于 Altmetrics 的学术论文影响力评价可以作为对传统影响力评价在社会影响力层面的补充, 具有预测高被引论文的作用; 郝若扬^[20]认为, 高 Altmetrics 指标分数的论文与其被引频次存在一定的正相关, 高 Altmetrics 指标论文同时具有较高的学术影响力; 王睿等^[21]认为, 作为一种早期指标, 高 Altmetrics 指标在一定程度上能够被视作文章在未来获得高被引的风向标。

1.3 从学术论文内容本身评价学术论文创新性

随着研究的深入, 研究人员、科研人员开始关注学术论文本身, 基于学术论文内容展开学术论文创新性识别的研究, 其主要包括基于主题词层面的学术论文创新性评价和基于句子层面的学术论文创新性评价研究。①基于主题词层面的学术论文创新性评价研究, Mishra S 等^[22]引入时间新颖性(主题概念在语料库中首次出现的时间)和数量新颖性(自第一个主题概念出现在语料库以来的文献数量)来衡量学术论文创新性; 杨京等^[23]构建了一种基于研究主题对比的单篇学术论文创新力评价方法; 曹树金等^[24]结合 LDA、SVM 方法, 识别潜在主题, 判断论文创新性; ②基于句子层面的学术论文创新

性评价研究, 侯剑华等^[25]提出了一种基于 SAO-ADV 模型及语义相似度的论文创新性测度方法; 周海晨等^[26]提出一个深度学习与规则结合的学术创新贡献识别方法; 张吉玉等^[27]提出一种考虑时序的科技文献问题—方法矩阵用于评估单篇文献新颖性。

综上可知, 虽然同行评议方法是通用的一种主观评价方法, 但是在实际操作过程中引起了争议, 如同行专家可能会存在偏见或仅支持预期结果^[28]。引用、社交媒体数据在评价学术论文创新性时虽然易于操作, 但也存在着一定缺陷, 如 Macroberts M H 等^[29]概述了引文分析的未正式引用、偏见引用等七大问题; Liu J 等^[30]论述了 Altmetrics 面临着如何解读数据、该使用哪些数据源等五大挑战。从学术论文内容本身评价学术论文创新性成为近年来学者们关注的重点, 并成为学术论文创新性识别的重要途径之一。因此, 融合多种方式的学术论文创新性评价将成为研究的重点。在此背景下, 本研究将从学术论文研究主题新颖性程度、主题融合程度和主题影响力 3 个维度综合评价学术论文创新性, 为学术论文创新性评价相关研究提供借鉴。

2 主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价设计

2.1 学术论文创新性评价流程设计

一篇文献的关键词或主题词是其核心内容的浓缩和提炼, 代表该文献的研究主题^[31]。杨建林等^[32]认为, 一对共现关键词在量化某文档的主题新颖度时的价值会随着在该文档之前发表的、包含该对共现关键词的文档数量的增加而降低, Luhn H P^[33]认为, 能够反映文档的主题新颖性的关键词是文档中具有较高词频的词汇, 李长玲等^[34]认为, 科学文献的价值与质量主要通过被引数据相关的指标进行测度。基于此, 本研究设计了主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价流程, 如图 1 所示。

1) 数据清洗。潘玮等^[35]认为, 数据清洗的对象有: 缺失、错误、不一致的关键词以及重复或相似重复的记录等。对关键词的清洗尤为重要, 其会影响到研究结果的准确性。

2) 数据统计。根据清洗后的关键词, 自编程

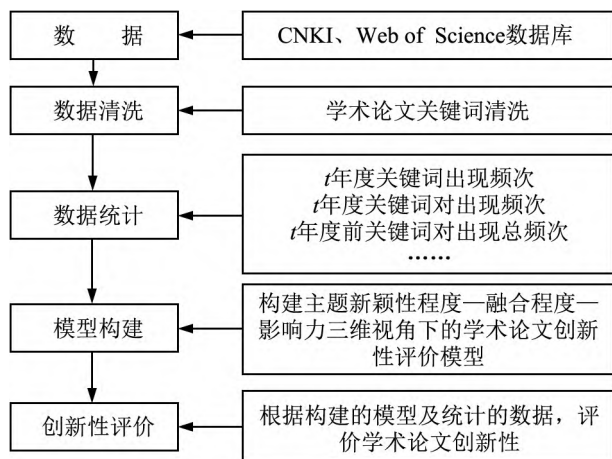


图 1 主题新颖性程度—融合程度—影响力三维视角下的
学术论文创新性评价流程

Fig.1 Process of Evaluating the Innovativeness of Academic
Papers Under Three-dimensional Perspective of the Degree of
Novelty-Integration-Influence of the Topic

序，依据学术论文和关键词间的隶属关系，配对关键词对，并统计 t 年度的关键词出现频次、 t 年度的关键词对出现频次、 t 年度前关键词对出现总次数、 t 年度关键词对被引用次数等数据。

3) 模型构建。从主题新颖性程度、主题融合程度、主题影响力三维，构建学术论文创新性评价模型。

4) 创新性评价。根据统计的数据、构建的模型，评价学术论文创新性。

2.2 学术论文创新性评价模型构建

2.2.1 学术论文主题新颖性维度(Novelty)

在创新性研究中，新颖性是将现有知识以一种全新的、前所未有的方式结合起来的成果，可以是一种渐进式的技术进步，也可以是一种跃进性的技术革新^[36]。本研究主要通过关键词对的出现次数及关键词对累积出现次数反映学术论文主题新颖性程度，如式 (1) 所示：

$$Novelty = \frac{\sum_{n=1}^N \left(\frac{T < k_i, k_j >_t}{1 + \sum_{i=0}^{t-1} T < k_i, k_j >_i} \right)}{N} \quad (1)$$

式 (1) 中， $< k_i, k_j >$ 表示关键词 k_i 和 k_j 对， $T < k_i, k_j >_t$ 表示在 t 年度的关键词 k_i 和 k_j 对出现的次数， $\sum_{i=0}^{t-1} T < k_i, k_j >_i$ 表示关键词 k_i 和 k_j 对在 t 年度前出现的累积次数， N 表示某篇学术论文关键词对的对数，“+1”避免关键词 k_i 和 k_j 对在 t 年度前累积出现次数为 0 的情形。

对数，“+1”避免关键词 k_i 和 k_j 对在 t 年度前累积出现次数为 0 的情形。

2.2.2 学术论文主题融合程度维度(Fusion)

韩芳等^[37]认为，如何定量识别技术交叉现象，把握技术融合发展方向是识别突破性创新的一个关键问题。因此，本研究主要通过关键词对所反映主题在 t 年度的融合程度，揭示主题的融合与分裂，反映主题的继承与发展程度，即其值越接近 1 说明关键词组成的关键词对越集中，该关键词对较为稳定，主题融合程度较好；越接近 0 说明该关键词组成的关键词对较为分散，该关键词对稳定性弱，主题融合程度较差，以此来反映学术论文主题融合、分裂程度，其测度结果值在区间 $(0, 1]$ 内，如式 (2) 所示：

$$Fusion = \frac{\sum_{n=1}^N \left(\frac{T < k_i, k_j >_t}{\sqrt{T < k_i >_t \times T < k_j >_t}} \right)}{N} \quad (2)$$

式 (2) 中， $T < k_i >_t$ 表示在 t 年度的关键词 k_i 出现的次数， $T < k_j >_t$ 表示在 t 年度的关键词 k_j 出现的次数， N 表示某篇学术论文关键词对的对数。

2.2.3 学术论文主题影响力维度(Influence)

学术论文主题影响力能够反映出该研究主题对研究人员的关注度、吸引力。本研究主要利用学术论文引用次数量化关键词对的引用程度，以反映学术论文主题影响力，其测度结果值在区间 $[0, 1]$ 内，如式 (3) 所示：

$$Influence = \frac{\sum_{n=1}^N \left(\frac{C < k_i, k_j >_t}{1 + \sum C < k_i, k_j >_i} \right)}{N} \quad (3)$$

式 (3) 中， $C < k_i, k_j >_t$ 表示某篇学术论文在 t 年度的关键词 k_i 和 k_j 对被引用次数， $\sum C < k_i, k_j >_i$ 表示关键词 k_i 和 k_j 对在 t 年度被引用次数， N 表示某篇学术论文关键词对的对数，“+1”避免关键词 k_i 和 k_j 对在 t 年度被引用次数为 0 的情形。

2.2.4 学术论文创新性评价模型(Model)

本研究采用 TOPSIS 法构建主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价模型。TOPSIS 法是 Hwang C 等^[38]于 1981 年首次提出，TOPSIS 法是根据有限个评价对象与理想化目标的接近程度进行排序的方法，是在现有的对

象中进行相对优劣的评价。TOPSIS 法的具体步骤如下:

假设有 n 个要评价对象, m 个评价指标, 构成正向化矩阵 X :

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

数据标准化处理, 为了消除不同的数据指标量纲的影响, 需要将矩阵 X 进行标准化处理, 记标准化后的矩阵为 Z , 其中 $z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}}$, 标准化矩

阵 Z :

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ z_{n1} & z_{n2} & \cdots & z_{nm} \end{bmatrix}$$

优解与最劣解计算, 取出每个指标, 即每一列中最大的数, 构成理想最优解向量, 即:

$$z^+ = [z_1^+, \cdots, z_m^+] = [\max\{z_{11}, \cdots, z_{n1}\}, \cdots, \max\{z_{1m}, \cdots, z_{nm}\}]$$

同理, 取每一列中最小的数计算理想最劣解向量, 即:

$$z^- = [z_1^-, \cdots, z_m^-] = [\min\{z_{11}, \cdots, z_{n1}\}, \cdots, \min\{z_{1m}, \cdots, z_{nm}\}]$$

TOPSIS 评分计算, 对于第 z_i 个方案, 计算它与最优解的距离:

$$d_i^+ = \sqrt{\sum_{j=1}^m (z_j^+ - z_{ij})^2} \quad (4)$$

与最劣解的距离:

$$d_i^- = \sqrt{\sum_{j=1}^m (z_j^- - z_{ij})^2} \quad (5)$$

根据最优解与最劣解计算得分:

$$S_i = \frac{d_i^-}{d_i^+ + d_i^-} \quad (6)$$

学术论文创新性评价模型(Model):

$$Model = S_{Novelty} \times Novelty + S_{Fusion} \times Fusion + S_{Influence} \times Value \quad (7)$$

3 实证分析——以人文社会科学的区块链为例

3.1 数据来源与数据处理

3.1.1 数据来源

区块链起源于比特币(Bitcoin) ^[39], 区块链技术的核心优势是去中心化, 能够通过运用数据加密、时间戳、分布式共识和经济激励等手段, 在节点无需互相信任的分布式系统中实现基于去中心化信用的点对点交易、协调与协作, 从而为解决中心化机构普遍存在的高成本、低效率和数据存储不安全等问题提供了解决方案^[40], 并成为政府部门、金融机构、科技企业和资本市场关注的焦点。因此, 本研究关注人文社会科学中区块链的相关研究问题, 并以人文社会科学中区块链相关研究文献为例, 从主题新颖性程度—融合程度—影响力三维视角下评价学术论文创新性, 以促进区块链在人文社会科学的发展与实际应用。

CSSCI 是南京大学自主研发的人文社会科学引文索引数据库, 遵循文献计量学规律, 采取定量与定性评价相结合的方法, 从全国2 700余种中文人文社会科学学术性期刊中精选出学术性强、编辑规范的期刊作为来源期刊, 在人文社会科学领域的知识创新、前沿研判、综合评价, 对提升中文学术影响力有着重要的作用。本研究数据来源为中国知识学术期刊数据库, 检索策略如下: 主题为“区块链”, 来源类型限定为 CSSCI, 出版年度截至 2021 年, 检索时间为 2022 年 9 月 6 日。经检索, 共检索到1 409条文献题录。本研究以 2021 年度学术论文为研究对象, 测度 2021 年度学术论文创新性, 故在检索文献题录时采集了 2021 年度学术论文的引用数据, 以备用。

3.1.2 数据处理

1) 在数据清洗方面, 删除会议通知、编辑约稿说明等无效文献题录, 总计有1 395条文献题录, 具体文献题录出版年度分布如图 2 所示。抽取1 395篇学术论文题录的关键词, 并对关键词进行清洗(如“LDA”和“LDA 主题模型”统一为“LDA”, “版权”和“著作权”统一为“版权”等)。

由图 2 可知, 我国人文社会科学中区块链研究最早来源于 2014 年的《比特币: 运行原理、典型

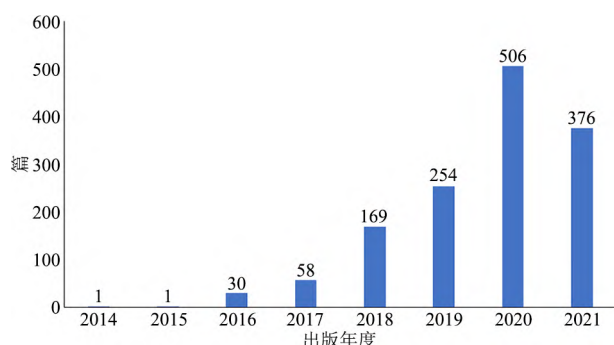


图 2 人文社会科学中区块链出版年度分布
Fig. 2 The Annual Distribution of Blockchain Publishing in the Humanities and Social Sciences

特征与前景展望》。该篇学术论文阐述了比特币的运行原理，剖析了比特币的典型特征，并展望了比

特币的可能前景^[41]，为后续比特币的研究提供了借鉴与参考。随后，其出版年度篇数大致呈增长趋势：2018 年首次超过 100 篇，为 169 篇；2019 年有持续增长，突破 250 篇；2020 年达到增长峰值，突破 500 篇。虽然 2021 年增长速度有所放缓，但仍维持 376 篇的年度出版篇数。

2) 在数据统计方面，自编程序，统计 2021 年学术论文关键词出现次数、关键词对出现次数、关键词对引用次数，以及 2021 年前(即 2014—2020 年度)的学术论文关键词对出现次数等数据。2021 年学术论文关键词对和 2021 年前的学术论文关键词对累积出现次数如表 1 所示。

表 1 关键词对出现次数(a~b)

Tab. 1 The Number of Occurrences of Keyword Pairs (a~b)

2021 年学术论文关键词对(a)			2021 年前的学术论文关键词对(b)		
关键词 1	关键词 2	出现次数	关键词 1	关键词 2	累积出现次数
区块链	智能合约	22	区块链	智能合约	60
区块链	数字经济	11	区块链	人工智能	39
区块链	人工智能	8	大数据	区块链	38
电子证据	区块链	7	区块链	去中心化	36
区块链	去中心化	7	比特币	区块链	28
供应链	区块链	6	区块链	数字货币	23
版权	区块链	6	区块链	图书馆	20
区块链	数字货币	6	金融科技	区块链	19
.....

由表 1 可知，关键词对“区块链—智能合约”无论在 2021 年前的累积出现次数还是在 2021 年的出现次数均位居前列，是当前研究人员关注与研究的重点。2021 年，研究人员对“区块链—数字经济”“电子证据—区块链”等关键词对的关注有所提高，而对“区块链—人工智能”“区块链—去中

心化”等关键词对的关注略有下降。可见，研究人员的关注度随研究年度略有变化。

3.2 学术论文创新性评价

根据式(1)~(7)及统计的数据，测度学术论文创新性，测度结果如表 2 所示。

表 2 学术论文创新性测度结果

Tab. 2 Results of Academic Paper Innovativeness Measures

题 名	第一作者	主题新颖程度	主题融合程度	主题影响力	创新性
智能技术生态驱动未来教育发展	杨现民	0.0402	0.0552	0.0491	0.1445
面向健康医疗大数据安全保护的医疗区块链模型构建	李洪晨	0.0402	0.0552	0.0462	0.1416
基于区块链的纠纷解决机制研究	杨锦帆	0.0402	0.0552	0.0455	0.1409

表 2 (续)

题 名	第一作者	主题新颖程度	主题融合程度	主题影响力	创新性
基于区块链的开放政府数据个人隐私保护模型及实现机理研究	陈兰杰	0.0402	0.0552	0.0455	0.1409
基于多链式区块链的互联网金融犯罪情报共享模型研究	张成虎	0.0402	0.0552	0.0433	0.1387
基于区块链技术应用的反垄断法律规制研究	蔡莉妍	0.0402	0.0552	0.0433	0.1387
身体、空间与城市：5G 时代智能城市的媒介化重构	白 龙	0.0402	0.0552	0.0433	0.1387
发达国家数字政府建设的探索与经验借鉴	胡税根	0.0402	0.0471	0.0501	0.1374
工具结构和产业发展：政策工具视角下中国区块链政策的文本分析	董石桃	0.0402	0.0552	0.0416	0.1370
教育新基建时代数字教育资源的建设方向与发展路径	柯清超	0.0402	0.0471	0.0482	0.1356
.....

由表 2 可知，杨现民等在《现代远程教育研究》上发表的《智能技术生态驱动未来教育发展》一文创新性最高(0.1445)，其认为运用人工智能、5G 通信、区块链等新一代信息技术促进教育变革已成为全球共识，而技术间加速融合、协同创新所形成的智能技术生态是驱动未来教育发展的关键，其含有“智能技术生态”“未来教育”“教育变革”“教育样态”4 个关键词，引用次数为 17 次，学术论文主题新颖性程度为 0.0402，学术论文主题融合程度为 0.0552，学术论文主题影响力为 0.0491；李洪晨等在《图书情报工作》上发表的《面向健康医疗大数据安全保护的医疗区块链模型构建》一文创新性次之(0.1416)，其构建了基于区块链技术的健康医疗大数据信息安全保护模型和系统架构，推动健康医疗数据化的发展，其含有“健康医疗大数据”“区块链去中心化”“信息安全保护”“PBFT 算法”4 个关键词，引用次数为 8 次，学术论文主题新颖性程度为 0.0402，学术论文主题融合程度为 0.0552，学术论文主题影响力为 0.0462。

本研究按照“高一较高—一般—较低—低”5 个等级由高到低对学术论文主题新颖程度、主题融合程度、主题影响力、创新性进行均等划分，并绘制了堆积柱形图，如图 3 所示。

由图 3 可知，我国人文社会科学中区块链研究的大多数学术论文主题新颖性程度位于“一般”等级(约占 71.54%)；学术论文的主题融合程度大多位于“低”等级(约占 43.62%)；学术论文的主题影响力大多位于“较高”等级(约占 37.77%)或

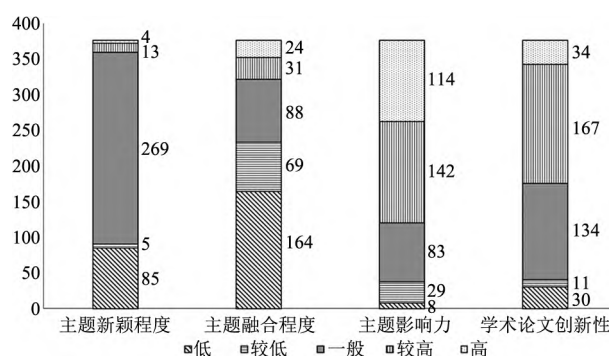


图 3 “高一较高—一般—较低—低”等级堆积柱形图
Fig. 3 Histogram of “High-High-Average-Low-Low” Rank Stacking

“高”等级(约占 30.32%)；而对于学术论文创新性而言，约 35.64% 学术论文位于“一般”等级，约 44.41% 的学术论文位于“较高”等级，仅有 9.04% 的学术论文被列为“高”等级。可见，我国人文社会科学中区块链研究的学术论文主题新颖程度维度、主题融合程度维度和主题影响力维度的测度结果等级相互交叉，形成了 46 种组合(如主题新颖程度“一般”—主题融合程度“低”—主题影响力“较高”组合的学术论文最多为 48 篇，主题新颖程度“一般”—主题融合程度“低”—主题影响力“高”组合的学术论文次之，为 36 篇)。不同维度的测度结果相互作用，提高了学术论文的创新性，并与学术论文创新性测度结果等级相互交叉，形成了 60 组组合(如主题新颖程度“一般”—主题融合程度“低”—主题影响力“高”—学术论文创新性“较高”组合的学术论文最多为 35 篇，主题新颖程度“一般”—主题融合程度“低”—主题影响力“较高”—学术论文创新性“较高”组合和主题新

颖程度“一般”—主题融合程度“低”—主题影响力“一般”—学术论文创新性“一般”组合的学术论文次之为 27 篇)。

Kendall 秩相关系数是计算多个等级变量相关程度的一种相关量,采用的是非参数检验方法度量定序型变量间的线性相关关系。它利用变量的秩计算一致对数目(U) 和非一致对数目(V)。本研究运用 Kendall 秩相关系数分析学术论文创新性主题新颖性程度、主题融合程度、主题影响力、学术论文被引频次间的定序结果的一致性。Kendall 秩相关系数结果如表 3 所示。由表 3 可知,学术论文创新性与主题新颖性程度的定序结果的 Kendall 的 tau_b 相关系数为 0.330,呈现弱相关性;学术论文创新性与主题融合程度、主题影响力的定序结果的 Ken-

dall 的 tau_b 相关系数在 0.5~0.6 之间,呈现中等程度相关性。这在一定程度上表明:学术论文主题新颖性程度、主题融合程度和影响力程度对学术论文创新性的重要性均发挥了一定作用,是 3 个维度相互作用、融合的结果,任意一个维度都不能代替学术论文创新性测度结果。学术论文创新性与学术论文被引频次的定序结果的 Kendall 的 tau_b 相关系数为 0.401,呈现中等程度相关性。学术论文被引频次被用来作为研究活动和绩效的衡量标准^[42],这在一定程度上表明,本研究构建的主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价模型与学术论文被引频次的定序结果有所差异,两者不能互相替代,可作为研究活动和绩效的衡量标准。

表 3 学术论文创新性与主题新颖性程度、主题融合程度、主题影响力、学术论文被引频次的 Kendall 秩相关系数
Tab.3 Kendall Rank Correlation Coefficients of Academic Paper Innovativeness with Degree of Subject Novelty , Degree of Subject Integration , Subject Impact , and Citation Frequency of Academic Papers

			主题新颖程度	主题融合程度	主题影响力	被引频次
Kendall 的 tau_b	学术论文创新性	相关系数	0.330**	0.524**	0.535**	0.401**
		Sig.(双侧)	0.000	0.000	0.000	0.000
		N	376	376	376	376

注: ** . 在置信度(双侧) 为 0.01 时,相关性是显著的。

4 结 语

新时代,学术论文评价的重点是评价其创新水平和科学价值。本研究融合学术论文主题新颖程度、主题融合程度和主题影响力 3 个维度,构建主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价模型,综合评价学术论文创新性,研究发现:

1) 当前,我国人文社会科学区块链研究维持着较高的增长态势。自 2014 年我国人文社会科学关注区块链研究后,近两年维持 300 篇以上的学术论文增长量。这与国家的宏观指导密不可分,如 2016 年 12 月 15 日,国务院关于印发《“十三五”国家信息化规划》的通知中明确指出,加强量子通信……区块链等新技术基础研发和前沿布局^[43]; 2020 年 1 月 17 日,发布的《国务院办公厅关于支持国家级新区深化改革创新加快推动高质量发展的指导意见》中指出,加快推动区块链技术和产业

创新发展,探索“区块链+”模式,促进区块链和实体经济深度融合^[44]等。上述通知、指导意见等从宏观层面体现我国政府鼓励探索研究区块链,服务实体经济,从而加快了包括人文社会科学研究人员对区块链相关领域的探索、创新与应用研究,研究成果维持较高增长态势。

2) 我国人文社会科学区块链研究的关注度随研究年度略有变化。经统计,2021 年出现频次在 4 次及以上的学术论文关键词对在 2021 年前均已出现过; 2021 年出现频次在 3 次的学术论文关键词对共有 20 对,但仅有 6 对并未在 2021 年前的学术论文关键词对中出现。可见,我国人文社会科学区块链研究关注度在稳定的同时略有变化。这正与罗棋等^[45]对国际区块链研究结果相一致,即国际区块链领域主题热度分布近年来逐渐趋于均衡。

3) 主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价模型可作为研究活动

和绩效的衡量标准。该模型将主题的内容、主题的价值有机融合,从学术论文主题新颖程度、主题融合程度和主题影响力 3 个维度测度学术论文创新性,丰富了学术论文创新性评价方式。该模型不仅考虑了学术论文关键词对的出现次数及关键词对累积出现次数对主题新颖性的影响程度,考虑了关键词对的集中、分散程度,还考虑了量化的关键词对反映的学术论文主题学术影响力,从而形成了主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价模型。

然而,主题新颖性程度—融合程度—影响力三维视角下的学术论文创新性评价研究,仅以人文社会科学区块链主题领域的学术论文为例,研究样本较少,研究维度局限于主题层面,在后续研究中要进一步增加研究样本的规模并扩展到不同的研究领域,深入到学术论文文本的句子层面,更全面深入地基于学术论文所取得的新进展、新知识、新价值与新应用等评价学术论文创新性。

参 考 文 献

- [1] 贺婉莹. 基于机器学习的论文学术创新力评价研究 [D]. 南京: 南京大学, 2019.
- [2] 沈超. 基于文本挖掘的学术论文创新力评价研究 [D]. 杭州: 杭州电子科技大学, 2022.
- [3] 汪雪锋, 于慧妍, 郑思佳, 等. 学术论文创新质量评价研究——以多能干细胞技术为例 [J/OL]. 数据分析与知识发现, 1-21 [2023-12-07] <http://kns.cnki.net/kcms/detail/10.1478.G2.20230608.1604.002.html>.
- [4] 黄晨, 赵星, 卞杨奕, 等. 测量学术贡献的关键词分析法探析 [J]. 中国图书馆学报, 2019, 45 (6): 84-99.
- [5] 钱玲飞, 贺婉莹, 杨建林. 论文学术创新力特征指标体系研究 [J]. 情报科学, 2021, 39 (1): 56-64.
- [6] 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 (4): 46-55.
- [7] 钟细军. 论科技学术论文创新性的初审评价 [J]. 编辑学报, 2010, 22 (2): 108-110.
- [8] 陈建青. 对我国学术论文创新性评审的几点思考 [J]. 青年记者, 2013, (18): 33-35.
- [9] 魏绪秋, 申力旭. 学术论文创新性研究述评 [J]. 图书情报知识, 2022, 39 (4): 68-79.
- [10] Lee C J, Sugimoto C R, Zhang G, et al. Bias in Peer Review [J]. Journal of the American Society for Information Science and Technology, 2013, 61 (1): 2-17.
- [11] 朱大明. 初审编辑鉴审科技论文创新性的几个途径 [J]. 中国编辑, 2007, (1): 42-43.
- [12] 曾群, 龚胜生, 刘建超. 论科技期刊论文审稿人的学术评价 [J]. 编辑学报, 2018, 30 (3): 234-236.
- [13] Garfield E. Patent Citation Indexing and the Notions of Novelty, Similarity, and Relevance [J]. Journal of Chemical Documentation, 1966, 6 (2): 63-65.
- [14] Uzzi B, Mukherjee S, Stringer M, et al. Atypical Combinations and Scientific Impact [J]. Science, 2013, 342 (6157): 468-472.
- [15] 姚飞, 石琴, 陈移峰, 等. 引文分析方法在期刊学术质量评价中的应用 [J]. 重庆大学学报 (自然科学版), 2003, 26 (8): 154-156.
- [16] Chen C, Mayanglambam S D, Hsu F, et al. Novelty Paper Recommendation Using Citation Authority Diffusion [C] // International Conference on Technologies & Applications of Artificial Intelligence, 2012.
- [17] Matsumoto K, Shibayama S, Kang B, et al. Introducing a Novelty Indicator for Scientific Research: Validating the Knowledge-based Combinatorial Approach [J]. Scientometrics, 2021, 126 (8): 6891-6915.
- [18] Hammarfelt B. Using Altmetrics for Assessing Research Impact in the Humanities [J]. Scientometrics, 2014, 101 (2): 1419-1430.
- [19] 赵蓉英, 郭凤娇, 谭洁. 基于 Altmetrics 的学术论文影响力评价研究——以汉语言文学学科为例 [J]. 中国图书馆学报, 2016, 42 (1): 96-108.
- [20] 郝若扬. 高 Altmetrics 指标论文的特征分析及影响力分析 [J]. 图书情报工作, 2018, 62 (8): 107-114.
- [21] 王睿, 胡文静, 郭玮. 高 Altmetrics 指标科技论文学术影响力研究 [J]. 图书情报工作, 2014, 58 (21): 92-98.
- [22] Mishra S, Torvik V I. Quantifying Conceptual Novelty in the Biomedical Literature [J]. Dlib Mag, 2016, 22 (9/10): 1-21.
- [23] 杨京, 王芳, 白如江. 一种基于研究主题对比的单篇学术论文创新力评价方法 [J]. 图书情报工作, 2018, 62 (17): 75-83.
- [24] 曹树金, 曹茹烨. 情报学论文创新性评价研究——LDA 和 SVM 融合方法的应用 [J]. 图书情报知识, 2022, 39 (4): 56-67.
- [25] 侯剑华, 王东毅. 基于 SAO-ADV 模型的学术论文创新性的测度方法研究 [J]. 情报理论与实践, 2020, 43 (11): 129-136.
- [26] 周海晨, 郑德俊, 酆天宇. 学术全文本的学术创新贡献识别探索 [J]. 情报学报, 2020, 39 (8): 845-851.
- [27] 张吉玉, 张均胜. 考虑时序的单篇科技文献新颖性评估方法 [J]. 图书情报工作, 2022, 66 (17): 93-105.
- [28] Alberts B, Hanson B, Kelner K L. Reviewing Peer Review [J]. Nature Immunology, 2008, 321 (5885): 15.
- [29] Macroberts M H, Macroberts B R. Problems of Citation Analysis: A Critical Review [J]. Journal of the Association for Information Science & Technology, 2010, 40 (5): 342-349.

- [50] Sexton A , Shepherd E , Duke-Williams O , et al. The Role and Nature of Consent in Government Administrative Data [J]. Big Data & Society , 2018 , 5 (2) : 1-17.
- [51] Kang H , Zhang Z , Dong J , et al. BeepTrace for COVID-19 Pandemic: A Demo [C] //IEEE. Proceedings of the 2021 3rd Conference on Blockchain Research and Applications for Innovative Networks and Services , BRAINS 2021. Paris , France , 2021: 6-7.
- [52] 최계영. Personal Information Protection in the Medical Field-Centering on the Legislative Systems of the European Union and the USA [J]. 경제규제와 법 , 2016 , 9 (2) : 206-223.
- [53] Ali-Eldin A M T , Zuiderwijk A , Janssen M. Opening More Data a New Privacy Risk Scoring Model For Open Data [C] //SciTePress. Proceedings of the 7th International Symposium on Business Modeling and Software Design , BMSD 2017. Barcelona , Spain , 2017: 146-154.
- [54] Huang H H , Lin J W , Lin C H. Data Re-IdentificationA Case of Retrieving Masked Data from Electronic Toll Collection [J]. Symmetry-Basel , 2019 , 11 (4) : 550.
- [55] Zouinina S , Bennani Y , Rogovschi N , et al. Data Anonymization Through Collaborative Multi-view Microaggregation [J]. Journal of Intelligent Systems , 2021 , 30 (1) : 327-345.
- [56] Badu-Marfo G , Farooq B , Patterson Z. Perturbation Methods for Protection of Sensitive Location Data: Smartphone Travel Survey Case Study [J]. Transportation Research Record Journal of the Transportation Research Board , 2019 , 2673 (12) : 244-255.
- [57] Ali S , Osman T , Mannan M , et al. On Privacy Risks of Public WiFi Captive Portals [Z] //Pérez-Solà C , Navarro-Arribas G , Biryukov A , et al. Data Privacy Management , Cryptocurrencies and Blockchain Technology: ESORICS 2019 International Workshops , DPM 2019 and CBT 2019. Luxembourg; Springer International Publishing , 2019: 80-98. https://doi.org/10.1007/978-3-030-31500-9_6.
- [58] Li W , Meng P , Hong Y , et al. Using Deep Learning to Preserve Data Confidentiality [J]. Applied Intelligence , 2020 , 50 (2) : 341-353.
- [59] Whittington J , Calo R , Simon M , et al. Push , Pull , and Spill: A Transdisciplinary Case Study in Municipal Open Government [J]. Berkeley Technology Law Journal , 2015 , 30: 1899.

(责任编辑: 郭沫含)

(上接第 139 页)

- [30] Liu J , Euan A. Five Challenges in Altmetrics: A Toolmaker's Perspective [J]. Bulletin of the Association for Information Science & Technology , 2013 , 39 (4) : 31-34.
- [31] 荣国阳, 李长玲, 范晴晴, 等. 主题热度加速度指数——学科研究热点识别新方法 [J]. 图书情报工作 , 2021 , 65 (20) : 59-67.
- [32] 杨建林, 钱玲飞. 基于关键词对逆文档频率的主题新颖度度量方法 [J]. 情报理论与实践 , 2013 , 36 (3) : 99-102.
- [33] Luhn H P. The Automatic Creation of Literature Abstracts [J]. IBM Journal of Research and Development , 1958 , 2 (2) : 159-165.
- [34] 李长玲, 高峰, 牌艳欣. 试论跨学科潜在知识生长点及其识别方法 [J]. 科学学研究 , 2021 , 39 (6) : 1007-1014.
- [35] 潘玮, 牟冬梅, 李茵, 等. 关键词共现方法识别领域研究热点过程中的数据清洗方法 [J]. 图书情报工作 , 2017 , 61 (7) : 111-117.
- [36] 高楠, 周庆山. 新兴技术概念辨析与识别方法研究进展 [J]. 现代情报 , 2023 , 43 (4) : 150-164.
- [37] 韩芳, 张生太, 冯凌子, 等. 基于专利文献技术融合测度的突破性创新主题识别——以太阳能光伏领域为例 [J]. 数据分析与知识发现 , 2021 , 5 (12) : 137-147.
- [38] Hwang C , Yoon K. Methods for Multiple Attribute Decision Making [M]. Hwang C , Yoon K. Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey. Berlin , Heidelberg: Springer Berlin Heidelberg , 1981: 58-191.
- [39] Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System [EB/OL]. <https://bitcoin.org/bitcoin.pdf> , 2022-09-08.
- [40] 袁勇, 王飞跃. 区块链技术发展现状与展望 [J]. 自动化学报 , 2016 , 42 (4) : 481-494.
- [41] 杨晓晨, 张明. 比特币: 运行原理、典型特征与前景展望 [J]. 金融评论 , 2014 , 6 (1) : 38-53.
- [42] Garfield E. Citation Frequency as a Measure of Research Activity and Performance [J]. Essays of an Information Scientist , 1973 , 1 (73) : 406-408.
- [43] 中华人民共和国国务院. 国务院关于印发“十三五”国家信息化规划的通知 [EB/OL]. https://www.gov.cn/zhengce/content/2016-12/27/content_5153411.htm , 2022-09-27.
- [44] 中华人民共和国国务院办公厅. 国务院办公厅关于支持国家级新区深化改革创新加快推动高质量发展的指导意见 [EB/OL]. http://www.gov.cn/zhengce/content/2020-01/17/content_5470203.htm , 2022-09-27.
- [45] 罗棋, 闵超, 颜嘉麒, 等. 国际区块链研究主题挖掘及演化分析 [J]. 现代情报 , 2021 , 41 (9) : 157-166.

(责任编辑: 杨丰侨)