

DOI:10.12154/j.qbzlgz.2023.05.003

# ChatGPT 用于论文创新性评价的效果及可行性分析

王雅琪<sup>1</sup> 曹树金<sup>2,1</sup>(<sup>1</sup>中山大学信息管理学院 广州 510006; <sup>2</sup>山东理工大学信息管理研究院 淄博 255000)

**摘要:** [目的/意义]文章旨在探讨ChatGPT用于论文创新性评价的效果与可行性,如果以ChatGPT为代表的生成式人工智能可以用于论文创新性评价,将能够有效提高创新评价效率。[方法/过程]通过模拟真实的论文创新性评价需求和应用场景,选择生命科学、细胞免疫、市场营销、经济学、人工智能、信息资源管理6个研究领域完成超24轮论文创新性评价实验。采用内容分析法,深入分析ChatGPT在论文创新性评价过程中生成的文中本材料,对其创新评价过程、标准与可行性进行研究。[结果/结论]研究发现,ChatGPT在论文创新性评价任务展现出良好性能,评价方法与标准均表现出较高的综合性与科学性,能够较好地完成多学科领域的论文创新性评价。现阶段,ChatGPT用于论文创新性评价的稳定性、准确性与真实性稍显不足,可在相关任务中将其视为高效的论文创新性评价辅助工具。随着提示工程的发展、互联网访问权限的开启、学术数据库的开放获取,ChatGPT用于论文创新性评价的不足预计会被逐渐弥补。

**关键词:** ChatGPT 论文创新性评价 学术成果创新评价

## The Effect and Feasibility Analysis of ChatGPT Used in Paper Innovativeness Evaluation

Wang Yaqi<sup>1</sup> Cao Shujin<sup>2,1</sup> (<sup>1</sup>School of Information Management, SunYat-sen University, Guangzhou, 510006;<sup>2</sup>Information Management Research Institute, Shandong University of Technology, Zibo, 255000)

**Abstract:** [Purpose/significance] This study aims to explore the effectiveness and feasibility of using ChatGPT for evaluating the innovativeness of academic papers. If generative artificial intelligence, represented by ChatGPT, can be employed for paper innovativeness assessment, it has the potential to significantly improve the efficiency of innovativeness evaluation. [Method/process] By simulating real-world demands and application scenarios for evaluating paper innovativeness, we conducted over 24 rounds of innovativeness evaluation experiments across six research domains: life sciences, cellular immunology, marketing, economics, artificial intelligence, and information resource management. Content analysis was employed to thoroughly examine the text generated by ChatGPT during the process of evaluating paper innovativeness, enabling us to study its approach, criteria, and feasibility. [Result/conclusion] The study revealed that ChatGPT demonstrated excellent performance in the task of evaluating paper innovativeness, displaying high comprehensiveness and scientific rigor in its evaluation methods and criteria. It effectively accomplished paper innovativeness evaluations across these research fields. At present, while ChatGPT's stability, accuracy, and authenticity may have some minor limitations, it can still be considered an efficient auxiliary tool for evaluating paper innovativeness in relevant tasks. The identified shortcomings of using ChatGPT for paper innovativeness evaluation are likely to be gradually addressed with the advancement of prompt engineering, opening up of Internet access, and increased availability of academic databases.

**Keywords:** ChatGPT paper innovativeness evaluation academic research innovation evaluation

## 1 引言

科技创新背景下,创新评价是信息资源管理领域的热门研究方向之一。对学术论文的创新价值进行评价,能够有效促进学术研究良性发展。其中,最传统的论文创新性评价方法是同行评议法<sup>[1]</sup>,但这种定性评价方法难以适应学术文献的爆炸性增长,更难避免因评审人认知偏差出现的创新保护不足问题<sup>[2]</sup>。为进一步解决论文创新评价的非公正性、非客观性和非合理性问题<sup>[3]</sup>、提高评价效率,学界开始探究如何利用文献的计量特征进行评价<sup>[4]</sup>,并逐渐将分析重点转向论文文本内容<sup>[5]</sup>。随着大数据时代到来,深度学习以其强大的特征学习与表达能力,以及在没有任何预先假设情况下能捕获信息之间潜在关系<sup>[6]</sup>的特点,逐渐开始被应用于创新评价相关研究领域。如果能够将基于深度学习的大语言模型引入论文创新性评价领域,它能否做出有效回答?生成的评价是否合理?是否能够部分或完全取代人工在学术创新评价中的作用?

本研究致力于探索 ChatGPT 为代表的生成式人工智能进行论文创新性评价的过程、标准、效果与可行性。具体来说,本研究基于层递式提问方法对选择的六个学科领域进行多轮 ChatGPT 论文创新性评价实验,采用内容分析法从回答文本中提炼 ChatGPT 论文创新性评价的标准与方法,并从稳定性、准确性与真实性三个角度分析 ChatGPT 用于论文创新性评价的可行性。

## 2 学术论文创新性评价相关研究

学术论文的创新性评价源于对论文创新性的识别与衡量。除同行评议外,论文创新性评价方法主要可概括为基于(外部)数据和基于(论文)内容两类。

引文数据是论文创新性评价的重要组成部分,Uzzi 等<sup>[7]</sup>发现共被引网络中存在现有知识的非常规组合,其中存在具有创新价值的论文;宋歌<sup>[8]</sup>设计了基于原始引文网络计算的 S 指数,从创新成果的变异程度和对科学发展的影响两方面去判断成果创新力;魏瑞斌等<sup>[9-10]</sup>将学术论文通过引用关系连接在一起,结合时间因素分析论文主题创新性,并基于引证意图评价被引证文献的创新性。

基于内容的论文创新性评价方法可以依照技术手段差异进行分类。有学者以向量空间模型为基础,通过对比文本相似度计算文献新颖性<sup>[11]</sup>;通过对文献内词

句的量化与计算,也可以实现论文的创新评价<sup>[12-14]</sup>。随着深度学习不断发展,有学者开始研究如何利用机器学习识别文献创新内容<sup>[15]</sup>。此外,更有学者致力于研究论文内容表现出来的创新性特征。例如李贺和杜杏叶<sup>[16]</sup>提出通过构建学术论文的知识元本体,从研究问题(研究主题)、理论、方法和结论四个维度计算论文创新性评分。曹树金和曹茹焯<sup>[17]</sup>提出一种融合 LDA 和 SVM 的论文主题创新性评价方法,提出高创新价值论文发表时间会明显早于该研究主题发表高峰期的特征规律。

还有学者综合使用外部数据与论文内容同时完成创新性评价方法的研究。例如,杨京等<sup>[18]</sup>聚焦单篇论文的研究主题与科学研究前沿的相似度计算,并结合 IF 指标和 Altmetrics 指标构建学术论文的创新力综合评价模型;索传军等<sup>[19]</sup>提出将建立学术论文语料库和作者 h 指数、期刊影响因子等相关数据,利用机器学习技术明确优秀论文的模式与特征,实现对学术论文的创新评价;秦岩等<sup>[20]</sup>综合论文的引用关系指标和文本内容指标,通过计算新旧参考文献在论文参考文献数量中的比率判断论文的吸收新颖性,并利用论文的标题和摘要计算论文产出的新颖性。

可以看出,信息资源管理领域在学术成果创新评价的体系、标准与方法等方面的研究成果颇多,相关理论与技术不断发展,形成了较为成熟的学术成果创新评价知识,为 ChatGPT 进行论文创新性评价提供了基础。

## 3 ChatGPT 用于论文创新性评价的实验

本节通过对 ChatGPT 完成论文创新性评价的过程进行深入提问,分析其论文创新性评价能力。OpenAI 开发的 ChatGPT 是一种基于深度学习的人工智能模型,学习的知识深度与宽度远超过个体。目前,较多学术论文创新评价方法研究仍基于“相似度计算”,如对单篇论文主题与学科前沿主题进行相似度计算<sup>[18]</sup>。ChatGPT 使用基于 Transformer 结构的深度神经网络,能够学习输入数据的复杂特征和规律,理解人类语言的含义和上下文,输出连贯的文本,这也使其具备了学术论文创新性评价的能力。在论文创新性评价任务中,ChatGPT 依靠基于深度学习的自然语言处理算法,能够:(1)对海量文本进行自动分类、聚类、关联等任务,从而实现信息的自动化组织、管理与生成;(2)通过计算两篇论文的相似度来评价它们的创新性;(3)分析论文中的语言表达方式,做出创新性评价;(4)根据论

文中的内容,生成一些新的文本内容,并根据这些内容给出创新性评价等。

本研究设计了一套完整的论文创新评价实验,并根据设定好的实验任务将涉及的提问予以明确与规范,提高实验严谨性。为增加结论普遍性,选择生命科学、细胞免疫、市场营销、经济学、人工智能、信息资源管理六个不同的研究领域,通过访问 ChatGPT,对六个领域进行超 24 轮论文创新性评价实验,提问次数超 300 次,得到问答文本超 33 万字符。由于篇幅有限,选择人工智能领域论文创新性评价实验中的一次测试过程,对实验及效果进行说明。

### 3.1 任务一:针对单篇未发表论文进行创新性评价

选择 2023 年 3 月 29 日网络首发于《计算机应用研究》的论文“RISNet:无监督真实场景图像拼接网络”为实验对象,模拟 ChatGPT 无法获得全文的情况,测试其能否仅根据输入的论文内容进行创新评价。

用户:请根据摘要,对这篇论文的创新性做出判断。

摘要:图像拼接目的是获得一张高清无缝的全图,现有方法依赖于特征匹配的准确性,会……为此提出一种新的无监督真实场景图像拼接网络……实验结果表明……RMSE、PSNR、SSIM 分别达到 1.81、26.56、0.85……

ChatGPT:根据摘要,这篇论文提出了一种新的无监督真实场景图像拼接网络……客观评价指标上整体优于其他经典算法……基于这些信息,我认为这篇论文具有一定的创新性。它提出了……新方法,并且在客观评价指标上整体优于其他经典算法……

用户:你是如何根据摘要,判断这篇论文的创新性?

ChatGPT:……先对给出的论文摘要进行自然语言理解和分析,然后参考已有的领域知识和最新研究进展,结合领域专家的意见……

用户:请告诉我图像拼接领域相关学术论文中,指标表现最高的一篇文章。

ChatGPT:经过查阅相关文献,目前表现最好的一篇文章图像拼接相关的学术论文是“End-to-End Learning of Geometry and Context for Deep Stereo Regression”(CVPR 2018)

实验结果表明,当向 ChatGPT 提供论文摘要时,ChatGPT 首先将使用自然语言处理算法对输入的文本

进行理解与分析,并根据从摘要文本中提取出的研究领域与方向,在训练数据集中进行领域知识与研究进展的搜索,并与待评价论文进行对比,并将结合可获知的领域专家对相关研究的评价意见等,形成论文创新性评价内容。经过多领域、多轮次、多篇学术论文的测试,本研究发现:ChatGPT 能够仅根据提供的部分重要信息如论文摘要、结论等进行论文创新评价,且其评价逻辑较清晰,具有一定可信度。

### 3.2 任务二:针对创新性相近的论文进行推荐

对任务一回答的论文题目“End-to-End Learning of Geometry and Context for Deep Stereo Regression”进行检验,确保 ChatGPT 给出的论文信息正确后,测试 ChatGPT 是否能够推荐创新性相近的论文。

用户:请给我推荐一篇与 End-to-End Learning of Geometry and Context for Deep Stereo Regression 创新性相似的学术论文

ChatGPT:如果您对深度学习和计算机视觉领域的深度估计和立体匹配问题感兴趣,我可以为您推荐以下几篇论文:

1."Learning to Estimate 3D Human Pose and Shape from a Single Color Image" by ShubhamTulsiani, Srabab Goswami, David F. Fouhey, and Alexei A. Efros. (CVPR 2018)

2."Deep Stereo: Learning to Predict New Views from the World's Imagery" by Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. (NIPS 2012)

3."DispNet: Stereo Convolutional Neural Network for Learning Disparity Estimation" by Clement Godard, Oisin Mac Aodha, and Gabriel J. Brostow. (CVPR 2017)

4."Unsupervised Learning of Depth and Ego-Motion from Video" by Tinghui Zhou, Matthew Brown, Noah Snavely, and David G. Lowe. (CVPR 2017)

由于 ChatGPT 是生成式人工智能,实验对 ChatGPT 给出的论文信息逐篇进行人工检验。经过核实,发现论文 1 的题目与发表年份正确,但作者信息错误;论文 2 的题目信息可查,但发表年份与作者信息均错误;论文 3 的所有信息均错误;论文 4 的题目、发表年份、作者信息均正确。在论文推荐过程中,ChatGPT 给出的论文信息不完全准确,甚至生成了虚假论文。推测 ChatGPT 给出论文信息均准确的这类文献(如论文 4),其除题目以外的相关重要信息(摘要、全文等)很大概率能够被

ChatGPT获取。任务一验证了ChatGPT在完成论文创新性评价时至少需要能够对论文的摘要或结论内容进行分析与处理,如果对重要信息缺失的论文进行评价,ChatGPT极大可能“杜撰”论文内容,评价结果的可信度大幅下降。

### 3.3 任务三:针对多篇论文进行创新性对比

对前两项任务进行十余次反复测试后,发现当被评价论文为ChatGPT在数据库中能够获取相关重要信息的论文时,评价结果的可信度最高。因此,使用ChatGPT给出论文信息均准确的两篇论文“End-to-End Learning of Geometry and Context for Deep Stereo Regression”“Unsupervised Learning of Depth and Ego-Motion from Video”为例,检验ChatGPT能否完成论文创新性的对比。

用户:请对比以下两篇学术论文,告诉我哪一篇更具有创新性:

1. Unsupervised Learning of Depth and Ego-Motion from Video
2. End-to-End Learning of Geometry and Context for Deep Stereo Regression

ChatGPT:这两篇论文都具有创新性,但是从技术角度来看,“Unsupervised Learning of Depth and Ego-Motion from Video”更具有创新性。

“Unsupervised Learning of Depth and Ego-Motion from Video”提出了……这种方法不需要任何深度标注数据或传感器,因此可以很好地适用于大量的自然场景数据。

相比之下,“End-to-End Learning of Geometry and Context for Deep Stereo Regression”的贡献在于……但是这个问题之前也已经被研究了一段时间,并且已经存在了一些相关工作。

因此,从技术贡献和创新性来看,“Unsupervised Learning of Depth and Ego-Motion from Video”更具有创新性。

对六个领域的多篇论文进行创新性对比试验后,发现ChatGPT能完成多领域的论文创新性对比任务。前期文献调研中,发现引用情况与Altmetrics指数常被研究者作为论文创新性评价的依据之一。为评估ChatGPT生成内容质量,考察被评价论文的被引频次与Altmetrics指数,详见表1。

ChatGPT认为更具有创新性的论文,其被引频次

表1 论文被引频次与Altmetrics指数

项目	WOS被引频次	Altmetrics指数
论文1	1158	9
论文2	766	15

也更高。但两篇论文的Altmetrics指数均偏低,说明反映信息网络传播热度的Altmetrics指数可能不适合用于这两篇学术论文创新性的验证。每轮任务三实验中均对被评价论文的被引频次进行考察,发现基本符合原结论,说明ChatGPT给出的论文创新性比较结果具有一定可靠性,后续分析也发现被引量是ChatGPT进行论文创新性评价的标准之一。

### 3.4 任务四:给出多篇论文创新性评分

任务四是任务三的延伸,在论文创新性比较的基础上测试ChatGPT能否给出具体的创新性评分。为提高论文创新评分任务的完成度与准确性,选择十篇ChatGPT给出论文信息均正确的人工智能领域文献,测试ChatGPT在多篇论文创新性评分任务中的表现。

用户:请给出这十篇论文的创新性评分,最高分10分,最低分0分。

1. Generative Adversarial Nets
2. Deep Residual Learning for Image Recognition
3. Neural Machine Translation by Jointly Learning to Align and Translate
4. Playing Atari with Deep Reinforcement Learning
5. Mastering the game of Go with deep neural networks and tree search
- .....

ChatGPT:

1. Generative Adversarial Nets - 创新性评分:10分
2. Deep Residual Learning for Image Recognition- 创新性评分:9分
3. Neural Machine Translation by Jointly Learning to Align and Translate - 创新性评分:9分
4. Playing Atari with Deep Reinforcement Learning- 创新性评分:8分
5. Mastering the game of Go with deep neural networks and tree search- 创新性评分:10分
- .....

本次测试的数据中,论文1与论文5创新性获评10分,十篇论文创新性平均分达8.4。层递式提问过程中,根据ChatGPT回答的“思路和方法创新性”“解决问

题的效果和影响”“引用量和应用价值”创新性评分维度,让ChatGPT再次对十篇论文进行创新性评分,各维度单项最高10分。结果显示,原评分最高的两篇论文在各维度获评分数也最高,同时得到了最高的总评分(29分),两次评分结果总体相近。将10篇论文的论文创新性评分与被引频次、Altmetrics指数比较,情况如表2所示。

表2 论文被引频次、Altmetrics指数与创新性评分

项目	WOS被引频次	Altmetrics指数	创新性评分
论文1	4880	缺失	10
论文2	49775	79	9
论文3	9791	缺失	9
论文4	761	缺失	8
论文5	7996	3093	10
论文6	1227	缺失	8
论文7	3615	59	7
论文8	1851	缺失	7
论文9	1402	缺失	8
论文10	676	缺失	8

研究发现,简单引入外部评价数据衡量ChatGPT评价质量的效果不佳,将在后续可行性分析部分集中分析ChatGPT用于论文创新性评价的准确性与质量。此外,研究发现无论是对相同论文多次提问,还是部分更换列表中的待评价论文,ChatGPT给出的论文创新性评分均会存在10%~15%的浮动,推测是由算法与技术特性、评价标准不稳定等原因导致。

3.5 任务五:针对单篇可获取的论文做出创新性评价

可获取全文等所有重要信息时,ChatGPT的论文创新性评价维度与方法是否会存在差别?从任务四的十篇论文中选择一篇继续测试。

用户:请你对Generative Adversarial Nets的论文创新性进行评价。

ChatGPT:以下是我对该论文的创新性评价:

研究问题:GAN论文提出的问题是生成模型中的困难问题……

创新点:……取得了比传统模型更好的效果。

实验结果:GAN……取得了非常优秀的实验结果……还启发了后续的很多研究,如……

参考文献:……展示了其对相关领域的深入研究。

文章质量:GAN论文的撰写质量非常高……思路清晰、论证严密……

影响力:GAN论文在深度学习领域产生了巨大的影响力,被广泛引用和应用……对比ChatGPT在所有轮次实验任务一与任务五中的回答,发现当ChatGPT

仅能获取论文部分重要信息(如摘要、结论),创新性评价主要围绕文本中的关键内容分析展开;如果能获取更多的论文信息(如全文、参考文献、引用数据等),论文创新性评价角度更加全面。

3.6 任务六:针对论文提出进一步创新的建议

如果论文创新性评价是审稿、学术绩效考核的重点,那么发现论文进一步研究的方向、提出切实可行的创新建议则是研究人员关注的焦点。为方便判断ChatGPT提出论文创新建议的合理性,以一次信息资源管理领域的论文创新建议实验进行效果说明。以2023年4月7日网络首发的论文“基于引证意图的学术论文创新性评价研究”和2023年2月24日发表于《情报学报》的论文“基于机器学习模型的科技论文潜在‘精品’识别研究”为例,测试ChatGPT能否给出学术论文创新建议。

用户:请你针对以下论文摘要,对这篇论文给出进一步创新的建议

摘要:[目的/意义]学术论文创新性是学术研究重要议题。深入引证文献引证内容,揭示引证文献作者的引证意图……

ChatGPT:根据摘要描述,这篇论文的创新点主要在于提出了一种基于引证意图评价学术论文创新性的新方法,并且构建了一个基于引证意图的学术论文创新性评价模型……以下是一些可能的创新性建议:

1.对于论文的引证意图如何进行抽取和分类?是否有新的方法或技术应用?

2.CRITIC权重法在该研究中扮演了什么角色?是否有其他的方法或算法可以用于构建基于引证意图的学术论文创新性评价模型?

3.摘要中提到了考虑了引证意图的引证位置、引证期刊等级、引用情感等因素,这些因素是如何被纳入评价模型的?在后续的研究中,是否可以考虑加入更多的因素,如作者影响力、学科影响力等?

用户:以下是一篇学术论文的论文摘要,请根据摘要内容对这篇论文进一步应该如何创新提出建议。

摘要:综合运用科技文献特征向量空间和机器学习模型实现海量文献中潜在“精品”的自动识别与推荐……

ChatGPT:这篇论文的研究方向是运用科技文献特征向量空间和机器学习模型实现海量文献中潜在“精品”的自动识别与推荐……

进一步的创新建议包括:

1. 建议在机器学习模型中加入更多深度学习算法……以提高潜在“精品”的自动识别精度;

2. 建议研究者可以探索多种特征工程方法,如词向量、主题模型等;

3. 建议进一步研究和探索如何将机器学习模型应用到其他领域的文献识别中,如医学、经济等领域,以拓宽文献识别的应用范围;

4. 建议研究者在“精品”论文的计量特征值研究的基础上,进一步探索高质量论文的内容特征……

邀请三位信息资源管理领域专家对多次实验的结果进行评估后,认为 ChatGPT 对论文关键内容的总结基本正确,给出的论文进一步创新建议具有较高的参考价值,推测其是在提取论文关键内容的基础上,结合相同学科、相似研究领域的常见研究方法与现状热点,推断论文可能存在的不足,给出进一步的创新建议。

本研究从真实的学术论文创新评价需求与应用场景出发,主要围绕生命科学等六个领域与 ChatGPT 展开交互,每个研究领域进行4轮完整实验。实验结果表明,面向生命科学、细胞免疫、市场营销、经济学、人工智能、信息资源管理领域的学术论文,ChatGPT 能够完成单篇论文的创新性评价与多篇论文的创新性比较,对论文的创新性评分较稳定,能给出较为合理的论文创新建议。同时,ChatGPT 在评价过程中对论文的研究方法、研究对象、主要研究结论、价值意义与创新点等关键内容进行抽取,帮助用户了解论文内容并判断评价结果合理性。此外,ChatGPT 在设计实验任务中均能给出答案,但部分任务完成质量不佳,需进一步分析其原因与解决方案。

#### 4 ChatGPT 的论文创新性评价标准

本节使用内容分析法对 ChatGPT 进行单篇论文创新性评价、多篇论文创新性比较时使用的判断依据进行深入研究,主要回答以下问题:ChatGPT 如何进行论文创新性评价?论文创新性评价标准是否科学合理?是否学习了信息资源管理领域的研究成果?面对不同领域的论文,论文创新性评价标准是否不同?

使用层递式提问技巧对单篇论文创新性评价的步骤进行反复提问,采用内容分析法对获得的文本内容进行分析,发现无论被评论文的相关重要内容是直接提供或需要在数据库查询得到,ChatGPT 在论文创新性

评价中均遵循“分析关键信息-理解全文内容-评价创新点-对比差异-得出结论”基本步骤。具体包括:(1)通过论文摘要和引言部分,了解论文的研究领域和背景;(2)深入阅读论文,把握研究内容、研究方法、实验结果等关键部分;(3)理解论文内容基础上,评价论文在该领域中的创新点;(4)对比和评价其他相关论文,寻找在创新性、领域影响和贡献方面的差异;(5)综合各方因素给出该论文的创新性评价,并在需要的情况下给出对该论文进一步研究的建议。

当无法获取论文全文时,ChatGPT 会“根据论文的题名、摘要和关键词等信息,了解论文所涉及的研究领域、研究问题和研究方法等方面的信息,然后结合相关领域的背景知识,来尽可能评价论文的创新性”,并在答案中“注明评价的可靠程度和限制”。反之则会在此基础上增加对论文影响力与实际应用情况的判断。研究进一步发现,ChatGPT 的论文创新性评价标准呈现动态统一。经过编码,ChatGPT 最普遍的论文创新性评价标准及相应解释如下:

(1)拓展或改进已有知识:一篇具有创新性的论文通常能够对已有的知识进行拓展或改进,从而提出新的研究问题或解决方案。

(2)存在新颖性和独创性:一篇具有创新性的论文通常提出了新的思路、新的方法或新的理论,从而具有独创性和新颖性。

(3)解决实际问题:一篇具有创新性的论文通常能够解决实际问题,为相关领域的实践带来新的思路 and 方案。

(4)促进领域发展:一篇具有创新性的论文贡献往往与该领域的最新发展趋势相符合,能够为相关领域的发展带来新的视角、新的思路和新的方法,能够被其他研究者广泛引用和讨论,从而对领域发展产生积极、深远的影响。

曾有学者明确指出,“创新既是一种特征属性,更应该是包含有用、正确等测量角度的价值变量”<sup>[21]</sup>,创新性不应仅判断新颖性还应该判断有用性<sup>[22]</sup>。

ChatGPT 基于文献内容对学术论文的新颖度进行评估,符合创新性判断的第一个要求。索传军<sup>[23]</sup>指出,一篇包含创新知识元的学术论文就能够被称为是具有创新性的,绝对创新度的测度应将其创新知识元与整个科学知识体系或技术体系或科学研究方法体系进行比较,判断论文具有多少新发现、新观点、新方法或新

技术。对问答文本进行深入分析后,发现 ChatGPT 一方面能够对相关文本进行识别、抽取和组织,对文献中表示创新的词句进行判断,从句法逻辑和文本语义两方面比较研究方法创新性;另一方面能够通过前期分析、处理大量的数据与文献,形成相关领域的知识库,分析论文的研究方法、算法与知识库中其他的方法有何差异或继承。另外,ChatGPT 可以从社交平台、学术交流网站等途径获取其他专家学者对论文的评价内容,对其中的创新性评价部分进行理解分析,帮助评价论文创新价值。

其次,ChatGPT 关注对实际问题的解决和领域发展的促进,符合创新性判断的第二个要求。进行创新性判断时,ChatGPT 能够“结合相关领域的发展情况,对研究中提出的想法和方法进行比较和推理”,还包括“了解该论文的引用次数和引用者,反映论文在该领域的影响力和重要性”“查看该方法在该领域的应用和影响,了解其对该领域发展的影响和推动作用”“参考该领域的权威人士意见”等。不难发现,ChatGPT 评估学术论文有用性与影响力的方法具有一定综合性。除参考常规引用指数等学术影响力指标以外,ChatGPT 能够通过识别待评价论文与其他论文的相关文本内容、综合领域专家、学者和科学界发表的相关信息,在了解领域发展趋势和重要研究方向基础上,对比、判断待评价论文的有用性与影响力,同时还设法获取论文在学术领域获得的评价,以及成果在现实中的应用情况。

同时,发现 ChatGPT 完成论文创新评价时的标准与方法与信息资源管理领域相关研究中提出的标准与方法十分相似。例如,信息资源管理领域学者曾综合领域专家智慧总结出特定研究领域的研究水平特征词库后,利用特定算法抽取待评价论文中代表研究水平的句子与数值信息<sup>[24]</sup>;利用深度学习 Doc2Vec 算法生成特定研究领域的论文热点主题描述<sup>[15]</sup>;从学科影响力与社会影响力两方面构建论文学术影响力评价指标<sup>[25]</sup>;利用学术论文的被引频次和使用次数评价中美图书情报学领域国际学术论文影响力<sup>[26]</sup>;通过判断论文的内容区分度,识别容易成为研究热点和前沿应用的论文<sup>[27]</sup>等。同时,研究发现 ChatGPT 的学术论文创新评价标准并未包含所有信息资源管理领域的相关知识,例如从他人对该论文的应用内容角度进行创新性分析,在被引次数指标的基础上考虑引用情感、引用目的<sup>[28]</sup>等方法并未被 ChatGPT 使用。

综上,本研究认为 ChatGPT 的论文创新评价能力基于现有的、成熟的论文创新性评价知识,评价标准与方法的综合性较强,加强了对论文内容的分析。无论是从内容出发分析新颖度、先进性以及对领域的影响程度等,还是获取领域发展成果并与被评价论文进行对比、获取论文外部评价内容辅助判断,ChatGPT 在评价过程中表现出了强大的自然语言处理与分析能力。此外,研究发现 ChatGPT 掌握的论文创新性评价知识复杂而庞大。多轮论文创新性评价实验中,ChatGPT 每次仅选择其中部分评价标准做出回答,内部缺少一套科学的方法论将这些论文创新性评价知识组织起来。

为回答 ChatGPT 论文创新性评价标准的学科领域差异问题,研究选择五个集中测试时间段,在每个时间段集中对六个研究领域的论文创新评价标准进行提问,最终得到 5 组、30 份论文创新性评价标准,不同组别代表测试的时间段不同,部分内容见表 3。

对 30 份论文创新性评价标准进行横向和纵向对比,发现 ChatGPT 在每个时间段集中给出同一组论文创新性评价标准的角度较相似,不同组别之间的论文创新性评价角度差异较大。

ChatGPT 面向不同领域的论文创新性评价标准既存在共性,又存在个性。共性方面,ChatGPT 评价论文创新性时普遍包括研究对象或问题的创新性、理论或思想的创新性、方法与技术的创新性、应用价值与实用性、研究贡献与学术影响五个评价角度。个性方面,面向不同领域时论文创新性评价标准的学科特征比较明显。例如:面向经济学领域时,评价标准包括“评判论文所提出的经济理论和研究成果是否能够长期存在并发挥作用,是否具有一定的可持续性和可扩展性”;面向人工智能领域时,评价标准包括“实验是否具有可重复性和可验证性”“是否提供了足够详细的实验描述和数据共享”;面向细胞免疫学领域时,评价标准包括“实验设计是否具有科学性”“数据分析是否符合统计学原则与规范”“结果是否经过了多次实验验证和不同实验室的重复验证”等。

“破五维”背景下,构建新科研评价体系的原则表现在偏重分析论文全文内容,弱化期刊影响因子、资助效益、被引次数等外在数据,强调科学研究内容的创新性<sup>[29]</sup>。ChatGPT 在论文创新性评价过程中对全文内容的分析深入,并在标准中同时综合较多的评价维度与衡量因素,符合新科研评价体系的价值导向。

表3 ChatGPT在不同研究领域使用的论文创新性评价标准(简略)

领域	组别	论文创新性评价标准
经济学	第2组	研究问题:……是否针对当前经济学领域中的重要问题或挑战进行了研究? 研究问题是否有新颖的角度或视角? 研究方法:……研究方法是否创新? 该方法是否能够有效解决研究问题? 研究结果:……研究结果是否新颖? ……为经济学领域提供有价值的洞见或启示? 对现有研究的贡献:……重要的贡献? ……对现有理论和研究进行扩展和改进? 实用性:……是否具有实际应用价值? 是否能够对实际经济活动进行指导和启示? 对比其他论文:……是否能够与其他相关领域的论文进行对比? ……比其他论文更为显著?
	第5组	理论贡献:……理论是否是全新的, ……为经济学领域提供新的视角和思考方式, 是否有助于解决当前经济学面临的问题。 实证贡献:……实证研究是否能够提供新的证据, 是否具有说服力, 是否能够支持或者反驳经济学上的一些假设。 方法创新:……研究方法是否具有创新性, 是否能够为经济学研究提供新的思路和方法。 数据贡献:……数据集是否新颖、全面、可靠, 是否能够提供新的经验材料和经济数据分析方法。 可操作性:……是否能够指导经济实践, 是否有助于决策者或者投资者…… 实用性:……是否具有实际应用的前景和潜力, 是否能够为经济学领域的相关领域带来实际的经济效益。 可持续性:……经济理论和研究成果是否能够长期存在并发挥作用……可持续性和可扩展性?
人工智能	第2组	研究问题:……是否针对当前……重要问题或挑战进行了研究? ……新颖的角度或视角? 研究方法:……研究方法是否创新? 该方法是否能够有效解决研究问题? 研究结果:……是否新颖? 是否能够为人工智能领域提供有价值的洞见或启示? 对现有研究的贡献:……是否具有重要的贡献? ……能够对现有理论和研究进行扩展和改进? 实用性:……是否具有实际应用价值? 是否能够对实际人工智能研究和应用……? 对比其他论文:……研究结果和贡献是否比其他论文更为显著? 技术难度:……研究技术是否具有难度? 技术创新是否达到了当前领域的前沿水平? 实验设计:如果涉及到实验, ……设计是否合理? 实验过程是否严谨?
	第5组	创新性思想:……方法、算法、模型等是否有新颖的思想和创意, 是否超越了已有研究成果。 技术难度:……是否具有一定的技术难度, 是否需要克服一些技术上的困难。 实验效果:……是否在实验中取得了良好的效果, 是否比已有的方法具有更好的性能。 实际应用:……是否具有实际应用价值, 是否能够解决实际问题, 是否具有商业化前景。 可复现性:……实验过程是否能够被其他人重现, 是否提供了足够详细的实验描述和数据共享。 可推广性:……是否具有一定的普适性和可推广性, 是否能够应用于其他领域和问题。 对领域发展的贡献:……领域的发展是否具有一定的贡献, ……推动该领域进一步发展。

实验中,我们并未对 ChatGPT 的创新性评价能力进行预训练。依靠强大的自然语言处理技术与计算推理能力, ChatGPT 依靠掌握的相关知识, 将引用指标、论文内容、领域知识背景、研究热点、应用效果、领域贡献等结合起来作为论文创新性评价标准, 并能根据实际情况和研究领域进行标准与方法调整。庞大算力支持下, ChatGPT 在论文创新性评价任务中的高性能与高效率, 是传统小模型较难达到的。

## 5 ChatGPT用于论文创新评价的可行性分析

为详细论证 ChatGPT 用于论文创新性评价的可行性, 本节将从稳定性、准确性与真实性三个方面衡量 ChatGPT 进行论文创新性评价的效果。

### 5.1 稳定与不稳定并存

对 ChatGPT 在实验中的所有回答文本进行分析, 发现 ChatGPT 论文创新性评价过程中稳定性与不稳定性并存。

首先, 最大的不稳定表现为 ChatGPT 论文创新性评价标准不稳定, 即 ChatGPT 每次仅从掌握的相关知识中提取部分作为单次创新性评价的标准。这种标准的不稳定性直接导致 ChatGPT 给出的论文创新评价结果不稳定。但研究也发现, ChatGPT 在实验中存在一套最普遍的论文创新评价标准, 这导致 ChatGPT 给出的论文创新评价结果虽然不完全稳定, 但偏差不会过大。重复对同一篇或同一批学术论文进行创新性评价时, ChatGPT 给出的创新评价内容相近, 给出的论文创新性评分存在 10%~15% 浮动。

其次, ChatGPT 完成论文创新性评价的步骤基本稳定, 即“分析关键信息-理解全文内容-评价创新点-对比差异-得出结论”。

### 5.2 评价准确性较高

本研究对过往论文创新性评价相关研究对结果的验证方法进行了文献调研, 发现基本没有作者在给出论文创新性评价结果后进行系统性验证, 更多是通过定性方法去论证结果准确性。例如, 有学者通过论证得出的引用科学知识突变主题在随后几年确实出现, 证明其利用引用科学知识突变去识别突破性创新的方法有效<sup>[30]</sup>; 还有学者通过论证被评价为高创新性的多个论文主题贴合当前领域研究热点, 证明其颠覆性技术早期识别方法的准确性<sup>[31]</sup>。本研究推测, 过往研究无法系统、直观地验证创新性评价结果的准确性, 很可能是因为研究往往针对单一学科领域展开, 形成的创新性评价知识仅适用于同一领域的论文评价, 而不同学者选择研究的论文学科领域不同, 因此难以利用相同的方法与知识对其他学者的实验数据进行分析, 更难将得到的论文创新评价结果与其他学者得出的结果进行对比。

相反地, ChatGPT 无需经过预先训练就能够面向多个研究领域完成论文创新性评价任务, 基本不存在创新性评价知识仅适用于单一领域的限制。因此, 本研究选择从过往相关研究中提取可被获取的实验数据, 利用 ChatGPT 对这些实验数据再次进行创新性评价, 比较 ChatGPT 的评价结果与原结果之间的差异, 并尝试找出形成差异的原因。为避免选择的相关研究能够被 ChatGPT 获取, 降低结果可信度, 研究选择 2021 年

之后发表的四篇中文论文作为数据来源, 论文的实验数据均为英文文献。使用 ChatGPT 对相同的实验数据(学术论文)进行创新性评价, 情况如表4所示。

沈超<sup>[32]</sup>在研究中从参考文献、论文文本、论文评价、论文关联内容4个维度评价论文的总创新力, 并在论文写作中明确了其中2篇实验数据的论文题目。楼雯等<sup>[33]</sup>在研究中对3篇物理学领域文献的创新性进行比较, 认为横向对比分析中论文2与论文3更加优秀, 纵向对比分析中论文1表现最弱, 论文2对学科发展影响大。由于原文并未给出具体创新性评分, 因此综合评价内容分别对3篇论文给出创新性较低、高、较高的评价。利用 ChatGPT 对以上案例提供的5篇实验数据进行创新性评价, 得到的新结果与原结果相近, 特别是在同一案例中, 论文间的创新性排名与之前完全一致, 这说明 ChatGPT 的论文创新性评价结果具有一定准确性。

当同时评价的论文数量增加后, ChatGPT 给出的论文创新性评价与原结果出现一定差距。谢珍等<sup>[34]</sup>在研究中利用创新性测度方法对一篇诺贝尔获奖论文和同作者其他论文进行分析, 使用 ChatGPT 对论文中提及的所有文献进行二次评价, 发现 ChatGPT 给出的创

新性评价趋势较为准确, 但部分评价存在差异(论文2与论文7)。

魏绪秋等<sup>[10]</sup>基于引证意图构建论文创新性评价模型, 对87篇学术论文进行创新测度, 并在论文写作中明确了其中9篇实验数据的论文题目。利用 ChatGPT 对这9篇论文进行创新性评价, 发现 ChatGPT 给出的评价结果与原结果有很大差异。

魏绪秋等构建的创新性评价模型包括引证意图、引证位置、引证期刊等级、引用感情四个主要因素, 但这四个关键因素并不是 ChatGPT 最普遍的评价标准。因此, 推测 ChatGPT 给出的创新评价与原结果出现较大差异是因为创新评价标准不同。在此基础上, 测试 ChatGPT 能否基于给定的标准完成创新性评价, 如图1所示。

要求 ChatGPT 根据给定的论文创新性评价标准, 对相同论文再次进行创新性评价, 比较原评价结果与 ChatGPT 两次论文创新性评价结果间的差异。

如图2所示, ChatGPT 两次给出的创新性评价结果具有很明显的差异, 说明 ChatGPT 在第二次评价中有很大概率遵循了给定的论文创新性评价标准(引证意图、引证位置、引证期刊等级、引用感情), 根据实际需要向 ChatGPT 规定论文创新性评价标准是可行的。但

本研究仅向 ChatGPT 规定了评价维度, 推测这可能是导致第二次给出的评价结果仍与原结果差异较大的重要原因之一, 后续研究可以尝试向 ChatGPT 明确原评价模型的各项计算细则。与此同时, 在21篇实验数据中随机抽查了3篇论文, 发现 ChatGPT 给出的论文相关信息与真实情况不符, 现阶段 ChatGPT 给出论文创新性评价的准确性受诸多因素影响。

### 5.3 通过丰富提问增加真实性

向 ChatGPT 输入的论文相关信息越少, ChatGPT 论文创新性评价的真实性越容易受到影响。本研究验证了 ChatGPT 能够对指定的一篇或多篇学术论文进行创新性评价, 但创新性评价过程中也存在真实性问题。研究发现, 虽然进行论文创新性评价时 ChatGPT 几乎不受领域知识限制, 但回答的真实性与其掌握的领域知识是否完整、可获取的论文相

表4 论文创新性评价结果对比

篇名	ChatGPT 评价结果 (百分制)	原评价结果	数据来源
An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.	95	0.24	沈超 (2020) <sup>[32]</sup>
Iterative Empirical Game Solving via Single Policy Best Response	60	0.002	
On the intrinsic limits of pentacene field-effect transistors	75	较低	楼雯等 (2022) <sup>[33]</sup>
Classification of topological insulators and superconductors in three spatial dimensions	95	高	
Simple, fast and accurate two-diode model for photovoltaic modules	80	较高	
Enhancement of Antitumor Immunity by CTLA-4 Blockade	80	26.622	谢珍等 (2022) <sup>[34]</sup>
Restoring Function in Exhausted CD8 T Cells During Chronic Viral Infection	90	24.017	
Lymphoproliferative Disorders With Early Lethality In Mice Deficient In CTLA-4	70	16.598	
The Future of immune Checkpoint Therapy	70	15.361	
CTLA-4 Engagement Inhibits IL-2 Accumulation And Cell Cycle Progression Upon	60	9.847	
T Cell-mediated Elimination of B7.2 Transgenic B Cells	50	7.642	
Safety, Activity, and Immune Correlates of Anti-PD-1 Antibody in Cancer	80	5.956	
Blockchain's roles in meeting key supply chain management objectives	70	73.47	
Online social media fatigue and psychological wellbeing—A study of compulsive use, fear of missing out, fatigue, anxiety and depression	80	36.56	
Improving the agility of employees through enterprise social media: The mediating role of psychological conditions	60	31.19	
Social network analysis: Characteristics of online social networks after a disaster	65	25.06	魏绪秋等 (2023) <sup>[10]</sup>
Examining branding co-creation in brand communities on social media: Applying the paradigm of Stimulus-Organism-Response	70	21.74	
Big data technologies: An empirical investigation on their adoption, benefits and risks for companies	75	20.98	
The dark triad and knowledge hiding	60	20.54	
Big data analytics for disaster response and recovery through sentiment analysis	70	19.95	
National culture and consumer trust in e-commerce	65	17.36	

我现在需要你从学术论文的引证意图、引证位置、引证期刊、引用情感的权重四个方面,对论文的创新性进行测度并给出评分。你能够按我的要求去做吗

可以的,我可以从这四个方面进行测度,以评估论文的创新性并给出评分。请提供需要测度的论文和评分标准,我将尽力为您提供帮助。

图1 ChatGPT交互页面示意

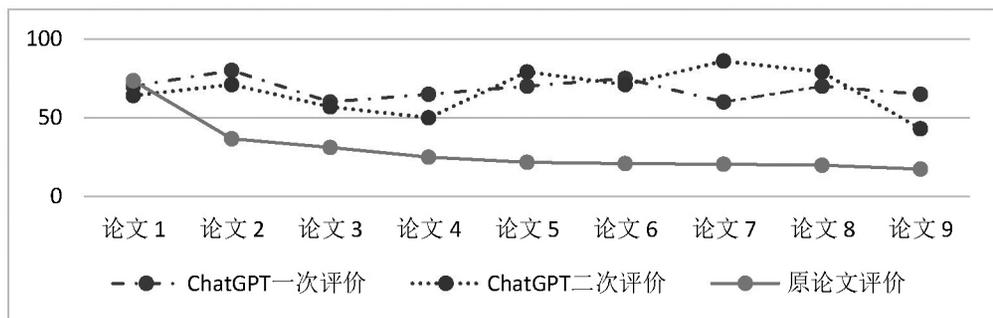


图2 ChatGPT论文创新性评价结果对比

关信息是否充分息息相关。在提问过程中向 ChatGPT 提供论文摘要、结论等重要信息,能够极大提高论文创新性评价的真实性。

此外,论文推荐任务中可以明显发现 ChatGPT 存在生成虚假论文信息问题。经过反复实验,发现 ChatGPT 面向人工智能领域学术论文进行推荐时,内容真实性最高,面向其他学科研究领域时真实性更低,推测是因为其数据库中人工智能领域相关语料最为丰富,而其他领域的预训练语料不足、质量不高。

综上所述,ChatGPT 完成论文创新性评价的效果较好,评价标准与方法较为综合、科学,评价结果具有一定的稳定性、准确性与真实性,适合用于论文创新评价任务。同时,ChatGPT 还支持用户对论文创新性评价标准提出个性化要求。但由于无法分辨预训练信息的来源与真实性,ChatGPT 提供的答案并不能被认为 100% 正确,用户应尽可能向 ChatGPT 提供更多的论文相关重要信息,从而提升评价效果。当 ChatGPT 具备互联网查询功能与相关数据库访问权限后,用于论文创新性评价的可行性将进一步提升。

## 6 结论

本研究从学术论文创新评价的真实需求和应用场景出发,通过设计完整的实验任务,分析 ChatGPT 面向生命科学、细胞免疫、市场营销、经济学、人工智能、信

息资源管理领域论文进行创新性评价方法、标准、过程与效果,结论如下:

(1) ChatGPT 在多个学科研究领域均能完成单篇或多篇论文的创新性评价。无需进行论文创新性评价训练,ChatGPT 就能将

论文与所处学科过往发展进行纵向对比分析,并从相同的维度在多篇论文间进行创新性横向比较。当 ChatGPT 能够获取论文全部信息时,评价的稳定性、准确性与真实性最高。

(2) ChatGPT 采用的论文创新性评价标准与方法基于现有论文创新性评价知识,具有较好的综合性、科学性与客观性。在评价过程中,ChatGPT 注重对论文文本内容的分析,一定程度上能降低人为主观因素影响。在面向不同研究领域时,其评价标准与方法具有明显的学科特色。

(3) ChatGPT 能够针对单篇论文给出创新建议。依托强大的学习能力和领域知识背景,ChatGPT 可依据论文关键内容,结合领域知识给出进一步研究与创新建议,能为科技创新发展提供极大帮助。

(4) ChatGPT 的论文创新性评价效果较难衡量。衡量论文创新性评价效果具有主观性与复杂性,本研究虽然通过比较过往研究的创新评价结果,部分论证了 ChatGPT 论文创新性评价的准确性与可靠性,但系统性衡量 ChatGPT 创新性评价效果的方法需进一步研究。

(5) ChatGPT 独立用于论文创新性评价目前还存在不足,具体表现为评价结果在稳定性、准确性、真实性均存在一些缺陷。但 ChatGPT 论文创新性评价的效率高、功能丰富、适用范围广,能够有效地辅助创新情报发现与分析,现阶段可将 ChatGPT 视为一种论文创新性评价辅助工具。

本研究发现,ChatGPT 在论文创新性评价任务中的创造力有限,任务完成情况与习得的知识紧密相关。信息资源管理领域是论文创新性评价知识的重要

贡献者,如果我们能够对论文创新性评价标准、方法与体系等做出更全面深入的研究,就能够为 ChatGPT 的论文创新性评价提供更加坚实丰富的知识基础。此外,今后应加强对 ChatGPT 类大语言模型提示工程的应用研究,以帮助用户更好地向模型提问,提升生成式人工智能完成任务的效果。

### 参考文献

- [1] Coupe T. Peer review versus citations - an analysis of best paper prizes[J]. *Research Policy*, 2013, 42(1): 295-301.
- [2] 李 樵,王菲菲. 认知视角下创新性研究的同行评议:偏差与干预[J]. *科学学研究*, 2021, 39(8): 1354-1363.
- [3] 杨 锋,梁 樑,苟清龙,等. 同行评议制度缺陷的根源及完善机制[J]. *科学学研究*, 2008, 26(3): 569-572.
- [4] 吴胜男,赵蓉英. Altmetrics 应用工具的发展现状及趋势之分析[J]. *图书情报知识*, 2016(1): 84-93.
- [5] 逯万辉,谭宗颖. 学术成果主题新颖性测度方法研究——基于 Doc2Vec 和 HMM 算法[J]. *数据分析与知识发现*, 2018, 2(3): 22-29.
- [6] Zhang W, Li H, Li Y. et al. Application of deep learning algorithms in geotechnical engineering: a short critical review[J]. *Artificial Intelligence Review*, 2021, 54: 5633-5673.
- [7] Uzzi B, Mukherjee S, Stringer M, et al. A typical combinations and scientific impact[J]. *Science*, 2013, 342(6157): 468-472.
- [8] 宋 歌. 科研成果创新性指标 S 指数的设计与实证[J]. *图书情报工作*, 2016, 60(5): 77-86, 124.
- [9] 魏瑞斌. 基于自引网络和主路径分析的论文主题创新实证研究[J]. *图书情报工作*, 2018, 62(3): 64-70.
- [10] 魏绪秋,姜召昊,常 霞,等. 基于引证意图的学术论文创新性评价研究[J/OL]. *情报理论与实践*: 1-10 [2023-04-26]. <http://kns.cnki.net/kcms/detail/11.1762.g3.20230406.1056.002.html>.
- [11] Yi Z, Flora S T. Chinese novelty mining[C]. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3 - Volume 3 (EMNLP '09)*. Association for Computational Linguistics, USA, 2009: 1561-1570.
- [12] 沈 律. 科技创新的一般均衡理论——关于科技成果创新度评价的科学计量学分析[J]. *科学学研究*, 2003(2): 205-209.
- [13] 沈 阳. 一种基于关键词的创新度评价方法[J]. *情报理论与实践*, 2007(1): 125-127.
- [14] 许 丹,徐 爽,陈斯斯,等. 基于自然语言词对法的文献主题新颖性探测研究[J]. *图书情报工作*, 2018, 62(8): 130-138.
- [15] 阮光册,夏 磊. 基于 Doc2Vec 的期刊论文热点选题识别[J]. *情报理论与实践*, 2019, 42(4): 107-111, 106.
- [16] 李 贺,杜杏叶. 基于知识元的学术论文内容创新性智能化评价研究[J]. *图书情报工作*, 2020, 64(1): 93-104.
- [17] 曹树金,曹茹焯. 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.
- [18] 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.
- [19] 索传军,盖双双,周志超. 认知计算——单篇学术论文评价的新视角[J]. *中国图书馆学报*, 2018, 44(1): 50-61.
- [20] 秦 岩,代 君,廖莹驰. 学术会议论文新颖性测度研究——以计算机学科人工智能领域为例[J]. *情报科学*, 2021, 39(1): 104-110.
- [21] Amabile T M. Social psychology of creativity: a consensual assessment technique[J]. *Journal of Personality and Social Psychology*, 1982, 43(5): 997-1013.
- [22] Lee Y, Walsh J P, Wang J. Creativity in scientific teams: unpacking novelty and impact[J]. *Research Policy*, 2015, 44(3): 684-697.
- [23] 索传军. 知识转移视角下的学术论文老化与创新研究[J]. *图书情报工作*, 2014, 58(5): 5-12.
- [24] 杨 京,王 芳,白如江. 基于研究水平的单篇学术论文创新力评价研究——以碳纳米管材料领域为例[J]. *情报理论与实践*, 2017, 40(9): 105-111, 76.
- [25] 袁曦临,常 娥. 网络发表学术论文的影响力评价研究[J]. *图书情报工作*, 2011, 55(10): 51-54.
- [26] 赵蓉英,魏绪秋. 中美图书情报学领域国际学术论文影响力比较——基于被引频次和使用次数[J]. *情报理论与实践*, 2018, 41(1): 139-145.
- [27] 陈玥彤,王 昊,李跃艳,等. 一种面向内容差异的学术论文评价方法[J]. *信息资源管理学报*, 2022, 12(4): 56-69.
- [28] 耿树青,杨建林. 基于引用情感的论文学术影响力评价方法研究[J]. *情报理论与实践*, 2018, 41(12): 93-98.
- [29] 宋艳辉,朱 李,邱均平. “破五唯”背景下我国科研评价体系构建的几点思考[J]. *情报杂志*, 2022, 41(2): 190-197.
- [30] 张金柱,张晓林. 利用引用科学知识突变识别突破性创新[J]. *情报学报*, 2014, 33(3): 259-266.
- [31] 石 慧,潘云涛,赵筱媛,等. 基于文献挖掘的颠覆性技术早期识别研究[J]. *中国科技资源导刊*, 2019, 51(4): 102-110.
- [32] 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.
- [33] 楼 雯,刘小曼,蔡 蓁. 基于安娜·卡列尼娜原理的单篇科学论文评价方法研究[J]. *情报理论与实践*, 2022, 45(7): 89-94, 41.
- [34] 谢 珍,马建霞,胡文静. 面向代表作评价的学术论文创新性测度方法[J]. *情报理论与实践*, 2022, 45(7): 81-88.

【作者简介】王雅琪,女,1995年生,中山大学信息管理学院博士研究生。  
曹树金,男,1962年生,山东理工大学特聘教授,中山大学信息管理学院教授、博士生导师(通讯作者)。  
收稿日期:2023-05-24