

# 认知计算——单篇学术论文评价的新视角

索传军 盖双双 周志超

**摘要** 长期以来,同行评议的主观性和低效率、文献计量学的人为操纵和不针对内容等问题始终没有得到很好的解决,学术论文评价面临不同评价方法之间难以相互弥补的矛盾,亟需新的评价理论和方法。认知计算是一个以一定的规模进行学习,带着某种目的进行推理,以此来与人类进行交互的系统。将认知计算引入学术论文评价领域,有望同时解决专家定性评价中的主观性和低效率,以及定量评价中不针对内容等缺陷。本文介绍了认知计算的相关研究和实践,探讨了学术论文评价的发展历程及其困境,揭示了学术论文评价的本质,重点论述了基于认知计算的学术论文评价新思路。研究发现,学术论文评价的发展与科学交流方式和技术变革密切相关。学术论文的评价本质是对其学术价值的评价,具体体现为创新性的测度。从学术论文语义内容角度,构建基于认知计算的学术论文评价系统是完善现有论文评价理论和方法的新视角,有望成为未来学术论文评价的重要发展方向之一。图3。参考文献43。

**关键词** 学术论文 论文评价 评价本质 评价方法 认知计算

**分类号** G250.2

## Cognitive Computing: A New Perspective for Evaluating the Individual Academic Paper

SUO Chuanjun, GAI Shuangshuang & ZHOU Zhichao

### ABSTRACT

Academic paper evaluation is a classic problem in the field of library science. Its main purpose is to help users find the excellent papers they need. The qualitative evaluation based on peer review and the quantitative evaluation based on bibliometrics are the most accepted evaluation methods of academic papers. Although qualitative evaluation is based on contents in each paper, it is inefficient and susceptible to expert subjectivity or other non-scientific factors, not suitable for the highly efficient evaluation of massive papers. Although quantitative evaluation is objective, efficient and operable, it is not directly related to the content of the paper and it is easily manipulated. All the time, although scholars at home and abroad have never stopped investigating the evaluative methods of papers, they have never achieved satisfactory results. There is contradiction that different evaluation methods could not be complemented by each other. Therefore, new theory and method for evaluation are in urgent need.

Cognitive computing is a system that interacts with human beings by learning with a certain scale and reasoning with some purpose. It has obvious advantages in understanding and processing unstructured text. Cognitive computing is mainly used to solve problems with ambiguity and uncertainty. The introduction of cognitive computing into the evaluation field of academic papers is expected to solve both subjectivity and

通信作者:索传军, Email: suocj@ruc.edu.cn, ORCID: 0000-0002-7416-1531 (Correspondence should be addressed to SUO Chuanjun, Email: suocj@ruc.edu.cn, ORCID: 0000-0002-7416-1531)

inefficiency in qualitative evaluation, as well as the lack of analyzing content in quantitative evaluation. Through the relevant document surveys, this paper discusses the research and practice of cognitive computing, analyzes the development process of academic paper evaluation, and discusses the development bottleneck that the current evaluation methods face with and evaluation essence. Mainly, this paper focuses on the new perspective based on cognitive calculation to evaluate academic papers, and we analyze the realization path of cognitive computing in the evaluation of academic papers and key issues in constructing a cognitive computing system for academic papers.

It is found that the development of academic paper evaluation is closely related to the changes in the way of scientific communication. The essence of academic paper evaluation is the evaluation of its academic value. Social value and economic value are the results of the application of academic value in different fields. The academic value and quality of an academic paper depend on its originality. From the perspective of semantic content, combing the multisource data sets and related knowledge base to construct the academic evaluation system based on the cognitive computing is expected to become one of the most important development directions of the evaluation of academic papers in the future.

The cognitive computing system proposed in this paper is a new idea for the evaluation of academic papers, and makes full use of the current advanced technology and big data thinking, which is of great value to optimize the existing evaluation theory and practice of academic papers. Firstly, the cognitive computing system of academic papers takes into account the multiple features including paper content, reference and citing papers, to realize the perfect combination of peer review and bibliometrics, and can make up its respective defects simultaneously. Secondly, the evaluation of academic papers based on cognitive computing is helpful in the fields of research management, discipline construction, reviews by editor and expert, and user's literature acquisition and reading experience et al. However, we are still in the early stage of cognitive system development, and there are several challenges to apply cognitive computing to academic paper evaluation. Among them, the datamation and semantization of the academic papers, and the machine learning model for the evaluation of academic papers are the two key issues that need to be solved in the construction of the cognitive computing system for academic papers. 3 figs. 43 refs.

## KEY WORDS

Academic paper. Paper evaluation. Evaluation essence. Evaluation method. Cognitive computing.

## 0 引言

近年来,科研水平成为综合国力的重要组成部分,国际科研竞争日趋激烈,各国对科研投入大幅增加,科研成果产出量爆发式增长。论文作为科研活动的阶段性知识形态成果,是主要的成果表现形式。据统计,过去几十年中,全球科研成果产出量正以每九年翻一倍的速度增长<sup>[1]</sup>。Science 杂志报道,平均每 20 秒就有 1 篇

学术论文发表<sup>[2]</sup>。学术论文的指数增长,一定程度上增加了用户的文献选择负担,影响了科研人员的工作效率和科研管理效率,学术论文评价因此成为全球科研管理界和学术界的热门话题。目前针对单篇学术论文评价的研究主要集中在以被引频次和期刊影响因子为核心的学术论文评价指标的构建、修正和验证,从定性和定量角度构建学术论文的评价指标体系,学术论文评价方法的探索和完善等方面。

学术论文评价的主要目的是便于用户找到

自己所需的优秀论文。基于专家同行评议的定性评价和基于文献计量学的定量评价一直是较为公认的学术论文评价方法。定性评价虽然能够针对论文内容逐篇评审,但效率低下,且易受专家主观性或其他非科学因素影响,不适宜高效的海量论文的评价。定量评价虽然较客观、效率高、可操作性强,但不直接针对论文内容,容易被人为操纵。一直以来,虽然人们从未停止对学术论文评价方法的探索,但始终没有取得令人满意的结果。无论同行评议还是文献计量学,其缺陷都是显而易见的,二者之间存在难以相互弥补的矛盾。要突破当前学术论文评价所面临的发展困境,必须转变思维,与当前的先进技术结合,寻找新的评价视角和思路。认知计算是大数据时代的发展产物,是信息分析技术和人工智能技术的结合,通过对各种类型数据的学习和理解,以一种自适应的方式解决现实世界中的复杂性和不确定性问题,对解决定性评价的主观性、低效率,以及定量评价不针对内容、指标广泛等问题具有潜在的应用价值。因而,从认知计算视角探索新的学术论文评价理论和方法显得极为重要。

## 1 认知计算的相关研究及实践

### 1.1 认知计算的产生

简单地讲,认知计算(Cognitive Computing)是一项使人类能够和机器合作的技术方法。认知计算这个术语来自于认知科学与人工智能,是借助认知科学理论构建算法,模拟人的客观认知和心理认知过程,使机器具备某种程度的“类脑”认知智能<sup>[3]</sup>。20世纪90年代后,人们开始使用认知计算一词。2013年,以IBM沃森为代表的认知计算系统实现了自主学习,并拥有了类似人脑的能力,能够按照用户需求从自然语言内容中搜寻关键知识,拉开了认知计算在各个领域应用的帷幕。

实际上,认知计算是一个多种技术的综合体,每种技术用不同的方法解决其领域内的问

题。IBM指出,认知计算是根据神经网络和深度学习来构建的,正在应用来自认知科学的知识来构建模拟人类思维过程的系统。具体来说,认知计算是指计算机拥有自我学习系统,可以像人类大脑一样学习,不仅可以处理结构化的数据,而且可以通过学习实现识别人类自然语言、图像、视频等以往都需要人类亲自操作的功能,建立一种能够摆脱人类干预并自行解决复杂问题的计算系统。认知计算最明显的特征和优势在于具备超强的自我学习能力、存储能力和计算能力。

### 1.2 认知计算的相关研究

(1)国外认知计算的相关研究。尽管认知计算还存在一些争议,但国外有关认知计算的研究已成为认知科学和人工智能领域的热点之一。目前的研究主要集中在对认知的可计算性的解释,认知系统的软硬件设施、原理及相关算法,在医疗、法律、教育等不同领域应用的探索,以及开发可以实现某种特定功能的认知计算系统等方面。2002年,美国国家基金会将认知科学看作21世纪四大前沿技术之一。2006年,IBM Almaden研究中心发起了认知计算国际会议,2007年又在加州大学召开了认知计算会议,最近的一次国际认知计算会议于2017年7月份在英国牛津大学顺利召开。现在,美国的宾夕法尼亚大学、麻省理工学院、布朗大学和乔治亚计算机学院等均建立了相关的研究所(或中心或实验室),都在开展相关研究。日本和德国也有相关的研究。

(2)国内认知计算的相关研究。虽然一些具有前瞻性的专家学者已经看到了认知计算的重要性,但国内对认知计算的研究尚处于起步阶段,主要集中在对认知计算的介绍方面,鲜有其在相关领域应用的探索。2008年,国家自然科学基金委员会发布了“视听信息的认知计算”重大研究计划,表明我国对认知计算这一国际前沿技术研究的认可,该重大计划的实施有力地推动了我国认知计算领域相关研究的发展。

2013年,北京举办以“从大数据到认知计算”为主题的认知计算研讨会,同年11月,中国自动化大会设立“生物信息与认知计算”专题会议,表明我国学者对认知计算技术方面研究的高度重视。另外,国内一些情报学学者从认知科学的角度对情报学相关理论进行探索和研究<sup>[4]</sup>。基于认知计算的专家系统、智能检索、决策支持系统、基于内容的个性化信息服务等都是情报学重要的研究领域。随着物联网的发展,有待处理的数据和信息量急剧增加,这些研究领域都需要与认知计算相结合,通过建立人机交互的认知计算系统,增强用户的服务体验。

### 1.3 认知计算的实践

认知计算是针对不同领域解决复杂问题的技术体系或方案,目前在医疗、金融和保险等领域已经开展了一些实践。在医疗领域,IBM智慧医疗的核心概念是认知计算,为用户提供个性化诊断和治疗。IBM沃森认知系统在癌症诊断和提供治疗方案方面已经取得显著成效,这主要得益于其超强的存储能力、运算能力和学习能力<sup>[5]</sup>。IBM沃森认知系统存储了300种以上医学专业期刊、250本以上的医学书籍、超过1500万页的资料和临床指南,能够在17秒内阅读3469本医学专著、248000篇论文、69种治疗方案、61540次实验数据和106000份临床报告,短期内学习并掌握医学领域的专业知识体系。

在生命科学领域,IBM沃森通过对超过7万份有关p53(一种涉及多种癌症的蛋白质)的科学论文的分析,花费不到一个月的时间,发现了6种可改善p53的蛋白质,其中前2种蛋白质已经得到实验室的验证。而同样的工作需要花费至少10位科学家14个月的时间才能完成<sup>[6]</sup>。由于认知计算在加快药物发现和再利用方面拥有巨大潜能,目前国际制药巨擘Johnson & Johnson正在进一步训练Watson阅读和理解科学论文,以评估治疗的临床试验结果。

从认知计算的相关研究和实践来看,其在

医疗、生物、金融和零售等行业都已进入初级应用阶段,教育和法律等领域也在积极开展探索性研究。尽管不同学科领域的学者都在积极从不同视角探索新的学术论文评价理论和方法,但目前尚未见认知计算用于单篇学术论文评价的研究和实践。因而,基于认知计算的学术论文评价是完善现有学术论文评价理论和实践的新视角。

## 2 学术论文评价方法的发展历程以及评价本质

### 2.1 学术论文评价方法的发展历程

#### 2.1.1 同行评议

学术论文的同行评议起源于科学共同体和科学期刊的诞生。1662年和1699年,皇家学会和巴黎皇家科学院分别成立,并分别创办了*Philosophical Transactions*和*Journal des Sçavans*两种内部期刊,逐渐取代过去通信交流实验报告和科学发现的方式<sup>[7]</sup>。1731年,爱丁堡皇家学会创办*Medical Essays and Observations*期刊,首次正式定义科学论文的同行评议流程<sup>[8-9]</sup>。尽管学术论文的同行评议已有200多年历史,但当前形式的同行评议直到19世纪才形成,标准于1967年得以确立,并被视为学术论文质量评价的金标准<sup>[10-11]</sup>。

21世纪以来,人们尝试利用自然语言处理和机器学习等技术探索学术论文中的创新点<sup>[12-15]</sup>、新知识声明<sup>[16-17]</sup>和亮点<sup>[18-19]</sup>等,试图采用技术手段识别和挖掘学术论文的价值,从而实现基于内容的学术论文评价,弥补专家同行评议的低效率问题。也有学者从知识元的角度探索新的学术论文评价视角。索传军<sup>[20]</sup>认为,一篇学术论文是n个知识元的逻辑组合,并从知识转移视角提出了基于创新知识元的学术论文评价方法。姜春林等提出,学术论文评价的基本单元是知识单元,对知识单元的把握和评价仍需借助同行专家的专业修养和思维意识,同时采用科学计量方法纠正同行评议的主

观偏差,要借助知识单元最终实现对论文作者思想和观点的认知和评价<sup>[21]</sup>。

### 2.1.2 文献计量学

20世纪50年代,引文分析的兴起催生了新的论文评价视角——文献计量学方法。1955年, Garfield 开创了从引文角度研究文献的新领域,提出被引频次评价论文影响力的思想<sup>[22]</sup>。1963年, Garfield 创办科学引文索引(SCI),促进了文献计量学在学术论文评价领域的应用和研究。1964年, Garfield 再次指出引文索引对审稿专家评价论文具有一定的辅助作用<sup>[23]</sup>。1967年, Margolis 探讨了引文索引的论文评价功能<sup>[24]</sup>。1975年,期刊引证报告(JCR)创办,影响因子被正式确立为期刊评价指标,随后被广泛用于单篇学术论文的评价<sup>[25]</sup>。

2010年, Priem 等正式提出评价论文社会影响力和短期影响力的替代计量学(Altmetrics)<sup>[26-27]</sup>。替代计量学是伴随科学交流网络化

兴起的,其实质是文献计量学与网络技术相结合的产物。替代计量学一经提出立即引起论文评价领域学者的关注,其在单篇论文评价中的应用和研究热度随之骤升,成为同行评议和传统文献计量学的有力补充。

随着信息技术的发展,基于引用内容的学术论文评价方法引起学者的强烈关注,一定程度上弥补了文献计量学不针对内容和引文不等效等缺陷<sup>[28-29]</sup>。印第安纳大学丁颖等人是基于引用内容的引文分析方法的倡导者,提出了基于引用内容的论文评价指标<sup>[30]</sup>。但实际上,人们很早就开始探索基于引用内容的论文评价,只是限于当时技术水平,只能手工分析小样本数据。如 Voos 等人早在 1976 年就对施引文献中的引用位置和引用强度进行了研究,指出由于引文在施引文献中被引用次数不同,引文是不等效的<sup>[31]</sup>。图 1 对学术论文评价方法的发展历程进行了简要描绘。

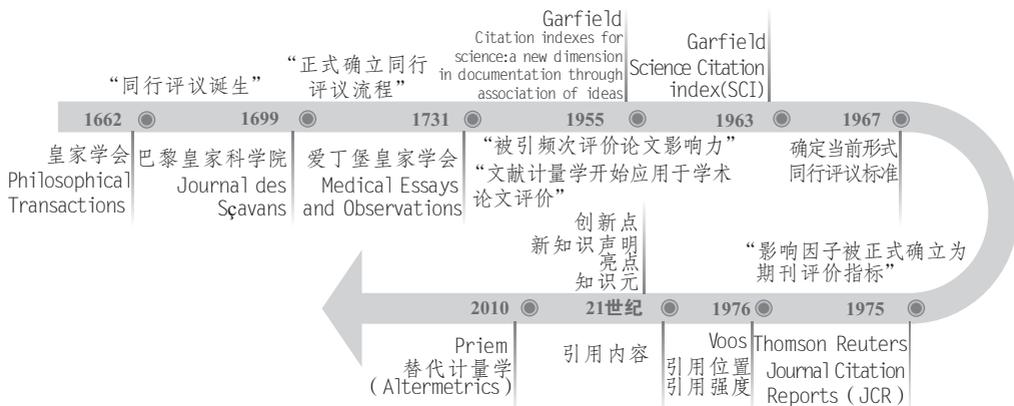


图 1 学术论文评价方法的发展历程

## 2.2 学术论文评价方法的发展困境

学术论文评价看似简单,实则为图书馆学学科领域的一个经典问题,一个受多种因素影响的复杂社会问题。既涉及学术问题,又涉及社会、人文和技术等多方面的问题。而且,学术论文类型多样、内容复杂、价值表现多样,以及实现缓慢、难以测度等问题,长期以来一直困扰着人们。

### 2.2.1 同行评议存在的主要问题

同行评议是学术论文质量的首要控制机制,是目前较为公认的论文评价方法。Ware 调查发现,93%的学者认为同行评议在学术论文评价中是必要的<sup>[32]</sup>。但由于易受个人和社会因素影响,同行评议一直备受质疑。首先,同行评议是一种主观判断,这种主观判断容易受专家学术视野、学术喜好、知识结构等个人因素以及人

情关系、竞争关系等社会因素的影响。其次,学术论文的同行评议不仅花费大量经费,还需要专家投入大量时间和精力,评价效率却往往不尽人意。最后,评价标准的缺失也使得评审过程和评价结果无据可循。虽然有学者提出论文定性评价标准,但其评价实践仍依赖于评价主体的主观判断,缺乏可操作性,而且目前尚没有哪一种标准被普遍接受或认可。然而,尽管同行评议存在诸多缺陷,但由于没有更合理的替代方法,仍是目前判断论文学术贡献的最佳方式。对科研人员来说,同行评议最重要的问题不是考虑是否抛弃它,而应该考虑如何改善它。

### 2.2.2 文献计量学存在的主要问题

虽然文献计量学仍是主流的论文评价方法,但在评价实践中一直备受诟病。首先,忽视论文内容是文献计量学评价存在的根本问题。文献计量学主要根据学术论文的外在特征对其进行间接评价,无法从内容揭示被引文献对施引文献的价值。由于容易被人为操纵,外在的文献计量学指标有时仅仅是一种表面现象。其次,文献计量学用于学术论文评价的理论基础是被引频次与论文质量之间存在相关关系,但在具体的评价实践中往往将相关关系视为因果关系。最后,忽视引文差异和个体差异是文献计量学在学术论文评价实践中的主要弊端。文献计量学指标的本质是被引频次,而简单的被引次数无法揭示引文的不同价值。而且有研究证明,期刊影响因子值与刊载的每篇论文的被引量之间并无正相关关系,期刊发表的论文被引频次往往呈非正态分布。因而,以整个期刊的平均值来衡量个体论文质量明显不够科学和准确<sup>[33-34]</sup>。

## 2.3 学术论文的评价本质

### 2.3.1 论文的评价本质是学术价值评价

学术水平、质量、价值和影响力是论文评价研究中出现最多的术语。质量和学术水平往往是评价的目标属性,价值和影响力通常是评价质量和学术水平的基本要素。然而,很多学者在构建指标时没有明确其评价功能和效应,混

淆了质量、价值和影响力之间的区别,其根本原因在于缺乏对学术论文评价本质的思考。

从学术论文的自身属性来看,一篇优秀学术论文的本质属性应该在于它的价值性。例如,一篇极具启发性的论文,即使其质量较差,如逻辑性较差、写作技巧欠佳等,但仍不能掩盖其显在或潜在的价值。相反,一篇符合规范、设计合理、逻辑严谨、缺乏创新的高质量的论文,却不能说其具有较高的价值<sup>[35]</sup>。另外,苏新宁指出,学术评价的终极目标是发现对科学发展有重要学术价值的成果<sup>[36]</sup>。因而,无论从学术论文的本质属性还是从学术评价的终极目标来看,学术论文的评价本质都在于对其价值的评价。进一步来分析,学术论文的价值由自身内容和所处环境共同决定,可以划分为社会价值、经济价值和学术价值。由于学术论文的主要功能是促进科学交流,其价值首先表现为学术价值。也就是说,学术论文的评价本质是对其学术价值的评价。社会价值和经济价值是学术论文学术价值在不同领域应用的结果。

### 2.3.2 论文的学术价值由创新性来测度

由于学科的差异,学术论文的形式、内容和价值具有多样性。也就是说,学术论文有价值和使用价值。论文被引频次是其使用价值的一种表现形式,社会价值和经济价值也体现了学术论文的使用价值,但学术论文不是普通商品,其价值不能由作者的简单劳动时间计算。一篇学术论文的学术价值是作者在长期科学研究过程中基于前人的研究成果(如参考文献就是前人研究成果的表现形式之一,参考文献的学术价值会部分转移到后续引用它的学术论文之中)产生的新思想、新方法、新观点等的具体体现,这与国家标准 GB7713—87《科学技术报告、学位论文和学术论文的编写格式》对学术论文的定义相吻合,即“某一学术课题在实验性、理论性或观测性上具有新的科学研究成果或创新见解和知识的科学记录;或是某种已知原理应用于实际中取得新进展的科学总结,用以提供学术会议上宣读、交流或讨论;或在学术刊物上发表;或作其他用途的书面文件”,并指出,学术

论文应提供新的科技信息,其内容应有所发现、有所发明、有所创造、有所前进,而不是重复、模仿、抄袭前人的工作<sup>[37]</sup>。因而,一篇学术论文的学术价值大小、质量高低,首先取决于其创新性的有无以及创新度的高低。

### 3 基于认知计算的单篇学术论文评价

#### 3.1 学术论文评价方法的探索

若干学者针对学术论文评价存在的问题开展了相应探索和尝试。为减少同行评议过程中专家的主观性和随意性,邱均平等<sup>[38-39]</sup>构建了定性评价指标体系,虽然一定程度上为专家评审过程提供了参照,但其评价实践仍依赖于评价主体的主观判断,可操作性不强。Small等<sup>[40-41]</sup>探索了基于引用内容的学术论文评价方法,深入语法和语义层面判断引文的贡献和价值,但引用动机的复杂性经常会掩盖引文的真实价值。替代计量学反映的主要是论文受关注度和流行度,而非影响力,同时存在数据源的有效性、一致性、可靠性难以验证以及不针对内容等缺陷。在信息技术飞速发展时代,面对海量的文献资源,与新技术结合是突破现有学术论文评价发展瓶颈的关键和唯一出路。我们认为,从学术论文语义内容角度,根据论文自身特征,以及学术论文的评价本质,利用与学术论文生产、传播、服务等相关的多源数据作为基础,采用认知计算技术方法构建学术论文评价的认知系统,是完善现有学术论文评价理论和方法的新思路。

#### 3.2 认知计算用于学术论文评价的科学性

认知计算方案是一个将阅读、推理和学习能力进行集成以回答问题或探索新关联的综合系统,基于类似人脑的认知和判断,试图解决生物系统中的模糊性和不确定性问题。认知计算系统对非结构化文本数据的理解和处理具有明显优势。首先,学术论文评价本身是一个复杂的不确定性问题的。其次,最理想的学术论文评价方法是,基于论文内容的同时,兼顾更多的相

关数据。因而,认知计算在学术论文评价领域具有较大的发展空间和应用潜能。此外,人类的认知是有限的,主要体现在可扩展性和偏见性两个方面。而认知系统在模仿人类思维过程的同时,增加了处理大量信息的能力和对信息进行无偏见评价的能力。因而,基于认知计算的学术论文评价有望同时解决同行评议的主观性和低效率,以及文献计量学不针对内容等问题。

#### 3.3 认知计算实现学术论文评价的路径

基于认知计算的学术论文评价需要针对学术论文自身特征,以及论文评价本质,基于语料库、知识库和相关数据集,利用自然语言处理、大数据分析和机器学习等新技术,构建学术论文认知计算系统。

具体地说,基于认知计算的学术论文评价的基本思路如下:首先,分析学术论文的特征,构建学术论文包含的各类知识元的数据模型。第二,通过自然语言处理对学术论文的文本内容进行分析和加工,建立学术论文的句子级语料库,并对语料库中的句子进行句法依存和语义依存分析,识别和抽取创新性或结论性句子。第三,依据创新知识元和结论知识元等数据模型,对学术论文的创新性句子和结论性句子进行实例化描述。基于学术论文的语义内容模型对学术论文进行语义标注,建立学术论文语料库。第四,建立学术论文作者h指数和期刊影响因子等相关数据集,以及有关专业领域的知识体系(包括专业领域体系结构、概念体系、重要知识点的知识图谱等)和知识谱系(包括专业领域知识发展的脉络,不同发展阶段的研究热点,重要研究组织与学者图谱等)等。第五,基于领域知识库、相关数据集和大量文本语料,利用机器学习技术训练认知系统,使认知系统在分析和训练中学习优秀论文的模式和特征。第六,认知系统对测试论文进行评价实践,领域专家判断评价结果的准确性和正确率,通过专家与认知系统的交互学习与反馈,不断修改和完善方案。具体路径如图2所示。

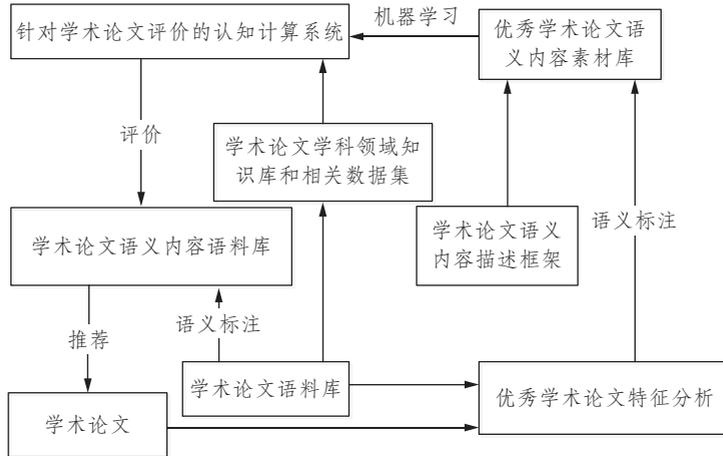


图2 基于认知计算的学术论文评价实现路径

## 4 学术论文认知计算系统构建的关键问题

### 4.1 学术论文文本的数据化和语义化

学术论文是一个结构化的、综合性的知识集合体,具有不同功能的内容组件包含了大量的知识实体,并通过组件间的关系进行组织<sup>[42-43]</sup>。换句话说,学术论文是由“内容组件”和“组件间的关系”两方面组成的。因而,我们假设,一篇学术论文是  $n$  个知识元的逻辑组合,一个知识元是  $n$  个语义三元组的逻辑组合。

简单地说,基于认知计算的学术论文评价,就是让计算机像领域专家一样,针对论文内容对其进行科学评价。因而,构建学术论文的认知系统首先要让计算机“读懂”论文(文本)、理解论文的语义。

认知系统理解学术论文,数据化和语义化的学科知识体系和学术论文文本语料是必备的,同时需要论文的内容组织框架。首先,学术论文的组织框架是针对学术论文的各类知识元(尤其是创新知识元或结论知识元),以及知识元之间的逻辑关系,构建学术论文的内容描述框架,实现对论文知识内容的组织和关联。其

次,认知计算最核心的能力是对复杂问题的“洞察力”,这个洞察力来自于对海量复杂学术论文的“深度数据分析”。深度数据分析,不仅要求待分析的数据具有大数据的4V特征,而且要求数据更加准确,且具有计算机能够理解的语义。因而,如何将海量的学术论文文本和学科知识体系加工成计算机可以理解的具有语义的形式化数据(各类知识元),是构建学术论文认知系统的关键问题之一,也是认知计算系统学习、理解、推理和判断论文创新性及其学术价值的基础。然而,目前关于知识元的定义、大小、类型、特征、表现形式以及学术论文内容的组织框架还没有形成统一的认识,尤其是学术论文中知识元之间的语义关系更是亟待探索的重要问题。

### 4.2 学术论文评价的机器学习模型

模型是机器学习中最核心的概念,针对一个问题,通常有大量机器学习模型可供选择。学术论文评价模型的质量决定了运用机器学习实现学术论文评价的成败。

机器学习是认知计算系统的关键技术之一,它所关注的问题是使用正确的特征来构建正确的模型,以完成既定的任务。其中,任务、模型和特征是其基本构成要素。任务是通过模型来完成的,是可通过机器学习解决的问题。

本研究中的任务是学术论文评价,涵盖了分类和回归两种任务。首先是分类问题。简单讲,学术论文的创新性可以分为0~1的原始性创新(如0表示无创新,1表示原始创新)和1~10的连续性创新(如应用创新)。原始性创新的判断属于二分类问题,连续性创新和质量的评价对应多分类问题,如可以将一篇学术论文划分为优秀、良好、一般和较差等若干类别,也可以根据使用价值将其分为优先推荐(优先阅读)、推荐(值得阅读)和不推荐等类别。在分类任务中,输出空间是一个由不同类别构成的集合。其次是回归问题。温有奎等指出,学术论文的创新点分布在论文的整个结构中,具体表现为主题中的创新点、技术背景中的创新点、技术方法中的创新点、论文结论中的创新点和总体创新点等<sup>[14]</sup>。同样,一篇优秀学术论文也是由创新性、规范性、逻辑性等若干内容特征以及来源期刊、作者、被引情况、参考文献等若干外部特征综合来体现的。因而,学术论文的创新性和

质量评价同时是一个涉及多元变量的回归问题,主要应用于学术论文的创新程度或质量评分过程。特征是一种用于描述问题域中相关对象的“语言”,本研究中的特征主要来源于论文的学术价值和质量的属性及其相关因素,具体体现为创新知识元的特征(学术价值体现为创新性的测度)和优秀论文的特征。学术论文评价的相关特征同时包含了属性特征、数量特征和有序特征等三种类型。模型是为解决某个既定问题而从数据中学习到的,应用于训练数据的某个机器学习算法的输出。本质上,朴素贝叶斯模型只能处理属性特征,许多几何模型只能处理数量特征。这意味着,运用机器学习实现学术论文评价需要同时结合多种模型以构建最适宜的分类器,而不能使用单一模型。因而,学术论文评价模型的构建和确定是运用机器学习实现论文评价的核心,也是构建学术论文认知计算系统的关键问题之一。运用机器学习实现学术论文评价的基本原理如图3所示。

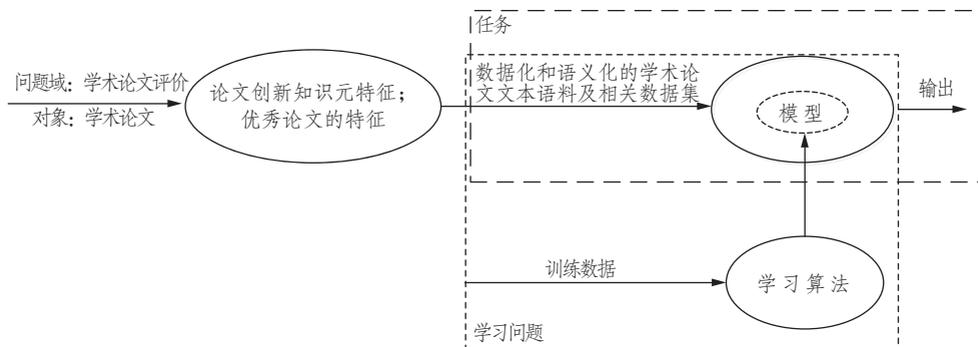


图3 机器学习实现学术论文评价的基本原理

## 5 结语

学术论文评价是一项自科学共同体和科学期刊诞生以来就一直存在并不断发展着的重要活动,同时是一个受学术、社会、人文和技术等多种因素影响的复杂问题。一直以来,同行评议的主观性和低效率以及文献计量学不针对内容等问题始终没有得到很好的解决,但由于尚

未找到更合理的替代性方法,具体的评价实践中,人们依然采用同行评议或文献计量学结果作为学术论文质量的判定标准。本文提出的学术论文认知计算系统是探索学术论文评价方法的新视角,充分利用了当前的先进技术和大数据思维,对于完善现有学术论文评价理论和方法具有重要的理论和应用价值。首先,学术论文的认知计算系统能够从论文内容、参考文献和施引文献等多个维度对论文进行评价,能够

实现同行评议和文献计量学的有效融合,同时弥补其各自缺陷。其次,基于认知计算的学术论文评价有助于服务科研管理、学科建设、编辑和审稿专家的评审工作、用户的文献获取及阅读体验等领域,有望成为未来学术论文评价的重要发展方向之一。

然而,像许多新思想、新方法和新技术一样,认知计算仍处于演进的早期阶段,构建学术论文的认知计算系统仍面临诸多困难和挑战。首先是基础设施的搭建和部署。一个认知系统由许多不同元素组成,涵盖了从硬件及部署模型到机器学习及应用的全部范围。学术论文认

知系统的构建首先需要根据自身需求,选择合适的元素和技术完成基础平台的搭建。其次,语料库的构建、维护和更新。认知计算是以数据为中心的,即语料库在学术论文认知系统中起着重要作用。利用分类法、主题词表和本体等相关知识库,以及学术论文的文本数据、反映论文质量的相关数据集等构建语料库是学术论文认知系统的重要问题之一。第三,优秀论文模式的学习和识别。如何利用机器学习等相关技术训练认知系统,使认知系统在训练和交互过程中发现和识别优秀论文的特征模式,也是认知系统实现学术论文评价的关键性问题。

## 参考文献

- [ 1 ] Bommann L, Mutz R. Growth rates of modern science: a bibliometric analysis based on the number of publications and cited references [ J ]. Journal of the Association for Information Science & Technology, 2015, 66 ( 11 ): 2215-2222.
- [ 2 ] Munroe R. The rise of open access [ J/OL ]. Science, 2013, 342 ( 6154 ): 58-59 [ 2017-04-20 ]. <https://www.sciencemag.org/content/342/6154/58.full>.
- [ 3 ] 陈敏. 认知计算导论 [ M ]. 武汉: 华中科技大学出版社, 2017: I. ( Chen Min. Introduction to cognitive computing [ M ]. Wuhan: Huazhong University of Science and Technology Press, 2017: I. )
- [ 4 ] 徐峰, 冷伏海. 认知计算及其对情报科学的影响 [ J ]. 情报杂志, 2009, 28 ( 6 ): 20-23. ( Xu Feng, Leng Fuhai. Cognitive computation and its impact on Information Science [ J ]. Journal of Intelligence, 2009, 28 ( 6 ): 20-23. )
- [ 5 ] IBM 商业价值研究院. IBM 商业价值报告: 认知计算与人工智能 [ M ]. 北京: 东方出版社, 2016: 27-28. ( IBM Institute for Business Value. IBM business value report: cognitive computing and artificial intelligence [ M ]. Beijing: Oriental Press, 2016: 27-28. )
- [ 6 ] Chen Y, Argentinis J E, Weber G. IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research [ J ]. Clinical Therapeutics, 2016, 38 ( 4 ): 688.
- [ 7 ] Lee C J, Sugimoto C R, Zhang G, et al. Bias in peer review [ J ]. Journal of the Association for Information Science & Technology, 2013, 64 ( 1 ): 2-17.
- [ 8 ] Spier R. The history of the peer-review process. [ J ]. Trends in Biotechnology, 2002, 20 ( 8 ): 357.
- [ 9 ] Kronick D A. Peer review in 18th-century scientific journalism [ J ]. Jama the Journal of the American Medical Association, 1990, 263 ( 10 ): 1321-1322.
- [ 10 ] Azam A P, Roger W. Peer review and the publication process: [ J ]. Nursing Open, 2016, 3 ( 4 ): 193-202.
- [ 11 ] Coupé T. Peer review versus citations-an analysis of best paper prizes [ J ]. Research Policy, 2010, 42 ( 1 ): 295-301.
- [ 12 ] Ronzano F, Saggion H. Knowledge extraction and modeling from scientific publications [ C ] // González-Beltrán A,

- Osborne F. Semantics, analytics, visualization. Enhancing scholarly data: second international workshop, SAVE-SD2016, Mont real, QC, Canada, 2016-04-11: 11-25.
- [13] Helen A, Purwarianti A, Widyantoro D H. Extraction and classification of rhetorical sentences of experimental technical paper based on section class[C]// International Conference on Information and Communication Technology. IEEE, 2014:419-424.
- [14] 温有奎, 吴广印. 碎片化科研创新点动态挖掘研究[J]. 数字图书馆论坛, 2014(7):25-32. (Wen Youkui, Wu Guangyin. Dynamic mining of fragmented scientific research innovation points[J]. Digital Library Forum, 2014(7):25-32.)
- [15] 毛琛瑜, 乐小虬. 领域内中文科技文献中新发现语言描述特征分析[J]. 现代图书情报技术, 2016, 32(5):47-55. (Mao Chenyu, Le Xiaoqiu. Linguistic features of new findings in Chinese scientific papers[J]. New Technology of Library and Information Service, 2016, 32(5):47-55.)
- [16] Dahl T. Contributing to the academic conversation; a study of new knowledge claims in economics and linguistics [J]. Journal of Pragmatics, 2008, 40(7):1184-1201.
- [17] Dahl T. The Linguistic Representation of rhetorical function: a study of how economists present their knowledge claims.[J]. Written Communication, 2009, 26(4):370-391.
- [18] Elsevier.Highlights[EB/OL]. [2017-09-19]. <https://www.elsevier.com/authors/journal-authors/highlights>.
- [19] Yang W. Evaluative language and interactive discourse in journal article highlights[J]. English for Specific Purposes, 2016, 42:89-103.
- [20] 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. (Suo Chuanjun. Study on obsolescence and innovation of academic papers from perspective of knowledge transfer[J]. Library and Information Service, 2014, 58(5):5-12.)
- [21] 姜春林, 张立伟, 谷丽, 等. 知识单元视角下学术论文评价研究[J]. 情报杂志, 2014(4):29-34. (Jiang Chunlin, Zhang Liwei, Gu Li, et al. Research on the evaluation of academic papers from the perspective of knowledge unit[J]. Journal of Intelligence, 2014(4):29-34.)
- [22] Garfield E. Citation indexes for science; a new dimension in documentation through association of ideas.[J]. Science, 1955, 122(3159):108.
- [23] Garfield E. "Science Citation Index"—a new dimension in indexing[J]. Science, 1964, 144(3619):649.
- [24] Margolis, J. Citation indexing and evaluation of scientific papers[J]. Science, 1967, 155(3767):1213-1219.
- [25] Campanario JM. Large increases and decreases in journal impact factors in only one year: the effect of journal self-citations[J]. Journal of the American Society for Information Science and Technology, 2011, 62(2):230-235.
- [26] Priem J, Hemminger B H. Scientometrics 2.0: new metrics of scholarly impact on the social Web[J]. First Monday, 2010, 15(7).
- [27] Priem J, Taraborelli D, Groth P, et al. Altmetrics: a manifesto[EB/OL]. [2017-09-19]. <http://altmetrics.org/manifesto>.
- [28] 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(3):583-592.
- [29] 刘盛博, 丁堃, 张春博. 基于引用内容性质的引文评价研究[J]. 情报理论与实践, 2015(3):77-81. (Liu Shengbo, Ding Kun, Zhang Chunbo. Research on the citation evaluation based on citation context nature[J]. Information Studies: Theory & Application, 2015(3):77-81.)

- [30] Ding Y, Zhang G, Chambers T, et al. Content-based citation analysis: the next generation of citation analysis[J]. Journal of the Association for Information Science & Technology, 2014, 65(9):2244-2248.
- [31] Voos H, Dagaev K S. Are all citations equal? or, did we op. cit. your idem?[J]. Journal of Academic Librarianship, 1976, 1(6):19-21.
- [32] Ware M. Peer review in scholarly journals: perspective of the scholarly community-an international study[J]. Information Services & Use, 2008, 28(2):109-112.
- [33] Seglen P O. Why the impact factor of journals should not be used for evaluating research[J]. BMJ, 1997, 314(7079):498-502.
- [34] Hamilton d P. Who's uncited now? [J]. Science, 1991, 251(4989):25.
- [35] 李睿明. 基于动机、价值、质量的科学论文评价观[J]. 科学学研究, 2008, 26(5):921-926. (Li Ruiming. The concepts on scientific paper evaluation: based on motivation, valuation and quality[J]. Studies in Science of Science, 2008, 26(5):921-926.)
- [36] 苏苏宁. 学术评价与学术考核奖励机制的辩证观[J]. 西南民族大学学报(人文社科版), 2017, 38(9):1-5. (Su Xinning. Dialectical view of academic evaluation and academic assessment incentive mechanism[J]. Journal of Southwest Minzu University(Humanities and Social Science), 2017, 38(9):1-5.)
- [37] 科学技术报告、学位论文和学术论文的编写格式[S/OL]. [2017-09-20]. <https://wenku.baidu.com/view/a9791ac12cc58bd63186bd11.html>. (The format of scientific and technical reports, dissertations and academic papers[S/OL]. [2017-09-20]. <https://wenku.baidu.com/view/a9791ac12cc58bd63186bd11.html>.)
- [38] 邱均平, 谭春辉, 任全娥. 人文社会科学评价理论与实践(上册)[M]. 武汉: 武汉大学出版社, 2012:503. (Qiu Junping, Tan Chunhui, Ren Quan'e. Evaluation theory and practice of Humanities and Social Sciences[M]. Wuhan: Wuhan University Press, 2012:503.)
- [39] 王凌峰, 杨小珍. 一种单篇研究论文直接评价新方法[J]. 情报杂志, 2016, 35(12):145-149. (Wang Lingfeng, Yang Xiaozhen. A new peer review method with multi ordinal evaluation indicator for one single research paper[J]. Journal of Intelligence, 2016, 35(12):145-149.)
- [40] Small H G. Cited Documents as concept symbols[J]. Social Studies of Science, 1978, 8(3):327-340.
- [41] 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 & Technology, 2014, 65(9):1929-1938.
- [42] 文庭孝, 刘晓英. 中文文本解构与知识发现研究[J]. 图书与情报, 2009(3):86-88. (Wen Tingxiao, Liu Xiaoying. Research on Chinese text structure analysis and knowledge discovery[J]. Library and Information, 2009(3):86-88.)
- [43] 王晓光, 宋宁远. 语义出版物的内容组织架构研究——基于纳米出版物和微型出版物的比较分析[J]. 出版科学, 2017, 25(4):20-27. (Wang Xiaoguang, Song Ningyuan. Content schema of semantic publication: a comparative analysis based on NanoPublication and MicroPublication[J]. Publishing Journal, 2017, 25(4):20-27.)

**索传军** 中国人民大学信息资源管理学院教授, 博士生导师。北京 100872。

**盖双双** 中国人民大学信息资源管理学院博士研究生。北京 100872。

**周志超** 中国人民大学信息资源管理学院博士研究生。北京 100872。

(收稿日期:2017-09-20;修回日期:2017-10-18)