



科学学与科学技术管理

Science of Science and Management of S. & T.

ISSN 1002-0241, CN 12-1117/G3

《科学学与科学技术管理》网络首发论文

题目：中国科技评价研究 40 年:历史演进及主题演化
作者：汪雪锋，张硕，刘玉琴，乔亚丽，韩晓彤，黄恒
收稿日期：2018-11-08
网络首发日期：2018-12-06
引用格式：汪雪锋，张硕，刘玉琴，乔亚丽，韩晓彤，黄恒. 中国科技评价研究 40 年:历史演进及主题演化[J/OL]. 科学学与科学技术管理.
<http://kns.cnki.net/kcms/detail/12.1117.g3.20181205.1033.004.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。



中国科技评价研究40年：历史演进及主题演化

汪雪峰¹ 张 硕¹ 刘玉琴² 乔亚丽¹ 韩晓彤¹ 黄恒¹

(1. 北京理工大学 管理与经济学院, 北京 100081;

2. 北京印刷学院 绿色印刷包装产业技术研究院, 北京 102600)

摘要:科技评价是政府进行科技管理的必要手段,是推进国家治理体系和治理能力现代化的重要工具。为厘清改革开放40年来中国学者围绕科技评价做了哪些研究工作以及研究主题的变迁过程是否与我国科技评价改革之路相伴而行,文章以CNKI数据库收录的期刊论文为样本数据,采用文献计量学和科学知识图谱方法,系统分析了科技评价领域的研究主题分布及其演化,进而挖掘其研究特征。分析表明我国在科技评价领域的研究总体上与科技评价改革方向一致,有关评价方法的研究一直贯穿始终,但不同阶段的研究主题存在明显差异,探索阶段(1979—1993年)的研究侧重于建设项目的经济性评估;规范阶段(1994—2006年)的研究热点从项目评估到绩效评价逐渐过渡,侧重于科学评价指标与评价体系的构建;而完善阶段(2007—至今)的研究侧重于绩效评价以及对各类评价方法的探索。在此基础上,结合中国科技评价改革动向就未来的研究侧重点进行了探讨,以期为中国科技评价领域的未来研究提供参考借鉴。

关键词:科技评价;历史演进;主题演化;知识图谱;文献计量学

中图分类号:F256 **文献标识码:**A **文章编号:**1002-0241(2018)00-0000-00

0 引言

科技评价是科技管理与市场机制相结合的产物,自20世纪初美国开展科技评价工作以来,科技评价已成为政府进行科技管理工作的重要手段和工具,在推动科技事业健康发展、促进资源优化配置、提高创新能力和科技管理水平等方面发挥着至关重要的作用。但近年来科技评价问题频发,成为困扰我国科技发展的关键问题之一,备受政府与科学界的关注,例如中科院院士沈文庆指出科技评价体系存在的四大问题阻碍了发展^[1],中科院院士武维华指出科技评价奖励制度存在五大问题^[2],使其与科技工作的发展不相匹配,未能充分发挥应有的积极作用。简而言之,以上各类问题相互影响,导致科学界科研不端现象频频发生,科

研人员热衷于项目申报、发表文章,却忽视了科学研究自身的价值和经济社会发展的影响,长此以往下去,科研经费得不到有效配置,不仅会对科研活动本身造成影响,还会对科学精神造成极大伤害。

与国外发达国家相比,中国的科技评价工作起步较晚、基础较差,虽在各级科技行政管理部门的不懈努力下取得了长足进步,但在评估体系、评估程序以及最终所取得的实际效果等方面还存在着较大差距。例如,教育部于2013年启动高等学校科技评价改革试点工作,与2003年《关于改进科学技术评价工作的决定》相比,面临的难题依然类似,整体内容也没有实质性突破^[3]。当前,科技评价已经被上升到推动国家创新驱动发展战略的高

收稿日期:2018-11-08

第一作者简介:汪雪峰(1977—),男,湖北荆门人,北京理工大学管理与经济学院教授,研究方向:技术创新管理、数据挖掘、科技评价。

通信作者:汪雪峰,wxfs122@bit.edu.cn

度,旨在提高科技创新的治理能力,探索更加符合创新规律的评价导向,建立符合创新规律的科技管理体制,推动科研管理向创新管理转变。同年,习近平总书记在视察中科院时指出要深化科技体制改革,坚决扫除阻碍科技创新能力提高的体制障碍,有力打通科技和经济转移转化的通道,优化科技政策供给,完善科技评价体系,营造良好创新环境^[4]。2017年科技部、财政部、人力资源社会保障部印发《中央级科研事业单位绩效评价暂行办法》,旨在深化中央级科研事业单位管理改革,建立绩效评价长效机制,发挥骨干引领作用^[5]。2018年7月4日,中共中央办公厅、国务院办公厅印发《关于深化项目评审、人才评价、机构评估改革的意见》,主要目标是形成以质量贡献绩效为导向的分类评价体系,激发科研机构 and 科研人员创新创业活力,使得科技资源配置更加高效,科技进步对经济社会发挥更大的作用^[6]。随即,国家自然科学基金委员会和教育部先后围绕人才帽子以及人才计划问题发声,足以表明中央对加快科技评价改革工作的决心以及高度重视。然而对于这些科技评价和科技评价改革问题,目前各方面的认识还存在差异,科技管理者、科学家、评价专家之间还未达成一个有效共识。因此,厘清中国科技评价工作的当前进展,探究科技评价主题变迁是否与科技评价改革之路相伴而行,找准未来研究侧重点,对于形成中国特色科技评价体系和跻身世界科技强国均具有十分重要的意义。

现有对科技评价发展阶段、改革突破口的研究偏向于采用内容分析法,且以对重要历史事件和政策文本解读的质性分析为主,无法对文本核心内容和文本内容间的关联关系进行挖掘。文章选取1979—2018年中国知网(CNKI)上科技评价领域的期刊论文,采用定性比较分析和数据挖掘相结合的方法,从多维视角对我国改革开放40年来科技评价领域的研究主题分布及演化进行探索,

并辅之以定性解读,相关研究成果不但可以对现有评价体系进行完善,还可以为科技评价领域内有关学者的未来研究提供参考借鉴。

1 中国科技评价的历史演进

1.1 数据来源

科技评价是指受托方根据委托方明确的目的,由专门的机构和人员依据大量的客观事实和数字,按照专门的规范、程序,遵循适用的原则和标准,运用科学的方法所进行的论证、评审、评议、评估、验收等活动。根据《科学技术评价办法(试行)》中的规定,科技评价的对象主要包括中央或地方财政资金资助的科学技术计划、项目、机构、人员和成果^[7]。文件中着重强调要对科学技术计划以及科学技术项目实行包括前期评价、中期评估以及后期绩效评价等在内的全程评价,并指出人才队伍建设和人才群体整体水平评价是未来的研究重点。

2010年教育部发布的《国家中长期教育改革和发展规划纲要(2010—2020年)》中对提高高等教育质量和加强质量保障与评估提出了一系列要求,并鼓励专门机构和社会中介机构对高等学校学科进行评估,形成具有中国特色的高校评价模式。2012年党的十八大报告指出要坚持科学决策、民主决策、依法决策,健全决策机制和程序,发挥思想库作用,建立决策问责和纠错制度,使得政策评估的意义再次凸显^[8]。故科技评价工作涉及计划评估、项目评估、人员评价、成果评估、机构评估、人才评估、学科评估、绩效评价、政策评估,而评价方法的研究则贯穿始终,同时考虑到我国部分学者对“评估”和“评价”并不做严格区分,为保证数据的查准率和查全率,在制定检索式时均涉及“评估”和“评价”。

为此,文章以中国知网(CNKI)收录的期刊论文数据作为数据源,按照TI=(科技评价 or 科技评估 or 评价方法 or 评估方法 or 人才评价 or 人才评估

or 成果评价 or 成果评估 or 项目评价 or 项目评估 or 绩效评估 or 绩效评价 or 学科评估 or 学科评价 or 政策评价 or 政策评估 or 人员评价 or 人员评估 or 计划评价 or 计划评估 or 机构评价 or 机构评估) 进行检索, 时间跨度为 1979 年 1 月 1 日—2018 年 6 月 30 日。为适当控制数据量, 同时兼顾论文质量, 选择 SCI 期刊、EI 期刊、核心期刊以及 CSSCI 期刊等高标准期刊作为检索数据源; 考虑到核心期刊和 CSSCI 检索期刊分别从 1992 年和 1998 年才开始收录论文, 故将检索时间划分为 1979—1991 年、1992—2018 年两个时间段, 第一阶段中从所有数据源中共检索到 1 586 条论文数据; 第二阶段则从 SCI 源期刊、EI 源期刊、核心期刊以及 CSSCI 源期刊中检索到 24 313 条论文数据, 最终获得样本数据 25 899 条。论文数量的年度变化趋势如图 1 所示 (由于检索时间截止到 2018 年 6 月 30 日, 因此 2018 年数据暂不完整)。

1.2 科技评价的历史演进

科技评价的改革一直伴随着科技事业的发展, 我国有关学者分析总结了我国科技评价工作在不同发展阶段的特性。代涛、李晓轩等^[9]将科技评价的发展历程划分为三个阶段: (1) 民国时期 (1912—1948 年), 该时期政府对科技投入较少, 对

科技的影响和干预也少, 科技管理以科学共同体内部治理为主; (2) 新中国成立到改革开放前 (1949—1978 年), 该时期科技管理学习苏联模式, 建立了高度集中型的科技管理体制, 政府在科技决策、经费配置、人员聘用等方面均处于主导地位, 科技评价发挥的作用非常有限; (3) 改革开放后 (1979—), 我国将科学技术作为第一生产力, 科技在国家经济社会发展中的地位不断得到提升, 成果转化、信息化工作推动、创新文化建设、体制改革等科技管理活动往往通过评价评比的方式进行, 科技评价的重要性日益突出, 成为我国科技管理的重要手段。郑晓娟^[10]则将我国科技评价工作的发展历程划分为四个阶段: (1) 1979—1993 年, 此时我国科技事业得到了前所未有的发展, 科技发展领域百花齐放、生机勃勃, 但政府各部门制定的各类计划间并没有有机的结合起来, 科技评价问题逐渐显现; (2) 1994—1997 年, 第三方评估机构出现, 整个科技评价体系趋向规范化和专业化; (3) 1998—2007 年, 科技评价体系建设成效显著, 同时国家通过出台法律条文, 明确了科技评价的法律地位; (4) 2011—2015 年, 中共中央、国务院出台的有关文件旨在推动科技评价工作朝着更符合创新规律的发展方向进行改革。

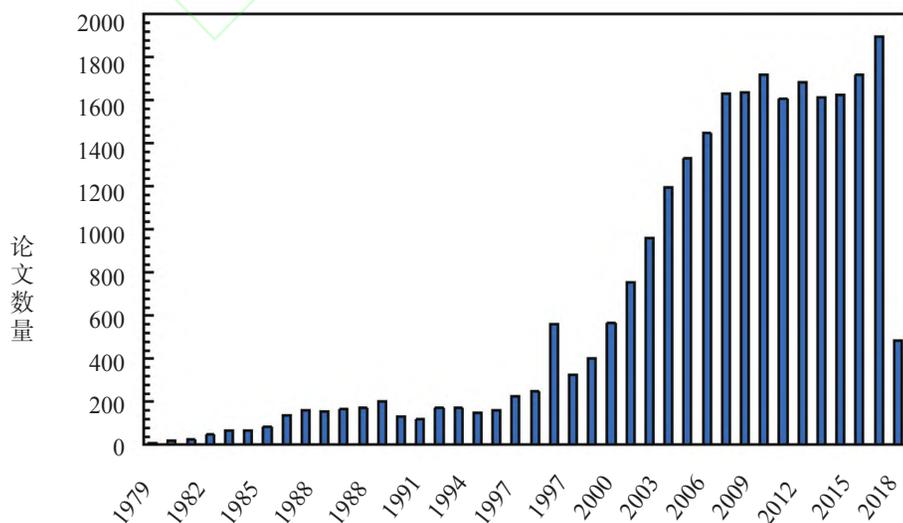


图1 1979—2018年科技评价领域论文数量年度变化趋势图

基于上述研究工作,重点探讨改革开放40年来我国科技评价研究工作的历史演进。这期间伴随着一些纲领性政策的出台:1982年我国设立中国科学院科学基金,1986年成立国家自然科学基金委员会,标志着由专家民主决策科学基金分配的同行评议机制正式引入我国。1994年,原国家科委率先提出用“第三只眼睛”对政府资助的科技计划、项目等开展独立评估。1998年,中科院启动实施知识创新工程,同行评议方法被正式应用到研究所的评价当中。2001年,我国科技评估活动中第一个行为规范和技术规范—《科技评估规范》出版,标志着我国科技评估活动步入专业化、规范化阶段。2004年,科技部等5部委联合发布《科学技术评价办法》,进一步明确了评价的目的、原则、分类方法、评价准则和监督机制等,针对科技活动中计划、项目、机构、人员和成果如何评价等相关问题进行了系统性的规范^[1]。而在2007年,国家修订通过《中华人民共和国科学技术进步法》,明确了科技评价的法律地位,指出国家应该建立和完善有利于自主创新的科技评价制度。基于上述研究成果以及有关政策,将科技评价在1979—2018年的发展历程划分为以下三个阶段:探索阶段、规范阶段、完善阶段。

1.2.1 探索阶段(1979—1993年)

改革开放后,我国将科学技术作为第一生产力,科技在国家经济社会发展中的地位不断提升。与市场经济改革相一致,科技管理中逐步破除“计划”方式,引入市场机制。1985年科技体制改革以后,科研事业单位事业费的资助被大幅度削减,取而代之的是各类竞争性项目资助逐渐成为科技投入的主要方式,项目的立项、结题、验收,科技成果评奖,科研机构、人员的绩效评价等也成为科技活动中必不可少的环节。在这一阶段,政府推出的各类资助计划(包括人才计划)虽然看起来百花齐放、生机勃勃,但各类计划间并没有有机

结合起来,以致于造成一种无序竞争状态。此外,对于政府管什么、评什么、科学共同体自主性的定义等,没有进行梳理,科技评价的问题开始显现。而此阶段中,美国、日本等发达国家的科技评价工作已经逐步步入正轨,远远领先我国,例如美国早在20世纪60年代就已经是评价领域的佼佼者,是最早以立法形式规定科技评价工作的国家^[2]。日本的科技评价工作最早也可追溯到20世纪40年代,到60年代初已经建立了比较完善的科技审议会制度,并于1986年编制了《研究评价指南》,规定了研究评估的标准。

探索阶段我国科技评价工作的研究主题主要包括以下几方面:

(1) 项目的经济性评价,关键词涵盖“科学管理、经济效益、国民经济、长远计划、综合经济”等。

(2) 建设项目的经济性评价:一是从企业的经济利益角度出发开展财务评价,关键词涵盖“内部收益率、利润、收益率、资本成本、贴现率、折现率、现金流、净现值、财务管理、投资回收期、基准收益率”等;二是从整个国家和社会的角度出发开展国民经济评价,关键词涵盖:“影子价格、会计价格、社会折现率、投资决策分析、通货膨胀率”等。

(3) 评价标准与方法的探索,关键词涵盖“综合评价、指标体系、评价体系、经济指标、指标权重、评价准则、评价标准、综合评价法、层次分析法、判断矩阵”等。

(4) 国外先进经验的借鉴,关键词涵盖“美国、日本、科学性、苏联”等。在早期探索阶段,美、日等发达国家从评价方法、评价体系以及评估机构等方面来说均已逐渐显现科学性雏形,通过对国外评价工作的梳理与借鉴,可以对我国的评价工作起到很好的推动作用。图2中“苏联”节点的出现,也可以印证这一点,因为我国在新中国成立后的一段时间内主要学习苏联科技管理模式

式,即建立高度集中型的管理体制,此时侧重于“项目评估”,“成果评估”、“绩效评价”、“政策评估”、“人才评估”等关键词出现频率较小。具体情况如图 2 所示。

1.2.2 规范阶段(1994—2006年)

此阶段政府努力摆脱“既是运动员又是裁判员”的角色,提出用“第三只眼睛”对科技计划进行评估,评价类型也从单一的科技成果鉴定转变为多元的公共研发项目管理,以彰显科技评价的公平、公正。同时,科技评价体系的组织管理运作从组建机构、建立人才队伍、制订规章制度、定出标准等环节开始,试图从原来的直接管理转变为间接管理。此时的变化集中在两个方面:一是体现科技评价规范化、专业化的要求,实行分类管理。分类管理一直是科技评价改革的核心之一,历年来

的相关政策大都强调应该注重基础研究、应用研究、科技产业化等不同类型的科学技术活动的自身特点,避免简单化、“一刀切”。二是加强评价的组织管理,促进评价活动的公平、公正、公开。对此相关部门也进行了积极探索,主要措施包括:引入第三方评估规避利益冲突、通过视频答辩等改进评估的组织、加强回避等相关制度的落实来加强评审管理、开展相关公示、专家自律等。整体而言,科技评价既是专业性很强、技术含量很高的研究活动,又是制度化、规范化的工作行为,仅仅凭借科技部颁布的《科技评估管理暂行办法》及《科技评估规范》对其进行约束不利于科技评估活动的健康有序发展。为此,需要借鉴西方发达国家的做法和经验,通过完善立法,使得科技评价工作较快的实现制度化、系统化和科学化的发展。

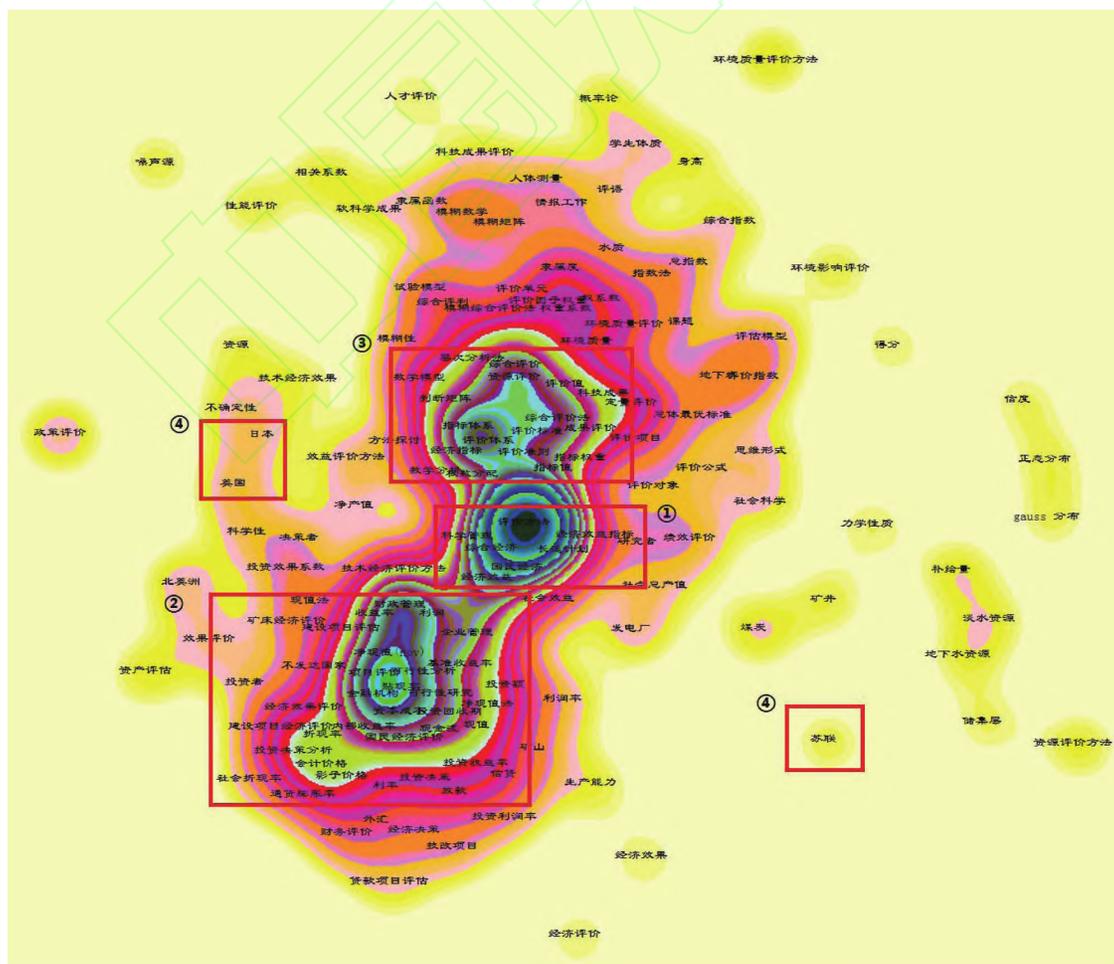


图2 探索阶段科技评价研究关键词分布

绩;二是评价结果与科技人员的工资、职称等直接挂钩过多,导致科研活动中容易产生急功近利思想,甚至成为科研不端行为和学术腐败的诱因;三是评价指标盲目套用,简单量化重科技人员发表论文的数量和刊物级别,而忽视了科研成果的质量和成果转化对社会的贡献^[13]。“十二五”期间,国家把“自主创新、重点跨越、支撑发展、引领未来”作为科技工作的战略方针,把自主创新能力建设提升到国家发展战略的核心位置,科技评价的改革也被提到落实国家创新驱动发展战略的高度,对战略的实施起着至关重要的。在此期间,主要任务表现为两个跨越的完成:一是从论文量化导向(包括以论文为代表的其他项目、经费、专利、人才、奖励等量化导向)到学术质量导向的跨越;二是从单纯的学术质量导向到学术贡献和影响力

导向的跨越^[14]。

完善阶段我国科技评价工作的研究主题主要包括以下几个方面:

(1) 以绩效为导向的评价体系构建,关键词涵盖“指标体系、评价体系、评价方法、绩效评价、绩效评价体系”等。

(2) 政府绩效研究,关键词涵盖“政府绩效评价、政府绩效、地方政府、电子政务、公共服务、有效性”等。

(3) 经济、环境绩效评价,关键词涵盖“环境绩效、财务绩效、低碳经济、环境绩效评估、财务绩效评价、企业绩效”等,而创新绩效、协同创新、技术创新作为国家新兴战略方针下的研究点也逐渐向针对不同评价对象的评价研究聚焦。较为明显的是,“项目评估”工作的研究热度开始大幅度骤

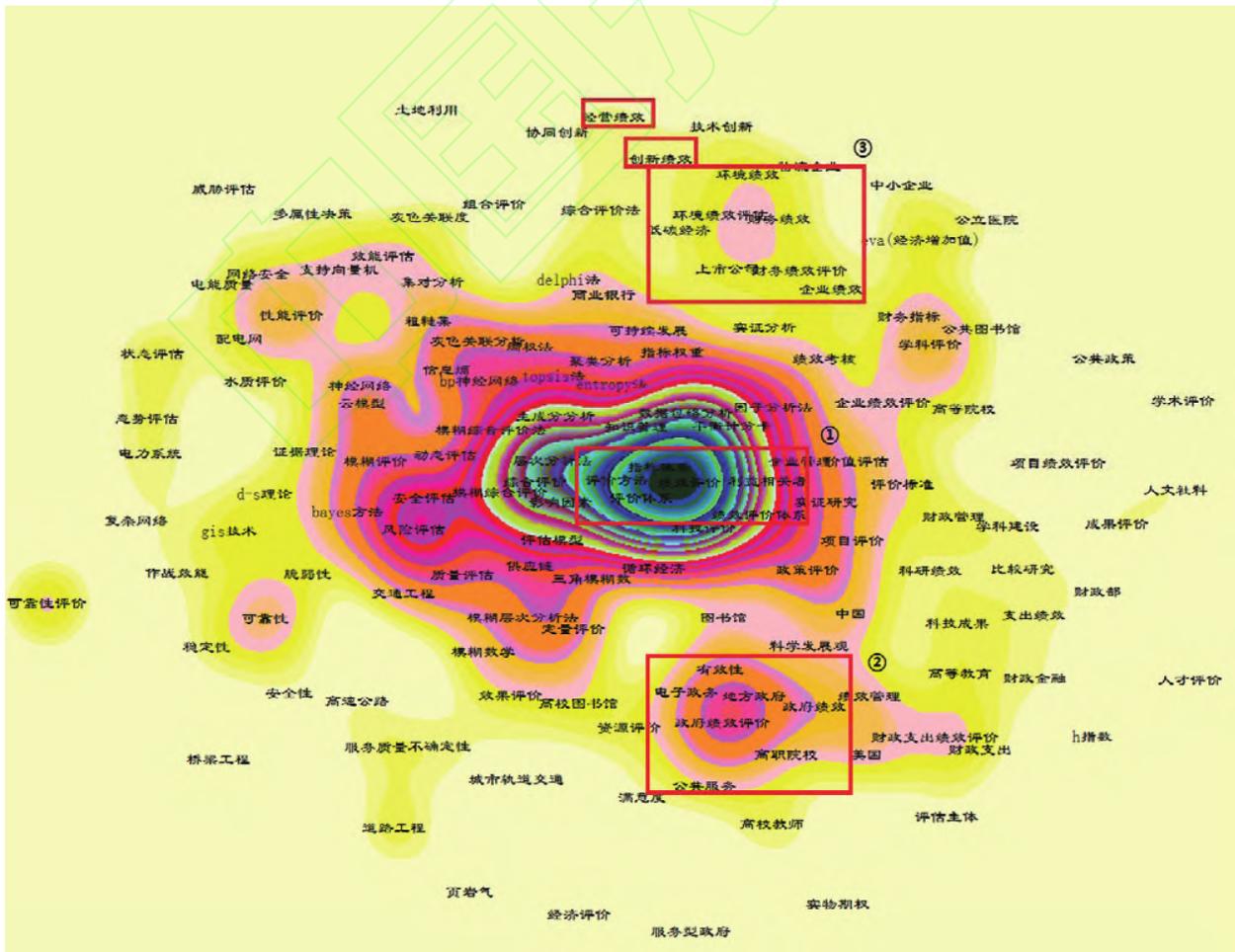


图4 完善阶段科技评价研究关键词分布

降,而“科技评价”、“政策评估”、“学科评估”等小众化研究主题的热度逐渐增大。“成果评估”、“人才评估”等检索词出现的频次较低。具体情况如图4所示:

综上所述,“评价方法”是每个阶段的研究热点,第一阶段的研究主要以“项目评估”为主;第二阶段主要侧重于科学化、多元化指标体系与评价体系的构建,“科技评价”、“学科评估”的研究热度有所增加,同时该阶段是一个承转阶段,面临着研究热点从“项目评估”到“绩效评价”的过渡;第三阶段中“绩效评价”的研究热度持续高涨排在首位,同时“项目评估”的研究趋于消亡,此阶段“科技评价”、“政策评估”、“学科评估”等小众化研究主题的热度也逐渐增大。

2 中国科技评价研究的主题演化

主题演化反映研究主题随时间的变化,包括主题内容或强度随时间的推移发生变化。对于科研人员来说跟踪主题的后续发展是一个非常重要的话题。如何有效地组织这些大规模文档数据,并且按照时间片来获取文档中研究主题的演化特征,非常具有实际意义。具体而言,主题演化的原理是使用固定时间窗口方法对检索到的论文数据进行划分,在每个时间段内以论文中的关键词作为学科主题表征,分别构建共词网络,接着对该网络进行处理,通过社区发现算法找出每个时段上的网络社区,并为每个社区赋予主题标识。利用相关性算法,探测前后时段中网络社区间的相似性以此确定社区间的演化关系,即整个研究主题的演化过程^[15]。本节应用文本分析与可视化软件ITGInsight^[16],筛选出每一阶段词频排在前十五位的关键词,绘制科技评价全阶段的研究主题演化图。该图可以展示主题产生、增强、减弱、聚合、裂变、消亡的过程,连线的粗细代表了主题间关系的强弱,不同颜色则代表不同研究主题^[17]。

从图5中可以看出,“评价方法”研究一直贯穿

始终,其中“项目评估”主要集中于第一、二阶段,从第三阶段起则趋于消亡;“绩效评价”在第一阶段开始出现,但研究热度很低,在第二、三阶段则一路攀升,最终排在首位。具体而言,第一阶段的研究主题比较分散,从第二阶段起研究主题开始聚焦,集中于指标体系、评价体系的构建,研究方法上以综合评价法和层次分析法较为突出,由此也体现出该阶段的两个特点,一是偏向于多元化指标的构建,二是偏向于分层交错且难以量化的指标的构建,该阶段对科学化、合理化、量化指标体系以及评价体系给予了足够重视。从第三阶段起,“绩效评价”凸显,学界对于科学化、合理化指标体系及评价体系的完善工作再次兴起,代表性事件是2014年国务院印发《关于改进加强中央财政科研项目和资金管理的若干意见》,重申对推进科技评价和奖励制度改革的相关要求,并从总体目标和基本原则两个方面对改进加强科研项目和资金管理提出要求和具体措施。该文件是建国以来国务院首次针对科研项目和资金这一具体工作提出指导意见,从而说明当前的科技评价体系仍然无法满足国家需求^[18]。此外,在第三阶段中,学界侧重于对各类评价方法进行探索来满足上述工作需求,例如平衡计分卡、数据包络分析是以绩效评价为导向,而综合评价、因子分析法、层次分析法、模糊综合评价、主成分分析则侧重于对多元化、分层化指标的研究,适用于各项分类评价。

为进一步探索不同阶段的研究主题演化情况,筛选出每一阶段词频排在前二十位的关键词,绘制不同阶段下科技评价的研究主题演化图,以此挖掘科技评价的阶段特征。由于评价方法研究一直贯穿始终,因此在具体分析时,将其剔除,以更好地显示不同阶段的差异。

2.1 探索阶段的主题演化

探索阶段最显著的主题是“项目评估”,在1983—1992年期间其研究热度都非常靠前。从定

义上看,项目评估是在可行性研究的基础上根据国家有关部门颁布的政策、法律条文,从项目、国民经济和社会的角度出发,由有关部门对拟建项目的必要性、建设条件、财务效益、经济效益、社会效益等进行全面分析与论证的过程。在我国项目评估的发展主要有三个关键性的时间节点,一是1981年成立中国投资银行,以转贷世行贷款为主要业务;二是1983年中国投资银行推出《工业贷款项目评估手册》;三是1987年国家计委和建设部发布《建设项目经济评价方法》。从图6中可知,1982—1990年期间,有关经济效益、财政管理的研究一直较为靠前,这从侧面说明了此阶段以项目的经济性评价为主要考量点,而作为项目评估主体的金融机构在该期间内也多次出现,探索阶段的主题演化图(图6)与主题聚类图相互印证(图2),关键词同样是围绕“经济”与“管理”进行展

开。除此以外,考虑到项目的经济效益对社会的影响,经济指标被越来越多地纳入评价体系中,以1989年和1991年两年最为明显。由于该阶段对于科学共同体以及评价工作如何展开并没有一个完整的定义,因此1988年后关于评价指标与评价标准的研究热度也呈上升趋势,具体体现为研究主题“指标体系”和“评价体系”出现频次明显增加。与此同时,为配合这种全方面、多层次的评价方式,层次分析法和综合评价法也成为该阶段主流的应用方法。整体上来看,探索阶段早期侧重于自然资源(土壤、水资源、煤炭等)的评价,中期则由于项目评估概念被引入我国并逐渐兴起,便以经济效益、财政管理以及建设项目经济评价为重点,而后侧重于整个体系框架的构建,为寻找一套行之有效的规范方法不断进行探索。具体情况如图6所示。

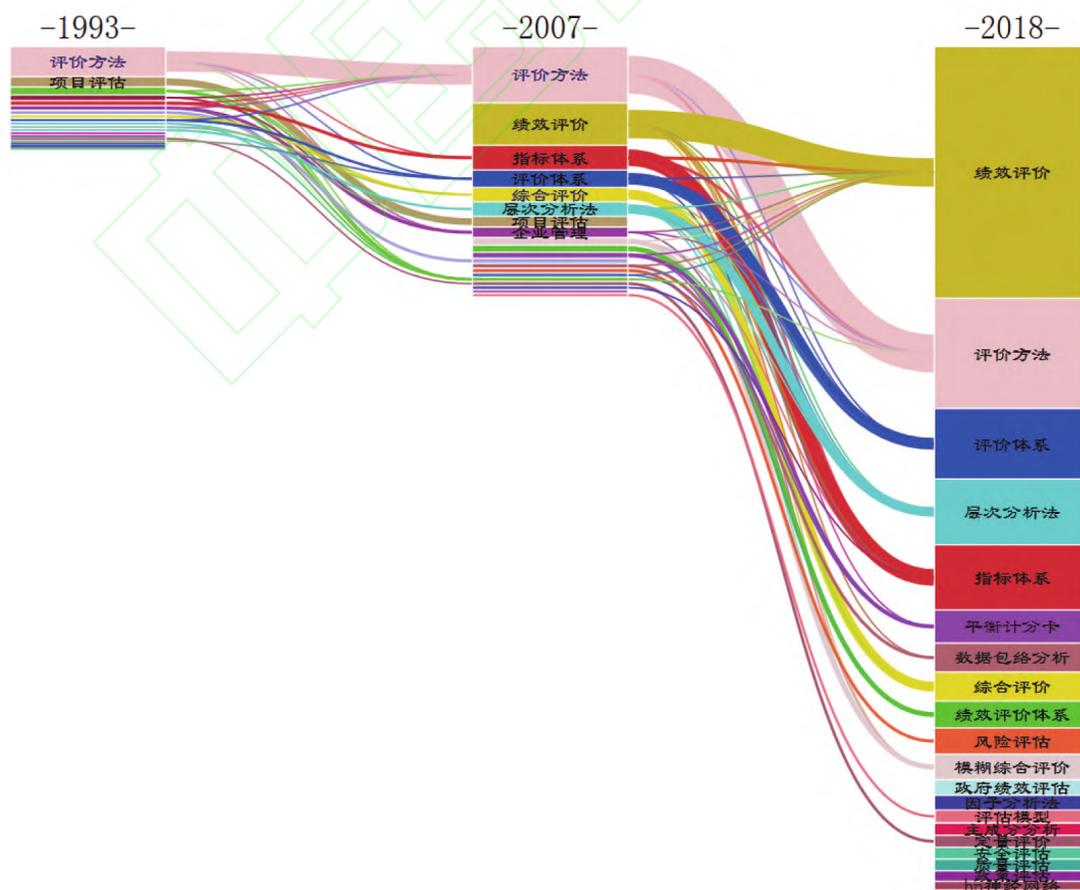


图5 科技评价研究主题演化

2.2 规范阶段的主题演化

规范阶段的研究主题“项目评估”在1994—1999年还保持着一定的研究热度,随后急剧下降并趋于消亡。与之相反,“绩效评价”的研究热度于2000年开始总体呈上升趋势,2003年起便排在首位,这可能是由于科技部于2000年发布的《科技评估暂行管理办法》所导致的,文件在规定科技评估工作的对象和范围时着重强调了对科技计划的执行情况与绩效的评价。相较于探索阶段而言,本阶段研究主题“指标体系”、“评价体系”的出现频次呈大幅上升趋势,这应该与有关政策的出台实施密不可分。2003年科技部出台了《科学技术评价办法(试行)》,明确了评价的目的、原则、分类方法、评价准则和监督机制,针对科技活动中计划、项目、机构、人员和成果如何评价的问题进行了系统性的规范^[6]。而后,国务院于2005年又在

《国家中长期科学和技术发展规划和纲要》中指出要进一步细分评价对象(科技计划、机构、人员),区别不同层次(国家、地方、部分),划分科技活动类型(基础研究、应用研究、科技产业化),依据自身特点,确立不同评价指标、内容和标准。因此,在具体构建评价指标时,“层次分析法”和“综合评价法”总是与指标体系共现,这也从侧面说明了本阶段的指标体系正朝着多元化、层次化的方向发展。一系列行政规章制度的出台,为中国科技评价工作实现规范化奠定了坚实基础,对推动中国科技评价体系的建立、评价工作的展开起到了十分重要的作用。具体情况如图7所示。

2.3 完善阶段的主题演化

科技评价活动是由国家主导的旨在解决基础科学、公共性技术的科学研究性活动,近年来中国的研发经费投入总量不断上升,先后超过英国、法

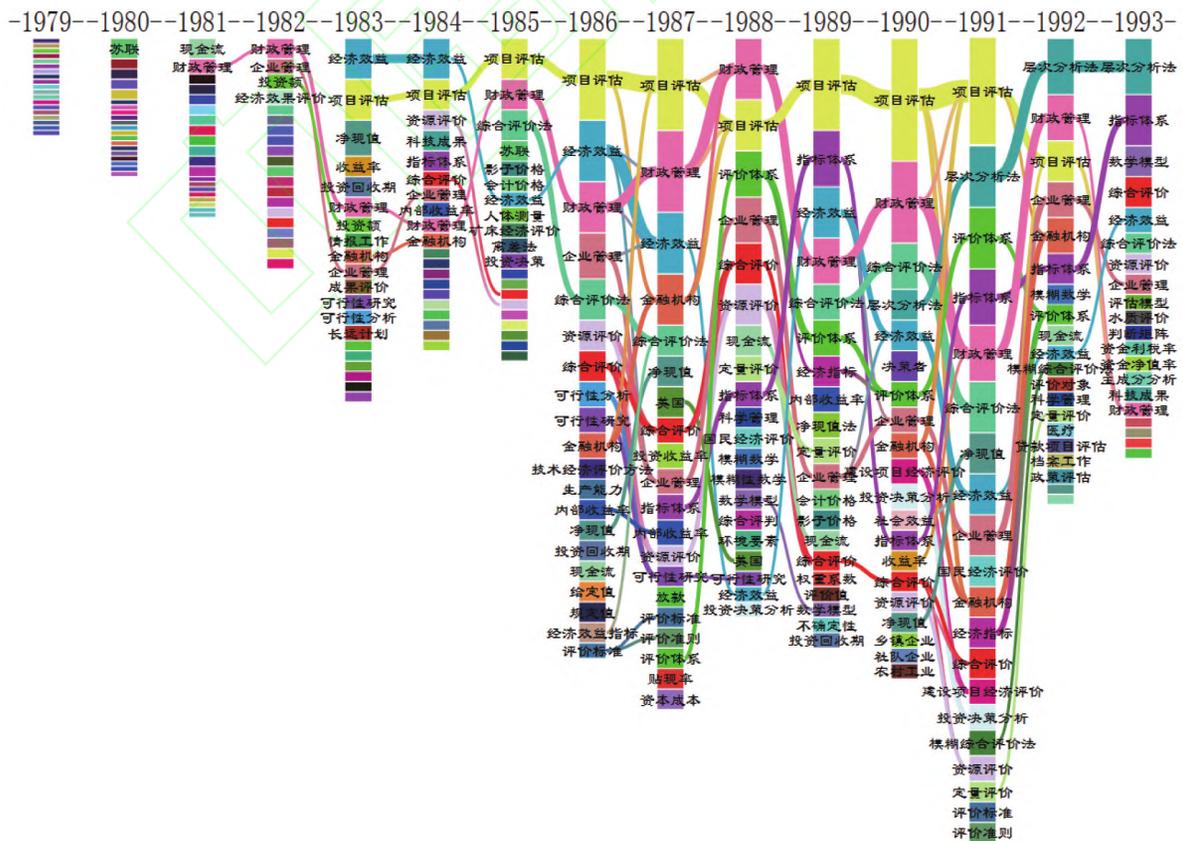


图6 科技评价研究主题演化(探索阶段)

国、德国和日本,成为仅次于美国的世界第二大科技经费投入国^[18],因此以资源配置以及使用效果为导向的绩效评价体系已经成为落实国家创新驱动发展战略的迫切需求。与前两个阶段相比,完善阶段每年的研究热点词基本保持不变,“绩效评价”牢牢占据第一位,且多与“指标体系”、“平衡计分卡”及“数据包络分析”共现,这表明绩效评价已经逐步被纳入到指标体系的构建中,而平衡计分卡与数据包络分析成为绩效评价领域的主流应用方法。此外,科技部于2009年颁布的《科技成果评价十个工作方案》和2011年颁布的《国家中长期科技人才发展规划(2010—2020年)》均旨在推动建设以重大产出成果为导向的分类评价体系,“综合评价”、“因子分析法”、“层次分析法”、“模糊综合评价”、“主成分分析”与“绩效评价”的共现也从侧面说明科学界对于分类评价体系构建的重视程

度。具体情况如图8所示。

3 中国科技评价领域的未来研究展望

总体而言,国内有关科技评价的研究已经取得较为丰硕的成果,总体上与科技评价的改革之路相伴而行,但必须认识到此类研究还存在着诸多不足,随着《关于深化项目评审、人才评价、机构评估改革的意见》(简称“三评”文件)文件的颁布,科技评价的新一轮改革工作再次被提上日程。“三评文件”旨在推进科技体制改革,提升我国科技创新能力,以激发科研人员的创造性为核心,以构建科学、规范、高效、诚信的科技评价体系为目标,统筹自然科学和哲学社会科学等不同学科门类,推进分类评价制度建设,发挥评价指挥棒和风向标作用,加快建设创新型国家和世界科技强国。结合中央发布的最新指导文件以及前述分析,我们认为研究工作的绩效评价还是未来的主旋律,与此

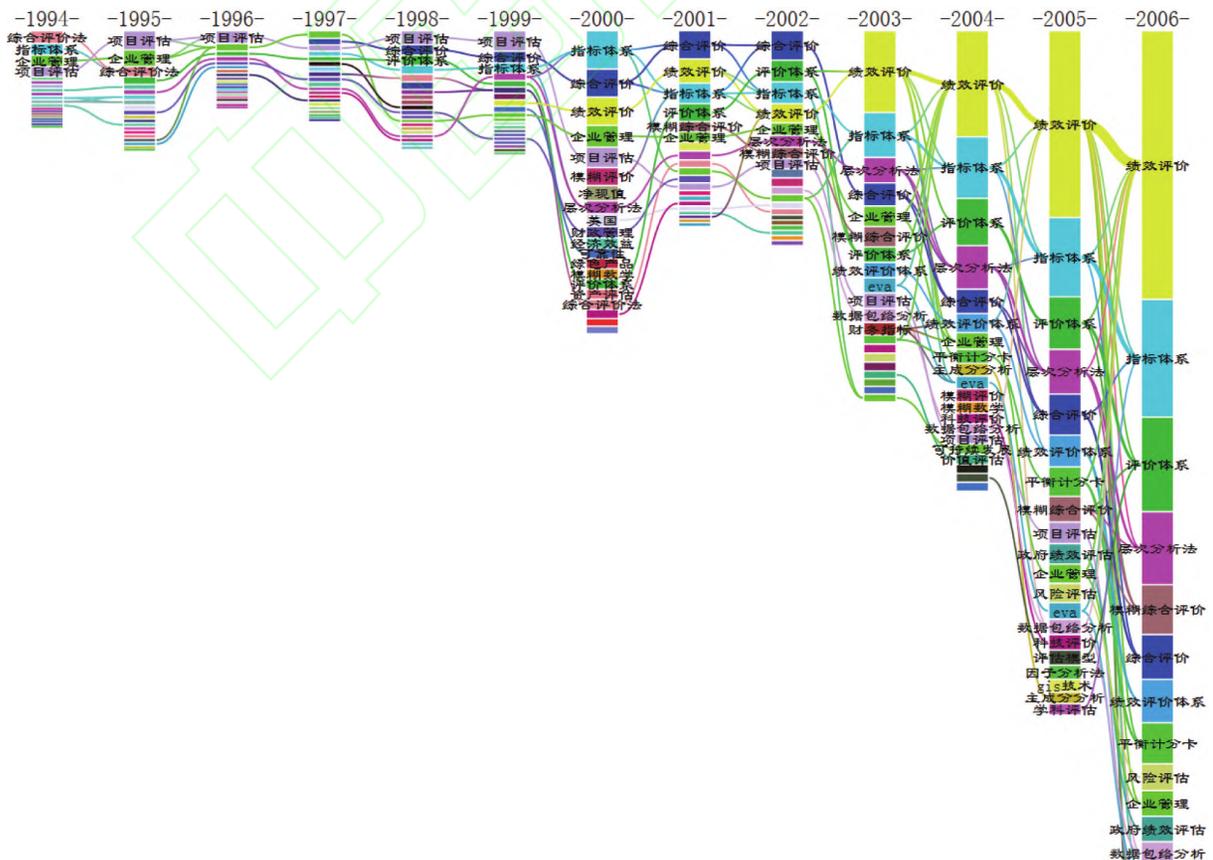


图7 科技评价研究主题演化(规范阶段)

重考察其实际效用,中国科学院一直在发展和完善的研究所评价体系^[23]就是在将评价重点转向对相关研究成果对经济社会起到贡献或作用上,例如针对基础类研究应该重点考察其原创性成果,应用类研究应该重点考察其成果转移转化效率和最终取得的实际应用效果,对于软科学类研究则应该重点考察其科研活动满足决策支撑需求的情况等,从第三阶段的主题演化图中数据包络分析、平衡记分卡所占比的上升,也能突出社会各界对于科研活动实际绩效的重视程度。

(3)完善评议专家库、细化评价制度。关于科技评价改革推进的研究中大多提到科技界过度定量的现象,但从对现有科技管理办法内容来看,大多数评价管理办法都没有对量化做明确要求,仍是以定性或是定性定量相结合的方法为主。这主要是因为在一些评价活动中,尽管请了行业内的领军专家,但他们在专业细分领域上也或多或少会存在差异,因此某种程度上并不能完全做到从研究贡献、研究水平等方面去判断一项研究工作的真正价值,还是要依靠定量的数据,比如论文引文数据(如h指标^[24],g指标^[25],R和AR指标^[26])、影响因子、专利申请数等等,在各种因素错综复杂的影响下,便会造成科研人员急功近利的思想愈演愈烈。要想解决此类问题,除了上面提到的分类评价,下放给科研人员更多的自主权以外,还应加强建设评审专家的遴选制度,建立同行评议库,细化整个评价流程、指标,给予评议专家一定的参照标准。从第三阶段主题演化图中主成分分析、综合评价、模糊综合评价、层次分析法所占比的上升,也能突出学界对于科学化评审制度的重视程度。

参考文献

- [1] 中央政协新闻网. 沈文庆:我国科技评价体系存在四大问题阻碍发展. [2011-6-23]. <http://cppcc.people.com.cn/GB/34961/225119/225122/14979169.html>.
- [2] 中国网. 中科院院士:科技评价奖励制度存在五大问题. [2012-3-10]. http://www.china.com.cn/news/2012lianghui/2012-03/10/content_24861088.htm.
- [3] 温才妃. 把脉科研评价体系改革. [2015-2-16]. <http://sci-tech.people.com.cn/n/2015/0226/c1007-26598556.html>.
- [4] 中央政府门户网站. 习近平:深化科技体制改革增强科技创新活力. [2013-7-17]. http://www.gov.cn/ldhd/2013-07/17/content_2449985.htm.
- [5] 科技部. 中央级科研事业单位绩效评价暂行办法. [2017-11-09]. http://www.gov.cn/xinwen/2017-11/09/content_5238210.htm.
- [6] 中国政府网. 关于深化项目评审、人才评价、机构评估改革的意见. [2018-7-4]. http://www.gov.cn/xinwen/2018-07/04/content_5303293.htm.
- [7] 科技部. 科学技术评价办法(试行). [2003-9-20]. <http://www.synl.ac.cn/org/non/nonequ/culture/way.pdf>.
- [8] 胡锦涛. 坚定不移沿着中国特色社会主义道路前进 为全面建成小康社会而奋斗:在中国共产党第十八次全国代表大会上的报告[R]. 北京:人民出版社,2012.
- [9] 代涛,李晓轩. 我国科技评价的问题分析与改革思路[J]. 中国科学院院刊,2013,28(6):750-755.
- [10] 郑晓娟. 论创新驱动发展战略下的科技评价体系改革[J]. 科技管理研究,2016,36(4):46-49.
- [11] 方衍,田德录. 中国特色科技评价体系建设研究[J]. 中国科技论坛,2010(7):12-15.
- [12] Brass C T. Changes to the government performance and results act (GPRA): Overview of the new framework of products and processes[J]. Congressional Research Service, 2012:1-33.
- [13] 杜鹏,李凤. 是自上而下的管理还是科学共同体的自治:对我国科技评价问题的重新审视[J]. 科学学研究,2016,34(5):641-646.
- [14] 徐芳,李晓轩. 跨越科技评价的“马拉河”[J]. 中国科学院院刊,2017,32(8):879-886.
- [15] 王晓光,程齐凯. 基于NEViewer的学科主题演化可视化分析[J]. 情报学报,2013,32(9):900-911.
- [16] 刘玉琴,汪雪峰,雷孝平. 科研关系构建与可视化系统设计与实现[J]. 图书情报工作,2015,59(8):103-110.

- [17] 刘玉琴, 逢金辉, 崔志成, 等. 一种简易的技术主题图绘制方法[J]. 图书情报工作, 2017, 61(13): 125-132.
- [18] 陈兆莹. 关于科技评价和科技评价改革的讨论[J]. 科学通报, 2018(7): 611-617.
- [19] 乔纳森·科尔, 斯蒂芬·科尔. 科学界的社会分层[M]. 中国香港: 华夏出版社, 1989.
- [20] Molasgallart J, Davies A. Toward theory-led evaluation: The experience of European science, technology, and innovation policies[J]. American Journal of Evaluation, 2006, 27(1): 64-82.
- [21] Corley K G, Gioia D A. Building theory about theory building: What constitutes a theoretical contribution?[J]. Academy of Management Review, 2011, 36(1): 12-32.
- [22] Xu F, Li X X, Meng W, et al. Ranking academic impact of world national research institutes: By the Chinese Academy of Sciences[J]. Research Evaluation, 2013, 22(5): 337-350.
- [23] Xu F, Li X. The changing role of metrics in research institute evaluations undertaken by the Chinese Academy of Sciences (CAS)[J]. Palgrave Communications, 2016, 2(1): 1-6.
- [24] Cronin B, Meho L. Using the h - index to rank influential information scientists[J]. Journal of the Association for Information Science & Technology, 2014, 57(9): 1275-1278.
- [25] Egghe L. Theory and practise of the g-index[J]. Scientometrics, 2006, 69(1): 131-152.
- [26] Jin B H, Liang I M, Rousseau R. et al. The R- and AR-indices: Complementing the h-index[J]. Chinese Science Bulletin, 2007, 52(6): 855-863.

Forty Years of Research on Science and Technology Evaluation in China: Historical and Theme evolution

WANG Xuefeng¹, ZHANG Shuo¹, LIU Yuqin², QIAO Yali¹, HAN Xiaotong¹, HUANG Heng¹

(1. School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China; 2. Academic of Printing and Packaging Industrial Technology, Beijing Institute of Graphic Communication, Beijing 102600, China)

Abstract: Science and technology evaluation has become the necessary means of modern science and technology management for government. It is also an important tool to promote the modernization of national governance system and governance capacity. In order to clarify what kind of research works have been done by Chinese scholars around science and technology evaluation, and whether the change process of research theme consistent with science and technology evaluation reform in China, the paper systematically analyzes the theme distribution and evolution of science and technology evaluation by using the bibliometrics and knowledge graph methods, taking the journal papers on science and technology evaluation area included in the CNKI as samples. The paper tries to exploit the research characteristic in this field. The results show that it is generally consistent with the direction of scientific and technological evaluation reform, and research on evaluation methods has always existed. But there are obvious differences at different stages. The exploration stage (1979-1993) focuses on construction project evaluation; The standard stage (1994-2006) focuses on the construction of scientific evaluation index and evaluation system. At the same time, this stage faces the change of research focus from project evaluation to performance evaluation; The perfection stage (2007-present) mainly focuses on performance evaluation and various evaluation methods. On this basis, this paper combines the trend of science and technology evaluation reform in China.

Keywords: S&T evaluation; historical evolution; theme evolution; knowledge graph; bibliometrics