

# 科技创新演化循证分析理论与方法研究

冷伏海 王建芳

(中国科学院国家科学图书馆 北京 100080)

**摘要** 数字科研环境的发展为推动科技创新各方面影响因素的提取提供了便利,使得支持科研决策的情报分析可以引入多方面的证据进行综合分析。本文尝试引入循证决策的方法,探索从科学数据资源和各类型文献资源中提取面向科技决策的证据基础,进而构建支持科研决策的科技创新演化循证分析理论与方法。

**关键词** 科技创新演化 循证决策 科学计量  
中图分类号 G353.1

## Evidence-based analysis theory and methodology for of science and technology evolution

Leng Fuhai Wang Jianfang,

(Library of Chinese Academy of Sciences Beijing 100080)

**Abstract** The development of digital research environment makes it possible to mine various of impact factors for science and technology innovation, to introduce multiple evidence to carry out information analysis, furthermore to support the decision making of science and technology developing programming. This paper tried to transplant method of evidence-based decision making, and explore making good use of data from both science data and literature as evidence for science and technology decision making. Then designed evidence-based analysis theories and methodology for innovation evolution to help science decision making.

**Keywords** science and technology evolution  
evidence-based decision making scientometrics

## 1 研究背景

科学技术发展到今天,其发展变化不再是科学技术系统自身的演化,而是越来越多地包含了社会经济等因素的制约和选择,是整个创新系统的演化。决策者未来科技发展战略的选择和规划作为科技管理的重要课题,需要从社会、经济等需求出发,科技发展也需要解决人类社会面临的主要挑战,并提高人们的生活质量,这就需要基于现实信息、数据和知识的综合科技情报分析来支持科技战略决策。

现有的支持科学技术发展战略选择的情报分析与研究主要限于科学或技术本身演化规律的揭示。在分析方法和思路上,主要是对科学领域中研究热点的转移研究为主,包括从研究前沿<sup>[1]</sup>、快速发展前沿、快速突破文献、热门文献和新兴研究前沿<sup>[2,3]</sup>等角度的演化分析。

以上分析视角所利用的方法,一般都是基于引文和文献特征的,如高被引或共引、共词,以及新兴研究前沿分析中主题提取和展现的文本挖掘方法和简单的统计分析方法。具体分析的方法是通过引用关系或概念词之间的关系构建知识网络,进而分析知识网络的演化特征来揭示科学技术的演化。基于引文的科学演化分析主要依据文献之间的引用关系建立科学知识关联,构建各层面科学知识的网络,来分析科学演化的特征<sup>[4]</sup>。这里的关系包括引文网络、

共引网络和耦合网络<sup>[5]</sup>。基于概念词关系的科学演化分析，主要是共词分析方法。共词分析原理主要是对具体某一组词分别两两统计他们在同一篇文献中出现的次数，在此基础上对这些词进行分层聚类，揭示出这些词之间的亲疏关系，进而分析这些词所代表的学科和主题结构演变<sup>[6, 7]</sup>。比较共引分析的科学结构揭示重点，共词分析更适合于微观层面学科或主题领域知识的演化分析。但共词分析存在标引者的影响、不同词描述同一现象、冗余度等问题，虽然共词图谱可以提供一个看得见的科学领域的结构和动态状况，但其利用率却远远低于共引分析。

定量解析科学技术演化规律经历了近半个世纪的方法探索和规律总结，已经取得了很大的进步，但计量方法的有效性仍然得不到认同，对决策支持的作用不明显。如有学者所说的，“无论我们怎样用统计手段来操纵处理引文和共词数据，科学知识图谱都不是科学领域的活动性和进展的适当表现。”因此，现有的科学演化揭示存在着理论支持不足、方法有效性不足、揭示的层次不够深入等问题，科学演化揭示能成为真正有用的科学研究的辅助方法，在理论与实践中还存在着一些值得深入思考和研究的问题。

数字科研环境的发展一方面推动了科研的快速进步，另一方面也扩大了支持情报分析的信息、数据来源，为推动科技创新的各方面影响因素的提取提供了便利，使得综合多方面的数据形成科研决策的证据，进而支撑面向需求的科技发展战略选择成为可能，可以引入医学领域广泛使用的循证决策方法进行科技情报分析，以揭示科技演化规律和未来发展趋向。

Davies 定义循证决策为在政策制定和执行过程中利用从外部可获得的最佳研究证据而帮助人们做出有关政策、项目和计划知情决策的方法<sup>[8]</sup>。循证决策是卫生决策者最常用的、客观的、也是最重要的一种卫生政策研究方法。其基于证据的决策思想也在其他领域得到了重视，并有人尝试用于学科发展热点的揭示，但仅限于基于文本分析的领域关键主题词所揭示的内容分析。<sup>[9, 10]</sup>

为此，从影响未来科技创新演化的互动因素出发，在现有科技演化计量分析方法的基础上，引入循证决策的方法，探索从科学数据资源和各类型文献资源中提取面向科技决策的证据基础，进而构建支持科研决策的创新演化循证分析理论与方法，具有重要的意义。

## 2 科技创新演化循证分析的理论与方法问题

科研及科研决策环境的改变使得科学技术演化的规律性、范式在改变，人们解读这种发展变化的理论和方法也需要逐步更新。科技演化的过程借用循证的思维进行分析，涉及循证决策分析方法在此的适用性，分析中作为证据的分析数据对象的选择、证据选择标准的建立，以及相应的分析方法的构建等方面的问题。

### 2.1 循证决策分析在科技创新演化分析中的适用性

循证决策分析中的关键是证据的获取，核心是证据的评价，目的是为科学决策提供依据，这也是科研决策中的必备要素。循证决策的过程是在问题提出的基础上，进行证据的收集并经过严格的评价，进而对证据进行综合的分析，然后通过不断的决策者评估、方案调整等，得到满意的结果。

目前科研及科研决策环境的改变，使得科技政策规划转向面向整体科研环境的科研事业规划，科学技术的选择更具战略性。现有的趋势或演化分析通常不能综合考虑影响未来科研发展的情景因素，而实际决策支持需要综合不同的信息和方法形成综合分析模型。循证决策

中的证据理论及综合分析思想恰好迎合了复杂科研环境下决策支持的要求,特别是科研决策中对于准确、可靠、全面证据的需求,需要引入循证决策中证据评价的思维和方法来提升赖以决策的信息的质量,而综合分析也是做出决策的有力依据。另一方面,目前数字科研环境的成熟,为科技演化描述中融入各层面的数字对象,进而形成支持科研决策的循证情报分析方法奠定了基础。因此,循证分析方法的各种要素与科技演化分析的内容和目的是一致的,虽然不同的决策环境中具体使用的方法不尽相同,但其循证的思维过程是必然适用的,数据对象的深度处理,也为深入的循证统计分析等提供了可能。

## 2.2 科技创新演化分析中的证据提取及选择标准

有别于以往完全依赖科学文献的视角,尝试从影响未来科技创新演化的互动因素出发,从社会经济需求、产品和工艺创新、技术开发和科学研究的创新价值链上获取演化分析的证据基础,并在现有科技演化文献计量分析方法的基础上,分析对象引入经济社会研究、政策文献和科学数据,从科学数据资源和各类型文献资源中提取面向科技决策的证据基础,这样可以大大拓展科技演化分析中的证据可利用资源范围,进而构建支持科研决策的创新演化循证分析的理论与方法。

影响循证决策效果的三个要素是:研究证据,可利用的资源以及政策的价值取向。因此,循证分析效用发挥的前提在于对研究课题进行系统全面的证据准备,确定数据纳入和剔除的标准,并对纳入的数据进行严格评价,在此基础上对结果进行定性定量分析才能得出可靠的决策依据。

作为支持科研决策的科技创新演化分析,需要明确社会经济需求、政策环境、产品创新、技术开发和科学研究各端点的现状和未来发展状况,及其相互作用关系,因此,作为循证分析证据的数据和资源的选择标准的制定是至关重要的,如包括:在统一学科领域定义下,各端点内涵外延的限定标准;在领域限定标准下可供分析的数据源的选择标准;数据或文献纳入和剔除标准,以及资料数据规范标准等,这些标准的建立将提供有力的研究证据和可用资源,为以下的指标和方法设计奠定基础。

## 2.3 证据指标的设计与分析方法的构建

在证据准备和选择完成后,最重要的是依据证据进行分析,包括从定性和定量两个方面,分析科技创新演化的规律性和创新价值链中各端点之间的关系,形成科研决策的判据。

### 2.3.1 创新价值链各端点的演化情况分析指标、方法设计

这里包含了社会经济需求、政策环境、产品或工艺创新、技术开发以及科学研究各端点知识演化的现状或未来关注趋向分析的指标和方法,如各端点的综合发展状态指标、基于领域知识特征的属性和内容发展特征指标的提取和状态分析。

这些指标可以从科学数据或科学文献、社会经济研究文献中提取领域共同的主题内容,分析属性划分后的主题词的统计或关联特征,基于语义的知识网络揭示领域知识演化的特征。其中主题词统计或关联特征是对领域发展各阶段所体现的定量特征的描述,此外,还可以计算领域在创新价值链各层面的知识熵、概念确定性、智力程度等方面的指标,找出知识演化质的发展阶段的评价指标,作为科学演化定量分析的依据。基于语义的知识网络分析从知识基因理论中知识的构成原理、变化特征的阐释出发,对表征知识的概念词和类属借用知识本体、分类体系等为语义不明确的概念词进行属性划分,在此基础上构建基于语义的知识网络,进而利用定量的指标、方法揭示微观的领域知识结构演化、领域关键研究问题的演化、

关键问题中知识关联的发展变化，提取定性定量的分析指标和内容。

### 2.3.2 系统关联机制的分析方法

通过创新链中各端点的关联分析，揭示科技创新中知识演进的互动关系。需要在各端点知识演化特征分析的基础上，利用特定的关联构建机制，分析创新系统视角下，多端点知识构成的创新网络的演化特征和各端点的相互作用关系。方法如依据核心主题词在不同数据源的共现和间接关系建立关联，形成创新网络，从创新网络、研究主题内容的核心属性标志的演化关系，来分析系统内部知识发展的演化过程。

### 2.3.3 基于判据指标的演化关系系统分析

系统分析，就是借鉴循证分析中系统综述的方法，即针对某一具体问题系统全面地收集相关资料，用统一的科学评价标准，筛选出符合标准、质量好的证据，用统计学方法进行综合，得到定量的结果，以得出可靠的结论。这里依据上述方法得到的多因素指标，作为循证决策的定量判据，利用系统分析的方法，判断创新发展的状况，挖掘以往科技创新发展中的空白点，尝试提出新的科学技术研究方向。

## 3 需要解决的关键问题

上述过程和方法的实现，需要解决一系列的理论和和技术问题，这里的关键问题，如制定分析数据源的选择标准与证据评价标准、用于循证分析的文献和科学数据指标的设计、知识网络的构建与分析，以及其中涉及的数据特征的提取等。

基于知识本体的数据特征提取是将多种数据用于系统分析的基础，特别是对各种文献作为证据对象进行深入分析，首先要提取适用的数据特征。这里需要利用知识本体(ontology)所规范的领域知识属性划分，对科学文献或数据的计量特征词进行自动的语义属性划分，或者对知识类别进行属性划分，进而利用基于语义的共现技术构造知识网络，形成循证分析的指标或论证基础。

用于循证分析的文献和科学数据指标的设计也是关键问题。要利用循证分析的方法揭示复杂科技创新系统中的知识演化特征，需要定量分析领域演化的现有水平，对领域不同端点的研究进行定量综合，揭示某方面研究的不确定性，提出新的演化方向，这里的指标设计决定了分析结果的成败。另一方面，系统分析的技术对结论的可靠性同样至关重要，要对多来源数据结果进行定量合并进行统计分析，涉及检验纳入分析的异质性、根据异质性检验结果选择模型进行统计分析，效应合并值的假设检验和统计推断，以及结果敏感性分析等问题，都需要进一步的验证和尝试。

此外，不同数据源和不同类型数据的融合，以及在此基础上创新网络的构建和分析也是本研究的一个挑战。对各层面知识网络及综合创新网络的分析，拟借用社会网络分析中的特征点和结构分析方法，进行特征的提取和演化关系揭示，进而形成定量分析指标。这里需要构建创新价值链不同端点的知识网络，以及整个创新系统构成的复杂创新网络，进而通过相似度的测度和演化特征的定量分析，如矩阵的向量距离测度、多维尺度分析、矩阵的相关与回归分析，以及包括关联性分析、中心性分析、凝聚子群分析、结构对等性分析、核心—边缘结构分析等在内的社会网络分析技术，进一步形成循证分析的指标和论证依据。

## 4.总结与展望

在以上思路的指导下，本研究期望在如下方面有所突破：

首先，支持科研决策的科技创新演化理论的拓展。支持科研决策的演化理论以往仅以科学或技术系统的演化理论总结为出发点，本课题适应新时期科研环境的变化、科学决策思维的变化，引入基于创新价值链的科学创新系统演化理论，作为演化分析的理论基础，可以更好地满足未来科研决策的需求。

其次，引入循证的情报研究方法。有别于以往仅从文献特征总结科学技术演化规律的研究，引入科学数据等分析对象，从循证的思维构建创新系统中科技创新演化发展状况的判据和分析方法，综合多种资源作为证据基础进行循证情报分析，是对情报研究方法的拓展。

再次，通过循证分析中指标、方法和流程的设计，在理论研究和案例实证的基础上，形成支持科技战略决策的循证情报分析的方法和流程，希望引导科技战略管理思想的新发展。

因此，以上理论和方法的完善，希望有助于在科学的宏观和微观环境变化条件下，探索演化机制揭示的新理论和方法，科技创新演化的新规律，为进一步深入研究科学演化的理论和方法提供思路，并全面提升科技情报研究辅助科学研究和决策的质量。

### 参考文献：

- [1] Chen C, Lin X, Zhu W. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature[J]. *Journal of the American Society for Information Science and Technology*. 2006, 57(3): 359-377.
- [2] Alkemade, F. M.P. Hekkert, Kleinschmidt, C., (2007), Analyzing emerging innovation systems: A functions approach to foresight, *Int. J. Foresight and Innovation Policy*, 2006,3(2):139 - 168.
- [3] Tugrul U. Daim. Forecasting emerging technologies: Use of bibliometrics and patent analysis. *Technological Forecasting & Social Change*, 2006 (73): 981–1012
- [4] Small H. Tracking and predicting growth areas in science[J]. *Scientometrics*. 2006, 68(3): 595-610.
- [5] Morris S A, Yen G, Wu Z, et al. Time Line Visualization of Research Fronts[J]. *Journal of the American society for information science and technology*. 2003, 54(5): 413-422.
- [6] Callon M, Courtial J P, Laville F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry[J]. *Scientometrics*. 1991, 22(1): 155-205.
- [7] Leydesdorff L. Why words and co-words cannot map the development of the sciences[J]. *Journal of the American Society for Information Science*, 1997, 48(5): 418-27.
- [8] Davies, P.T. ‘What is Evidence-Based Education?’, *British Journal of Educational Studies*, 1999, 47(2):108-121.
- [9] Wibowo, H. S. Critical themes in electronic commerce research: A meta-analysis. *Journal of Information Technology*, 2006, 20(1):1-19
- [10] Chen, C. and Yu, Y. Empirical studies of information visualization: A meta-analysis, *International Journal of Human-Computer Studies*, 2000(53): 851-866
- [11] Ronald N. Kostoff. Literature-related discovery (LRD): Methodology. *Technological Forecasting & Social Change*, 2008(75):186–202
- [12] Chen, C.. Searching for intellectual turning points: Progressive knowledge domain

- visualization[R]. Proceedings of the National Academy of Sciences, 101(Suppl. 1), 2004: 5303-5310
- [13] Leydesdorff L. 科学计量学的挑战——科学交流的发展、测度和自组织[M]. 北京: 科学技术文献出版社, 2003.
- [14] Cahlik T, Jirina M. Law of cumulative advantages in the evolution of scientific fields[J]. Scientometrics. 2006, 66(3): 441-449.
- [15] Jarneving B. A comparison of two bibliometric methods for mapping of the research front. Scientometrics, 2005, 65(2):245-263

作者简介: 冷伏海, 男, 1963 年生, 教授, 中国科学院国家科学图书馆博士生导师、情报研究部副主任。通讯地址: 北京市海淀区中关村北四环西路 33 号, 中国科学院国家科学图书馆情报研究部, 邮编 100190。电子信箱: lengfh@mail.las.ac.cn  
王建芳, 女, 1979 年生, 中国科学院国家科学图书馆助理研究员。通讯地址同上。电子信箱: wangjf@mail.las.ac.cn。