

 中国社会科学院创新工程学术出版资助项目

 全国博士后管理委员会资助项目

科学知识网络理论与实践

Scientific Knowledge Network Theory and Practice

吕鹏辉 著



科学知识网络理论与实践

Scientific Knowledge Network Theory and Practice

吕鹏辉 著



经济管理出版社

ECONOMY & MANAGEMENT PUBLISHING HOUSE

图书在版编目 (CIP) 数据

科学知识网络理论与实践/吕鹏辉著. —北京: 经济管理出版社, 2016. 12
ISBN 978 - 7 - 5096 - 4680 - 9

I. ①科… II. ①吕… III. ①科学知识—数据处理—研究 IV. ①G302

中国版本图书馆 CIP 数据核字(2016)第 262165 号

组稿编辑: 宋 娜

责任编辑: 魏晨红

责任印制: 黄章平

责任校对: 张 青

出版发行: 经济管理出版社

(北京市海淀区北蜂窝 8 号中雅大厦 A 座 11 层 100038)

网 址: www.E-mp.com.cn

电 话: (010) 51915602

印 刷: 三河市延风印装有限公司

经 销: 新华书店

开 本: 720mm × 1000mm/16

印 张: 16.75

字 数: 274 千字

版 次: 2016 年 12 月第 1 版 2016 年 12 月第 1 次印刷

书 号: ISBN 978 - 7 - 5096 - 4680 - 9

定 价: 88.00 元

· 版权所有 翻印必究 ·

凡购本社图书, 如有印装错误, 由本社读者服务部负责调换。

联系地址: 北京阜外月坛北小街 2 号

电话: (010) 68022974 邮编: 100836

第五批《中国社会科学博士后文库》 编委会及编辑部成员名单

(一) 编委会

主任：王京清

副主任：马援 张冠梓 俞家栋 夏文峰

秘书长：张国春 邱春雷 刘连军

成员（按姓氏笔画排序）：

卜宪群	方勇	王巍	王利明	王国刚
王建朗	邓纯东	史丹	刘伟	刘丹青
孙壮志	朱光磊	吴白乙	吴振武	张翼
张车伟	张世贤	张宇燕	张伯里	张星星
张顺洪	李平	李林	李永全	李向阳
李国强	杨光	杨忠	陆建德	陈众议
陈泽宪	陈春声	卓新平	房宁	罗卫东
郑秉文	赵天晓	赵剑英	高洪	高培勇
曹卫东	曹宏举	黄平	朝戈金	谢地坤
谢红星	谢寿光	谢维和	裴长洪	潘家华
冀祥德	魏后凯			

(二) 编辑部（按姓氏笔画排序）：

主任：张国春（兼）

副主任：刘丹华 曲建君 李晓琳 陈颖 薛万里

成员：王芳 王琪 刘杰 孙大伟 宋娜

苑淑娅 姚冬梅 郝丽 梅枚 章瑾

本书获中国博士后科学基金面上项目（项目编号：2015M581984）
以及国家自然科学基金创新研究群体项目（项目编号：71521001）
资助。

序 言

博士后制度在我国落地生根已逾 30 年，已经成为国家人才体系建设中的重要一环。30 多年来，博士后制度对推动我国人事人才体制机制改革、促进科技创新和经济社会发展发挥了重要的作用，也培养了一批国家急需的高层次创新型人才。

自 1986 年 1 月开始招收第一名博士后研究人员起，截至目前，国家已累计招收 14 万余名博士后研究人员，已经出站的博士后大多成为各领域的科研骨干和学术带头人。其中，已有 50 余位博士后当选两院院士；众多博士后入选各类人才计划，其中，国家百千万人才工程年入选率达 34.36%，国家杰出青年科学基金入选率平均达 21.04%，教育部“长江学者”入选率平均达 10% 左右。

2015 年底，国务院办公厅出台《关于改革完善博士后制度的意见》，要求各地各部门各设站单位按照党中央、国务院决策部署，牢固树立并切实贯彻创新、协调、绿色、开放、共享的发展理念，深入实施创新驱动发展战略和人才优先发展战略，完善体制机制，健全服务体系，推动博士后事业科学发展。这为我国博士后事业的进一步发展指明了方向，也为哲学社会科学领域博士后工作提出了新的研究方向。

习近平总书记在 2016 年 5 月 17 日全国哲学社会科学工作座谈会上发表重要讲话指出：一个国家的发展水平，既取决于自然科学发展水平，也取决于哲学社会发展水平。一个没有发达

的自然科学的国家不可能走在世界前列，一个没有繁荣的哲学社会科学的国家也不可能走在世界前列。坚持和发展中国特色社会主义，需要不断在实践和理论上进行探索、用发展着的理论指导发展着的实践。在这个过程中，哲学社会科学具有不可替代的重要地位，哲学社会科学工作者具有不可替代的重要作用。这是党和国家领导人对包括哲学社会科学博士后在内的所有哲学社会科学领域的研究者、工作者提出的殷切希望！

中国社会科学院是中央直属的国家哲学社会科学研究机构，在哲学社会科学博士后工作领域处于领军地位。为充分调动哲学社会科学博士后研究人员科研创新的积极性，展示哲学社会科学领域博士后的优秀成果，提高我国哲学社会科学发展的整体水平，中国社会科学院和全国博士后管理委员会于2012年联合推出了《中国社会科学博士后文库》（以下简称《文库》），每年在全国范围内择优出版博士后成果。经过多年的发展，《文库》已经成为集中、系统、全面反映我国哲学社会科学博士后优秀成果的高端学术平台，学术影响力和社会影响力逐年提高。

下一步，做好哲学社会科学博士后工作，做好《文库》工作，要认真学习领会习近平总书记系列重要讲话精神，自觉肩负起新的时代使命，锐意创新、发奋进取。为此，需做到：

第一，始终坚持马克思主义的指导地位。哲学社会科学研究离不开正确的世界观、方法论的指导。习近平总书记深刻指出：坚持以马克思主义为指导，是当代中国哲学社会科学区别于其他哲学社会科学的根本标志，必须旗帜鲜明加以坚持。马克思主义揭示了事物的本质、内在联系及发展规律，是“伟大的认识工具”，是人们观察世界、分析问题的有力思想武器。马克思主义尽管诞生在一个半多世纪之前，但在当今时代，马克思主义与新的时代实践结合起来，愈来愈显示出更加强大的生命力。哲学社会科学博士后研究人员应该更加自觉地坚持马克思主义在科研工作

中的指导地位，继续推进马克思主义中国化、时代化、大众化，继续发展 21 世纪马克思主义、当代中国马克思主义。要继续把《文库》建设成为马克思主义中国化最新理论成果宣传、展示、交流的平台，为中国特色社会主义建设提供强有力的理论支撑。

第二，逐步树立智库意识和品牌意识。哲学社会科学肩负着回答时代命题、规划未来道路的使命。当前中央对哲学社会科学愈发重视，尤其是提出要发挥哲学社会科学在治国理政、提高改革决策水平、推进国家治理体系和治理能力现代化中的作用。从 2015 年开始，中央已启动了国家高端智库的建设，这对哲学社会科学博士后工作提出了更高的针对性要求，也为哲学社会科学博士后研究提供了更为广阔的应用空间。《文库》依托中国社会科学院，面向全国哲学社会科学领域博士后科研流动站、工作站的博士后征集优秀成果，入选出版的著作也代表了哲学社会科学博士后最高的学术研究水平。因此，要善于把中国社会科学院服务党和国家决策的大智库功能与《文库》的小智库功能结合起来，进而以智库意识推动品牌意识建设，最终树立《文库》的智库意识和品牌意识。

第三，积极推动中国特色哲学社会科学学术体系和话语体系建设。改革开放 30 多年来，我国在经济建设、政治建设、文化建设、社会建设、生态文明建设和党的建设各个领域都取得了举世瞩目的成就，比历史上任何时期都更接近中华民族伟大复兴的目标。但正如习近平总书记所指出的那样：在解读中国实践、构建中国理论上，我们应该最有发言权，但实际上我国哲学社会科学在国际上的声音还比较小，还处于有理说不出、说了传不开的境地。这里问题的实质，就是中国特色、中国特质的哲学社会科学学术体系和话语体系的缺失和建设问题。具有中国特色、中国特质的学术体系和话语体系必然是由具有中国特色、中国特质的概念、范畴和学科等组成。这一切不是凭空想象得来的，而是在中

国化的马克思主义指导下，在参考我们民族特质、历史智慧的基础上再创造出来的。在这一过程中，积极吸纳儒、释、道、墨、名、法、农、杂、兵等各家学说的精髓，无疑是保持中国特色、中国特质的重要保证。换言之，不能站在历史、文化虚无主义立场搞研究。要通过《文库》积极引导哲学社会科学博士后研究人员：一方面，要积极吸收古今中外各种学术资源，坚持古为今用、洋为中用。另一方面，要以中国自己的实践为研究定位，围绕中国自己的问题，坚持问题导向，努力探索具备中国特色、中国特质的概念、范畴与理论体系，在体现继承性和民族性，体现原创性和时代性，体现系统性和专业性方面，不断加强和深化中国特色学术体系和话语体系建设。

新形势下，我国哲学社会科学地位更加重要、任务更加繁重。衷心希望广大哲学社会科学博士后工作者和博士后们，以《文库》系列著作的出版为契机，以习近平总书记在全国哲学社会科学座谈会上的讲话为根本遵循，将自身的研究工作与时代的需求结合起来，将自身的研究工作与国家和人民的召唤结合起来，以深厚的学识修养赢得尊重，以高尚的人格魅力引领风气，在为祖国、为人民立德立功立言中，在实现中华民族伟大复兴中国梦征程中，成就自我、实现价值。

是为序。

王康

中国社会科学院副院长

中国社会科学院博士后管理委员会主任

2016年12月1日

摘 要

为了对科学知识网络建立共性的理论基础,本书从科学大数据知识元共现的角度进行了深入探讨,将现有的几种科学知识网络纳入以共现为特征的理论体系之中,并对以共现为基础的新型知识网络做了研究展望。

本书的内容除引言部分外,主体框架可分为三大部分:

第一部分(第一章至第三章)是学科知识共现网络组合的理论与方法研究,在此理论上构建了系统的科学知识网络类型。目前,知识元的利用受到了学界广泛重视,知识元的共现关系又是科学知识之间普遍联系的依据,使用数学组合知识可以将这些科学知识元组合出多种知识网络类型。这个过程的实现需要靠数据库中科学数据元的重组来完成,科学文献具体字段之间的关系通过重构后的知识网络类型基本可以分成三类,其中常见的典型知识网络对科学创新路径与方法具有重要的启示。

第二部分(第四章至第八章)主要阐述了几种具体科学知识网络的类型、构建、演化及其与科学创新路径间的关系。第四章针对科学引文网络,总结了其发展历程、发展方向和误区,开发了其构建办法、流程并以管理科学的数据作为实证案例,在科学引文网络方面的继承与发展路径上,如层出不穷的新材料无一不是建立在对传统材料物理性质的深化研究之上,没有现有的传统材料,新材料的研发也是无从谈起的;第五章总结了主要的三种知识网络类型,并分别研究文献共被引网络、期刊共被引网络和著者共被引网络的发展历程、研究方法和与科学创新的关系,如每一次成功的共被引交叉均是原有理论与研究对象的完美结合;第六章从科研主体间的关系出发,总结了合作网络研究的历程、

方法和类型,接着分别研究了国家合作网络、机构合作网络与著者合作网络的结构特征及与科学创新之间的关系,如管理科学就是跨学科合作的例子,它成功把各自为政的两大学科知识通过交流合作,用来解决同一个问题对象;第七章从词汇提取与矩阵生成出发研究词汇网络的规律,总结了作者关键词共词网络和机标关键词共词网络两种类型,如熵理论目前已泛化至化学、生物学、材料科学甚至人文社会科学当中,但均使得熵概念在多学科中共同出现并成为前沿热点;第八章涉及语义知识网络,它是近年来新发展出的研究热点,但这种网络也是经过叙词表对科学文献处理后的标引词之间的共现网络组成的,这种网络对科学文献内容的解释与展示更为充分,它也是学科发展到一定阶段的必然产物,本书总结了语义网络的组合理论、构建方法和发展动力、关键难题,从这种层级关系的图谱上可以看出学科的发展前沿,发现潜在热点,预测下一个可能渗入的具体领域,提前对科学研究进行创新布局。

第三部分(第九章至第十章)重点对科学知识网络中的科学创新规律及未来趋势和存在问题进行思考,学科知识网络主要是沿着处理网络科学数据、发展混合知识网络以及深化研究手段等几个方向发展,其存在的主要问题是知识网络将会对现有知识进行整序、发展更大数据的处理技术和对知识管理产生潜在影响。第十章对全书做了总结,指出了本书的研究不足并对后续要进行的研究做了展望,希望在本书研究的基础上未来不但能够获得更为完整的理论体系,还能持续对科技创新有所启示。

本书研究内容的作用和意义主要体现在以下几个方面:

(1) 从数据库以及互联网获取来的科学大数据,按照学科或者领域建立学科数据库后,基于科学知识元共现关系对这些数据进行知识组织,借由知识组织与知识发现的方法面向学科或者领域科学知识进行整合、挖掘出更多的潜在新知识,使得领域知识得到增值,从而帮助科研人员实现学科或者领域内的科学得到创新,形成以科学数据驱动的知识创新发展模式。

(2) 在现有科学知识网络的基础上,深入挖掘了科学大数据中的基本共现关系,使用数学组合的方法进行了科学知识元组合,

形成了以共现理论为主的可扩展性较强的科学知识共现网络体系。该知识共现理论除了能使现有知识网络统一在同一个理论框架体系之内外,还可以继续组合获得更多的新型科学知识网络,形成系统的知识共现理论与方法体系。

(3) 在知识元共现理论框架下统一的科学引文网络、文献共被引网络、科研合作网络、学科共词网络和学科语义网络五大主要内容,组成了科学知识网络体系。在科学大数据背景下这些知识网络群能对科学创新路径与模式有所启示,使得科学知识网络带来科学上的潜在创新,并通过发现数据背后的创新规律,使得知识网络从静态数据走向动态管理,从历史数据推向未来预测,从已知知识扩大到未知领域。

因此,本书在科学知识元的基础上将学科知识元重组成知识网络,可以为科学创新带来更多的启示并使知识增值;对于科学知识创新路径和模式的发现,主要是通过知识共现网络来进行;在研究知识共现网络方面,又是沿着从知识创新的客体网络到主体网络,文献级别的网络到属性级别的网络,一般共性的规律到特殊个性的规律这样的逻辑次序依次展开。

关键词: 知识组织; 引文网络; 合作网络; 语义网络; 知识网络

Abstract

To establish a common theoretical basis for scientific knowledge network, this book applied co - occurrence theory study them from the perspective of scientific data knowledge element, summarized several existing scientific knowledge networks according the co - occurrence rules in a self - consistent theoretical system and pointed that the co - occurrence is the key to future knowledge networks as the prospect of the book.

This paper can be divided mainly into three parts except the introduction section:

The first part (Chapter 1 - 3) is the construction of knowledge co - occurrence networks theory and introduction of research methods under combination of knowledge units. A systematic scientific knowledge network type are built under the basis of the mathematical combination theory. The application of knowledge element called lots of academic attentions, and co - occurrence relationship between knowledge units is the key connection between the scientific knowledge, using the mathematical combination method the knowledge units can be combined into various types of knowledge networks. This process can be completed through knowledge units restructuring. Relationships between given fields reconstructed under new rules can be divided into three types, of which the typical knowledge networks have tremendous implications for the common scientific innovation path and method.

The second part (Chapter 4 - 8) mainly described several types of specific scientific knowledge networks, especially the evolution and in-

novation relationship between scientific knowledge networks laws. For the science citation network, we summarized its history, development direction and misunderstanding, developed the construction methods and took Chinese natural science data as empirical case studies, proving that the innovation law in science citation network nodes, such as the outcome of the power law distribution and citation delays; for the co-citation network we summarized the main three types of knowledge networks, and summarized literature in co-citation network, journal co-citation network, and author co-citation network relationship development process, research methods and the scientific innovation, and finally took quantum mechanics co-cited by chemistry and physics as example, illustrating the co-citation network played important role in innovation process; for scientific research cooperation network, we studied the relationship between the main departure, summed up its course, method and network types, and then investigated relationship between innovation as well as national cooperative network, the institution network and individual network structure, taking translational medicine as a case study of these three networks in interdisciplinary cooperation and at last, summarized several innovative revelation; for co-word network, we created this knowledge network matrix from the lexical rules, summarized the author co-word and keyword plus co-word keywords are two types of co-words networks, in which applied entropy theory co-word network from natural sciences to humanities and social sciences to summarize the concept of innovation; for semantic knowledge networks, this network is a thesaurus of the standard scientific literature network after treatment between scientific literature class content and presentation, the co-occurrence network can composed fully explained for the subject to a certain stage. As a new research focus in recent years, the semantic network theory, construction method and develop dynamic, the key object was discussed, and finally using the medical informatics literature data semantic network to show the innovative rule.

The third part (Chapter 9 - 10) focused on the law of scientific

knowledge networks in innovation on science and future trend. Knowledge networks will develop in the main way of processing future scientific big data network, the hybrid networks to deepen the study of the means of knowledge understanding and several other development directions. The main problems involve on the knowledge network will give the order for existing knowledge, development of big data processing technology and produce a greater potential impact on information science. Finally, after pointing out the shortcomings of this study and the follow-up study, we hope that this book will not only be able to obtain a more complete theoretical system for the future, but also have implications for continued innovation.

The main research contents of this book are mainly focused on the following aspects of innovation and function:

(1) The scientific big data under the rule of knowledge organization and knowledge discovery method in given subject or field based on scientific knowledge element co-occurrence relationships from the database and internet access into data warehouse, in accordance with the discipline or field established, could integrate scientific knowledge, mine more potential new knowledge, add the domain knowledge value, thereby helping researchers achieve scientific innovation on their subject or in their field, form and develop scientific data driven knowledge innovation model.

(2) The basic scientific data co-occurrence relations were mined for current scientific knowledge, and scientific knowledge element combination were studied using the method of combinatorial mathematics, forming a scalability strong scientific knowledge co-occurrence network system based on the co-occurrence theory. The knowledge co-occurrence theory not only unite the existing knowledge networks in a same consistent theoretical framework but also obtain more new scientific knowledge network by combined method to form systematic knowledge co-occurrence theory and new method system.

(3) The knowledge co-occurrence theory framework grouped sci-

ence citation network, co - citation network, scientific collaboration network, co - word network and semantic network as main five major contents of knowledge network system. Under the environment of scientific big data, the knowledge networks can shed some light on the scientific innovation path and mode, making scientific knowledge network bring potential scientific creative ability, and by discovering the innovation rule of data behind, take the knowledge network from the static data to dynamic management, from historical data toward future predictions and from the known knowledge expanding to the unknown ones.

The network knowledge in this book were re - composed on the basis of knowledge element, bring more inspiration and make knowledge value added for scientific innovation. We discovered scientific knowledge innovation path and pattern mainly by co - occurrence network knowledge, and crested knowledge co - occurrence network along such a logical order in turn from the knowledge network innovation network object to subject, the literature level network to unit level network, the common law to the special character.

Key Words: Knowledge Orgnization; Citation Networks; Cooperation Networks; Sematic Networks; Knowledge Networks

目 录

引 言	1
第一章 科学知识网络的研究历程及研究现状	3
第一节 科学知识网络的研究背景与意义	8
一、科学大数据增长态势迅猛	10
二、科学知识网络的研究背景	13
三、科学知识网络的研究意义	18
第二节 科学知识网络的国内外研究现状	22
一、国外科学知识网络的研究现状	22
二、国内科学知识网络的研究现状	27
三、科学知识网络国内外研究评述	30
第三节 科学知识网络基本原理与方法	32
一、科学知识网络研究目标	33
二、科学知识网络范围界定	33
三、科学知识网络研究方法	36
第二章 科学知识共现网络的组合理论与方法	43
第一节 科学知识共现关系及网络形成	45
一、何谓科学知识元共现关系	45
二、知识元共现关系理论基础	46
第二节 科学知识元共现理论与机制研究	50
一、知识元共现关系的内在机制	55
二、知识元共现关系的外部实现	56