

■ 冉从敬 著

知识产权学术演化研究

——基于信息可视化技术的视角



科学出版社

知识产权学术演化研究

——基于信息可视化技术的视角

冉从敬 著

科学出版社
北京

内容简介

本书利用信息可视化工具,将知识产权的研究文献按照知识产权总论、著作权、专利权和商标权进行了科学计量和可视化研究。具体探讨了知识产权的学术源流,透析了知识产权的学术集群,解读了知识产权的学术渗透,凸显了知识产权的学术热点。最后,本书在纵观全书的基础上,推演和预测了知识产权各领域的十大研究趋势。

本书可供知识产权和图书情报研究的高校教师、研究生及从事知识产权法律研究和实践的学者学习与参考。

图书在版编目(CIP)数据

知识产权学术演化研究:基于信息可视化技术的视角 / 冉从敬著. —北京:科学出版社, 2015

ISBN 978-7-03-046475-0

I. ①知… II. ①冉… III. ①知识产权—研究—中国 IV. ①D923.404

中国版本图书馆 CIP 数据核字 (2015) 第 282658 号

责任编辑:徐倩/责任校对:马显杰

责任印制:霍兵/封面设计:无极书装

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

文林印务有限公司印刷

科学出版社发行 各地新华书店经销

*

2015年11月第一版 开本:720×1000 1/16

2015年11月第一次印刷 印张:17 1/2

字数:352 000

定价:98.00元

(如有印装质量问题,我社负责调换)

作者简介

冉从敬，中国政法大学知识产权法专业博士后，武汉大学图书馆学系副主任，武汉大学信息资源管理专业博士，美国伊利诺伊大学香槟分校访问学者，武汉大学珞珈青年学者，武汉市黄鹤英才，中国国家专利信息师资人才，美国信息科学与技术协会（Association for Information Science and Technology, ASIS&T）会员。主持国家自然科学基金、国家社会科学基金、教育部人文社会科学规划项目、博士后科学基金特别资助、博士后科学基金一等资助、湖北省科技支撑计划软课题等重要项目 10 余项；参与国家科技支撑项目、国家社会科学基金重大项目、中国南方电网有限责任公司委托项目、百度公司委托项目等重要项目 30 余项；发表论文 70 余篇，出版专著 4 部，出版合著 1 部，参编著作 8 部。其主要研究方向为知识产权、信息咨询、专利分析。

前 言

保护知识产权,保护人类的智力成果,已经成为当代世界关注的热点。在知识经济条件下,知识发挥着重要的作用,而真正起主导作用的却是知识产权。近年来,我国对知识产权的重视程度明显提升。2008年6月5日,国务院颁布了《国家知识产权战略纲要》,并将知识产权工作上升到国家战略层面进行统筹部署和整体推进,为知识产权事业发展指出了明确方向。2014年基本实现了《国家知识产权战略纲要》确定的第一阶段五年目标,对促进经济社会发展发挥了重要的支撑作用。为了进一步推进《国家知识产权战略纲要》的实施,2014年,由国家知识产权局等部门联合制订了《深入实施国家知识产权战略行动计划(2014—2020年)》。根据该计划到2020年,国家知识产权法治环境将更加完善,创造、运用、保护和管理知识产权的能力将显著增强,知识产权意识深入人心,知识产权制度对经济发展、文化繁荣和社会建设的促进作用将会充分显现。按照《深入实施国家知识产权战略行动计划(2014—2020年)》的部署,国家知识产权局制订了《2015年国家知识产权战略实施推进计划》,该推进计划共确定了5方面的重点工作、80项的具体措施。据了解,2015年国家知识产权战略实施的重点工作包括强化知识产权保护,鼓励创新创造;促进知识产权创造运用,支撑产业转型升级;加强知识产权管理和保护,促进创新成果转移转化;拓展知识产权交流合作,推动国际竞争力提升;加大支持力度,提高知识产权战略实施保障水平等。

信息网络技术的不断发展,数字图书馆、网络经济等新的技术形态和经济形态对知识产权制度也提出了新的挑战。在近20年来,知识产权学术研究成了图书情报、企业管理、技术创新、法律实践等多个学科领域共同关注的焦点。目前国内在知识产权学术方面的研究已经有了很多的成果,著作权、专利权和商标权等各个知识产权研究的子领域都涌现出了大量的优秀文献,为我国知识产权的立法和实施提供了大量的指导性意见。目前我国已有一批学者利用计量学的方法和可视性技术,对知识产权领域的文献进行了分析,进而梳理出知识产权领域在某一时间段内的发展现状。

本书旨在全面的对知识产权学术研究进行归纳分析,在反映出知识产权这一交叉领域的研究全貌的同时,总结出知识产权在学术演化中存在的规律,运用更多元化的分析模式及可视化技术将知识产权学术研究成果、演化过程及规律更直观地展现出来。作为一种重要的法权和无形资产,知识产权作为无形资产在经济生活中具有可观的经济价值,作为一种法定的智力成果具有巨大的商业竞争价值,其不仅是一种重要的竞争资源,也是企业参与竞争求得生存与发展的基础。

对其学术的演化规律结合信息可视化技术进行研究能够厘清知识产权学术的演化过程,为之后的科学研究提供参考,也为我国企业的知识产权战略决策提供理论支持。

知识产权领域虽然和其他人文社会科学领域有很大差别,但也有很多共性,信息可视化技术在知识产权领域的应用为在其他学科中的应用提供了很好的参考和借鉴作用。本书通过对高质量的知识产权学术研究论文的量化分析,对知识产权学术史的研究、学术前沿的把握和学术生态圈的良性发展有着重要的作用,将区别于传统的定性研究方式。本书通过深入知识产权的实体内容之中,为知识产权制度的完善、知识产权战略的制定提供重要的参考依据。引导学者们在了解知识产权学科演进规律的基础上更高效、有针对性地进行下一步研究,推进知识产权体系的建设,促进知识产权价值的实现,推动企业创新和知识产权的资产化、资本化、产业化,激励高校和科研单位知识产权的创造。

本书在传统研究方法的基础上,注重采取定性分析与定量分析相结合的研究方式,尤其侧重与知识图谱相结合的定量分析,运用独特视角考察知识产权学术产生、发展的整个演化过程。主要运用引文分析法、共被引分析法、词频分析法、共词分析法和社会网络分析法等研究方法,对知识产权学术研究过程进行宏观分析,对具体的研究热点、学术团体、研究趋势等方面做微观细致的解读。本书利用信息可视化技术分析知识产权领域的学术演进历程,通过绘制知识产权领域的权威作者共被引图谱、经典文献共被引图谱、学术机构及作者集群图谱、术语突增图谱和术语时区图谱等知识图谱,清晰并形象地梳理和展现了知识产权领域的学术源流及发展进程。对近年来知识产权的研究重点和研究现状进行了总结归纳,并预测了未来知识产权研究的发展趋势。

在内容上,本书聚焦于如下方面:知识产权学术源流的可视化研究、知识产权学术集群的可视化研究、知识产权学科渗透的可视化研究、知识产权学术热点的可视化研究和知识产权学术趋势研究。在本书的撰写过程中,研究生和知识产权管理实务人士参与了研究和写作。各章的撰写人为第1章:陈一、殷娜、冉从敬;第2章:冉从敬、陈一梅;第3章:冉从敬、涂文艳;第4章:冉从敬、李天才;第5章:冉从敬、肖兰;第6章:冉从敬、郭晓婉、赵洋和李新来。

本书受博士后科学基金项目“基于信息可视化技术的知识产权学术演化规律研究”、国家社会科学基金重大项目“国家知识产权文献及信息资料库建设研究”、国家自然科学基金面上项目“基于专利计量的专利战略构建模型研究:以云计算产业为实证对象”的资助。科学出版社的责任编辑徐倩同志为本书的出版付出了辛勤的劳动,在此一并感谢!利用信息可视化技术来研究知识产权是一个新的尝试,书中的疏漏之处在所难免,敬请专家学者们批评指正!

冉从敬

目 录

第 1 章 学术演化探测：原理、工具与方法	1
1.1 原理	1
1.2 工具	5
1.3 方法	8
第 2 章 知识产权学术源流的可视化研究	15
2.1 知识产权文献共被引分析	15
2.2 著作权文献共被引分析	27
2.3 专利权文献共被引分析	36
2.4 商标权文献共被引分析	43
第 3 章 知识产权学术集群的可视化研究	53
3.1 研究背景及现状	53
3.2 知识产权领域合作网络分析	56
3.3 知识产权分领域作者合作图谱分析	64
3.4 知识产权分领域机构合作图谱分析	75
3.5 国内知识产权研究合作网络结构演变	84
第 4 章 知识产权学科渗透的可视化研究	113
4.1 医学中的知识产权	114
4.2 经济学中的知识产权	120
4.3 图书情报学中的知识产权	139
第 5 章 知识产权学术热点的可视化研究	157
5.1 知识产权热点分析	157
5.2 著作权研究热点分析	172
5.3 专利权研究热点分析	189
5.4 商标权研究热点分析	204
第 6 章 知识产权学术趋势研究	222
6.1 知识产权十大趋势推演和预测	222
6.2 著作权十大趋势推演和预测	231

6.3 专利权十大趋势推演和预测	241
6.4 商标权十大趋势推演和预测	251
参考文献	262

第1章 学术演化探测：原理、工具与方法

自人类出现伊始，人们就开始有意识地对知识进行不断的归纳、总结。近年来，由于电子存储技术、数据传输、信息检索和各种可视化技术的发展，使人们获取海量信息和处理海量信息的能力大大加强。知识图谱（mapping knowledge domains）的理论和实践也得到了空前的发展。多领域的专家学者利用这种直观、形象的研究方法梳理学术观点，探测学术走向。

1.1 原理

进行知识产权学术演化探测必须尊重一定的具有普遍意义的基本规律，这些规律是在大量观察、实践的基础上，经过归纳、概括而得出的。本书基于信息可视化技术的视角，具体以知识图谱的表现形式展开研究，而厘清知识产权学术化探测的原理，是开展研究的基础工作。

1.1.1 知识图谱的发展

知识图谱是显示科学知识的发展进程与结构关系的一种图形，用信息可视化技术描述人类随时间拥有的知识资源及其载体，绘制、挖掘、分析和显示科学技术知识及它们之间的相互联系，在组织内创造知识共享的环境以促进科学技术研究的合作和深入。由于知识图谱是以科学知识为计量的研究对象，因此，其属于科学计量学（scientometrics）的范畴。

知识图谱同绘图学和地图学也有一定关系，但知识图谱的概念与知识地图的概念并不完全相同。“地图”是一个以二维或三维空间的形式显示地形和人类活动及相关特征的地理学概念，知识地图最初是表现科学技术活动与知识的地理分布状况的地图；而“图谱”则是一个图像以一定空间形式在一定时间范围内展现变化的系统概念，虽然可以把知识地图作为知识图谱的一种形式，但知识图谱比知识地图更能揭示知识之间的联系及知识的进化规律（陈悦和刘则渊，2005）。

知识图谱的理论与方法近年来成为科学计量学研究的最前沿，科学计量学家们将应用数学、图形学、信息可视化技术及计算机科学等学科的理论和方法

传统的科学计量学引文分析方法结合起来,用信息可视化的知识图谱形象地展示科学学科的结构与发展。绘制知识图谱的目的是应用图示的方法解释潜在的学科结构。

从 20 世纪 50 年代开始,科学引文索引 (Science Citation Index, SCI) 开始了大规模的商业性应用。加菲尔德 (E.Garfield) 1955 年发表于 *Science* 杂志的关于引文索引的文献奠定了引文分析的基础。作为美国科学情报研究所 (Institute for Scientific Information, ISI) 的创始人,加菲尔德推动了代表学术共同体的多学科数据库——科学引文索引数据库的发展。不仅如此,加菲尔德还为研究科学的动态发展状况设计了一系列成熟的概念性工具。引文分析的概念已经成为当今科学计量学、文献计量学、信息计量学、网络计量学的基础。

加菲尔德的发明极大地改变了科学计量学家们研究科学共同体的方式,20 世纪 60 年代早期,加菲尔德等开始了基于引文数据的先锋性研究——在《应用引文数据撰写科学历史》(*The Use of Citation Data in Writing the History of Science*) 中绘制了脱氧核糖核酸 (deoxyribonucleic acid, DNA) 研究领域的历史发展图谱,不久之后,普赖斯用相同的数据在他一系列经典著作,如《巴比伦以来的科学》、《小科学,大科学》和《科学文献的网络》中进行了知识图谱绘制的开创性工作,尽管当时并没有使用知识图谱这一概念,但是实际上以引文分析为基础的知识图谱理论与方法已经应运而生了。

在知识图谱中,学科前沿之间的交互关系是以空间的形式展现出来的。研究发现,科学引文与被引文之间往往存在着学科内容上的联系。通过引文聚类分析,特别是通过引文间的网状关系研究,能够探明有关学科之间的亲缘关系和结构,划定某学科的作者集群,分析推测学科间的交叉、渗透和衍生趋势,还能对某一学科的产生背景、发展概貌、突破性成就、相互渗透和今后的发展方向进行分析,从而揭示科学的动态结构和某些发展规律 (邱均平, 2001)。

1.1.2 信息可视化与知识图谱

1987 年美国国家科学基金会举办的可视化会议,首次正式提出了可视化 (visualization) 这一术语。可视化最开始应用在科学计算中,把科学数据以直观的图形、图像显示出来,使随时间和空间变化的物理现象或物理量呈现在研究者面前。具体来说,可视化是指在人通过视觉观察并在头脑中形成客观事物的影像的过程,这是一个心智处理过程。可视化提高了人们对事物的观察能力及整体概念的形成等;可视化结果便于人的记忆和理解,同时其对于信息的处理和表达方式具有使用其他方法无法取代的优势。信息可视化技术以人们惯于接受图形、图像并辅之以信息处理技术,将被感知、被认知、被想象、被推理、被综合及被抽

象了的对象属性及其变化发展的形式和过程,通过形象化、模拟化、仿真化、现实化的技术手段表现出来。可视化不仅是客观现实的形象再现,也是客观规律、知识和信息的有机融合。可视化技术不仅可以用来表现静态的知识,同时也可用于动态地描述和表达客观对象的发展演化规律及进行动态知识的获取。简言之,信息可视化就是用图形正确的表现复杂的信息和逻辑之间的关系。

1989年,斯图尔特·卡德(Stuart K.Card)、约克·麦金利(Jock D.Mackinlay)首次提出了“信息可视化”这个概念(Robertson et al., 1989)。信息可视化的主要对象是非空间、非数值和高维的大数据集,其主要作用是对这些数据集进行交互式视觉表现,其目的是发现这些海量信息之间的相互关系,进而推理出其发展趋势。艾普勒(Eppler)和布卡(Burkhard)于2004年提出知识可视化(Eppler and Burkhard, 2004)。知识可视化运用视觉表征手段,促进群体知识的传播和创新。知识可视化的目标在于传播见解、经验、态度、价值观、期望和预测等,并以这种方式帮助他人正确的重构、记忆并应用这些知识(周宁等, 2007)。信息可视化是利用计算机支撑的、交互的、对抽象数据的可视表示。知识可视化是将数据挖掘和知识发现等方法所获得的知识与规律,尤其是知识的构成和知识之间的逻辑关系,以及具有复杂结构的知识,利用可视化的方法表现出来,使知识便于理解。知识图谱就是利用信息可视化技术,根据共被引分析、共现分析等理论基础,构建的一种知识之间的关系的网络图,常见的网络有时序网络、共引网络、共词网络、耦合网络、合作网络等。知识图谱包含了以文献等信息为节点、以它们之间关系为边的链型、树型、网型等结构图形。这些关系聚类图的可视化包括作者、文章、期刊、关键词、学科等类型,节点分别是作者、文章或期刊等。知识图谱和信息检索过程可视化、信息检索结果可视化一样都属于信息可视化的一个重要分支(秦长江和侯汉清, 2009)。

1.1.3 学科知识图谱

在现实中,知识图谱不仅可以应用在描述学科发展及动态状况,明确主要研究领域之间的内部联系、追踪学术前沿等方面,而且还可以应用在各个研究领域之间的合作网络、知识输入与知识输出的相互关系,识别“无形学院”等方面。其中,学科知识图谱是知识图谱的一种具体运用方式,即通过对某学科在特定时间段内发表的学术论文或者专著的作者、题名、关键词、作者机构等信息进行可视化分析,从而得出该学科在特定时期内形成的以作者、文献、期刊、机构等为节点的知识图谱,具有相似特征的节点在图谱中聚成一类,并对这些聚成一类的图谱进行学科解读,从直观上展现学科的整体情况。正如爱因斯坦所指出的那样:“科学必须创造自己的语言和自己的概念,供他自己使用。科学的概念最初是日

常生活中所使用的普通概念,但它经过发展就完全不同。它们已经变换过了,失去了普通语言所带有的含糊性质,从而获得了精确的定义,这样它们就能使用于科学的思维”(爱因斯坦和英费尔德,1962)。

学科知识图谱所描绘的对象主要包括从事科学技术活动和作为知识载体的主体,包括科学家、技术专家、项目组、实践团体或某一学科领域;显性或编码化的知识,如论文、专利、所学课程、数据库或类似的应用等;过程或方法,包括研究问题和解决问题的过程与方法、组织的业务流程,以及相关的知识投入等(杨虹,2008)。由于学科知识图谱具有相对成熟的理论基础,分析指标众多,可以快速处理大量学术文献,并且具有清晰直观等优点。本书就借此对知识产权整个学科发展状况和发展轨迹进行整体分析。学科知识图谱具有以下主要特征。

1. 可视化方法

学科知识图谱是对基于学科的知识结构进行可视化,它通过对学科中各个不同知识单元间关系的展示,揭示某一学科的知识发展规律。一些针对结构的可视化是从几何结构角度进行的,如针对超链接结构的可视化,但是从认知角度进行的可视化更加重要,它往往可以通过分析文献内容来进行。在知识经济的背景下如何实现对知识的获取、组织、检索及服务是非常迫切和必要的研究课题。学科知识图谱在计算机技术的支持下,对大量信息及信息间的关系进行抽象后,借助一定的图形化方式加以表现,揭示出隐藏在其中的模式和规律,以便更加有效、直观地与信息进行交互。“可视化的方法不仅在揭示学科信息资源的广度与深度上有很大的优势,而且能够将隐藏在信息资源内部的、复杂的、抽象的语义,以直观的图形方式呈现出来,为研究者直观地过滤、理解和方便获取大规模数据和信息提供了有效途径,从而发现信息之间的关系特征和规律”(张钢和倪旭东,2005)。

2. 静态的形式显示动态的过程

学科知识图谱的可视化是一个动态的过程,它通过学科各知识节点之间的关系,各知识节点的位置是由几何方法聚类而成,将隐藏在大量数据背后的学科发展信息转化成见解、经验、态度、价值观、期望和预测等,并以“一图胜千言”的表达方式帮助人们重构、记忆并应用这些知识,以静态的图谱表现隐藏在基础知识中的动态信息(赵国庆等,2005)。

3. 知识依赖性

学科知识图谱的表现形式可以是线性二维图形,也可以是立体三维图形,两者都是通过图示的展示方法揭示知识的空间结构。因此,学科知识图谱有别于一般的地图或图像,它是对知识产品的二次以上的加工,它的绘制需要建立在知识基础之上,但该基础必须是一次加工后的学科发展的客观精确的统计数据或文本

知识信息，不能是一般的数据或者人为编制的虚拟知识基础。

4. 空间关联性

学科知识图谱能反映学科体系内存在的“科学共同体”的结构与发展，而且知识单元之间的交互关系都能以空间的形式展现出来，这样就能够探明有关学科之间的亲缘关系和结构，划定学科的学者集群及“无形学院”，分析推测学科间的交叉、渗透和衍生趋势，对学科的发展历史、突破性成就、发展趋势进行分析，从而揭示出学科的动态结构。因此，二维或三维的学科知识图谱，都揭示了图谱中点与点的相互关联性，点与点的距离可以表示关系的亲疏，线条粗细与点之间的强度的大小，图谱的局部与整体之间的关系则可呈现出知识单元发展、变化的过程。

1.2 工具

工具是学术演化探测的基础，不同的数据库为学术研究提供了充实的数据支撑，各类可视化分析软件也具有不同的分析侧重点，为学术信息的挖掘提供工具支持。

1.2.1 数据库

1. 科学引文索引数据库

美国科学引文索引于1957年由美国科学情报研究所在美国费城创办。科学引文索引、工程索引(The Engineering Index, EI)、科技会议录索引(Index to Scientific&Technical Proceedings, ISTP)是世界著名三大科技文献检索系统，是国际公认的进行科学统计与科学评价的主要检索工具。

科学引文索引已经成为当今世界最为重要的大型数据库，被列在国际六大著名检索系统之首。它不仅是一部重要的检索工具书，而且也是科学研究成果评价的一项重要依据。它已成为目前国际上最具权威性的，用于基础研究和应用基础研究成果的重要评价体系。它是评价一个国家、一个科学研究机构、一所高等学校、一本期刊，乃至一个研究人员学术水平的重要指标之一。

科学引文索引以布拉德福(S. C. Bradford)文献离散律理论、加菲尔德引文分析理论为主要基础，通过论文的被引用频次等的统计，对学术期刊和科研成果进行多方位的评价研究，从而评判一个国家或地区、科研单位、个人的科研产出绩效，来反映其在国际上的学术水平。

科学引文索引以其独特的引证途径和综合全面的科学数据，通过统计大量的

引文得出某期刊、某论文在某学科内的影响因子、被引频次、即时指数等量化指标来对期刊、论文等进行排行。被引频次高,说明该论文的学术水平很高,在它所研究的领域里产生了巨大的影响,被国际同行重视。由于科学引文索引收录的论文主要是自然科学的基础研究领域,所以科学引文索引指标主要适用于评价基础研究的成果,而基础研究成果的主要表现形式是学术论文。所以,如何评价基础研究成果也就常常简化为如何评价论文所承载的内容对科学知识进展的影响。

2. 中国知网数据库

中国知网,即中国国家知识基础设施(China National Knowledge Infrastructure, CNKI),由世界银行于1998年提出。CNKI工程是以实现全社会知识资源传播共享与增值利用为目标的信息化建设项目,由清华大学、清华同方发起,始建于1999年6月。CNKI工程的具体目标:一是大规模集成整合知识信息资源,提高资源的综合性和增值利用价值;二是建设知识资源互联网传播扩散与增值服务平台,为全社会提供资源共享、数字化学习、知识创新信息化条件;三是建设知识资源的深度开发利用平台,为社会各方面提供知识管理与知识服务的信息化手段;四是为知识资源生产出版部门创造互联网出版发行的市场环境 with 商业机制,大力促进文化出版事业、产业的现代化建设与跨越式发展。

CNKI是全球信息量最大、最具价值的中文网站。据统计,CNKI网站的内容数量大于目前全世界所有中文网页内容的数量总和,可谓世界第一中文网。CNKI的信息内容是经过深度加工、编辑、整合、以数据库形式进行有序管理的,内容有明确的来源、出处,内容可信可靠,如期刊、报纸、博士学位论文及硕士学位论文、会议论文、图书和专利等。因此,CNKI的内容有极高的文献收藏价值和使用价值,可以作为学术研究、科学决策的依据。基于对文献内容的详细标引,CNKI文献搜索提供了对标题、作者、关键词、摘要、全文等数据项的搜索功能,保证了文献的检全率和检准率。

3. 中文社会科学引文索引数据库

中文社会科学引文索引(Chinese Social Sciences Citation Index, CSSCI),是由南京大学中国社会科学评价中心开发研制而成。CSSCI遵循文献计量学规律,采取定量与定性评价相结合的方法从全国2700余种中国人文社会科学学术性期刊中精选出学术性强、编辑规范的期刊作为来源期刊。目前收录包括法学、管理学、经济学、历史学、政治学等在内的25大类的500多种学术期刊。

1.2.2 软件

1. Bibexcel

Bibexcel 是由瑞典科学家佩尔松 (Persson) 开发的文献计量学研究软件, 它帮助用户分析文献数据或者是文本类型格式的数据, 实现引文分析。Bibexcel 处理的数据来自集成的 ISI Web of Knowledge 平台上的数据库, 包括 Web of Knowledge 数据库、Deiwent Innovation Index 数据库和 Medline 数据库等。Bibexcel 除了可以对数据的相关知识单元 (作者、关键词、参考文献等) 做频次分析和排序外, 还能实现知识单元的共现关系矩阵。

2. SPSS

SPSS (statistical product and service solutions, 统计产品与服务解决方案) 软件。最初软件的全称为“社会科学统计软件包” (solutions statistical package for the social sciences), 但是随着 SPSS 产品服务领域的扩大和服务深度的增加, SPSS 公司已于 2000 年正式将英文全称更改为“统计产品与服务解决方案”。它是世界上最早的统计分析软件, 由美国斯坦福大学三位研究生 Nie、Hull 和 Bent 于 1968 年研究开发成功。

SPSS 是世界上最早采用图形菜单驱动界面的统计软件, 它最突出的特点就是操作界面极为友好, 输出结果美观漂亮, 几乎所有的功能都以统一、规范的界面展现出来, 使用 Windows 的窗口方式展示各种管理和分析数据的功能, 对话框展示各种功能的选择项。SPSS 的基本功能包括数据管理、统计分析、图标分析、输出管理等。SPSS 统计分析过程包括描述性统计、均值比较、一般线性模型、相关分析、回归分析、对数线性模型、聚类分析、数据简化、生存分析、时间序列分析、多重相应等几大类, 每类中又分多个统计过程, 如回归分析中又分线性回归分析、曲线估计、Logistic 回归、Probit 回归、加权估计、两阶段最小二乘法、非线性回归等多个统计过程, 而且每个过程中又允许用户选择不同的方法及参数。SPSS 也有专门的绘图系统, 可以根据数据绘制各种图形。

3. Pajek

Pajek (斯洛文尼亚语中意为蜘蛛) 软件, 是一款基于 Windows 的大型网络分析和可视化软件, 由 Mravar 和 Batagelj 于 1996 年开发。不同于一般的网络分析软件, Pajek 软件可以处理拥有多达几百万节点的大型网络, 突破了很多网络分析软件只能处理较小规模数据的瓶颈。它可以从大规模网络中提取出若干小网络, 以便于使用经典算法实现更加细致的研究, 并通过强大的可视化功能将网络及分析结果展示出来。Pajek 软件具有灵活的输入方式, 可以直接定义一个小网络, 也可以从外面导入数据生成网络, 除了本身的数据格式之外, 它还支持很多其他软

件数据格式的导入。软件结构是建立在网络、分类、向量、排序、群和层级六种数据结构之上的(王柏等, 2007)。

4. Hist Cite

Hist Cite = history of cite, 意味着引文的历史, 或者被称为引文图谱分析软件。该软件是由加菲尔德倡导开发。设计 Hist Cite 的初衷是展示引用关系是如何随着时间而延续的。通过从 Web of Science 数据库得到的数据, 生成一系列表或图。在 Hist Cite 中, 整个 Web of Science 数据库中被引用的次数被称为全域被引次数(global citation score, GCS), 而在所下载的数据中的被引频次称为局域被引次数(local citation score, LCS)。

在引文历史图中, 文章按照其发表年份进行排列。发表最早的文章放在图的最上方, 发表越晚的文章位置越靠下。如果文章 A 引用了文章 B, 那么这种联系就用一条从 A 到 B 的连线表示。一篇文章的被引次数, 对应于其节点的大小。

1.3 方法

在具体的学术探测中, 为了实现不同的研究目标, 我们尝试利用不同的分析方法。各种分析方法各有侧重, 不同分析方法相互独立又互相补充, 共同构成了整体的知识产权学术探测研究方法体系。

1.3.1 引文分析理论与方法

所谓引文分析(citation analysis), 就是利用各种数学及统计学的方法和比较、归纳、抽象、概括等逻辑方法, 对科学期刊、论文、著者等各种分析对象的引用现象进行分析, 以便揭示其数量特征和内在规律的一种文献计量分析方法(邱均平, 2001)。随着分析对象的不同, 引文计量(citation measure)指标也有所不同。例如, 对科学期刊进行分析常用的计量指标主要有五种, 即引文率、影响因子、自引证率、自被引率和当年指标; 在对专业或学科结构进行研究时, 除用引文率外, 还可用引文耦合和共被引等测度指标进行分析。引文率是由计量指标来衡量的, 我们可以根据不同的需要规定出可从不同角度进行引文测度的各种指标。但这些指标都是由一篇篇科学论文之间的引用关系建立起来的, 或者说是由引文率演变而来的。因此, 引文率是各种引用分析中最基本的测度。引文分析主要有三个主要领域: ①对科学家、出版物和科学机构进行定性和定量评估; ②模拟科学技术的历史发展; ③信息搜索与检索(Zunde, 1971)。

Weinstock (1971) 指出, 文献被引用有以下 15 种动机: ①对开拓者表示尊

重；②对有关著作给予荣誉；③核对其所用的方法及仪器；④提供背景阅读材料；⑤对自己的著作予以更正；⑥对别人的著作予以更正；⑦评价以前的著作；⑧为自己的主张寻求充分的论证；⑨提供研究者现有的著作；⑩对未被传播、很少被引或未被引证的文献提供向导；⑪鉴定数据及物理常数等；⑫核对原始资料中某个观点或概念是否被讨论过；⑬核对原始资料或其他著作中的起因人物的某个概念或名词；⑭承认他人的著作或概念；⑮对他人的优先权要求提出争议。

通过文献引用频率的分析研究可以测定某一学科的影响和某一国家某些学科的重要性，科学引文与被引文之间往往有着学科内容上的联系。通过引文聚类分析，特别是通过引文间的网状关系研究，能够探明有关学科之间的亲缘关系和结构，划定某学科的作者集群，分析推测学科间的交叉、渗透和衍生趋势，还能对某一学科的产生背景、发展概貌、突破性成就、相互渗透和今后的发展方向进行分析，从而揭示科学的动态结构和某些发展规律。

从根本上说，文献的相互引用是由科学本身的发展规律和研究活动规律所决定的。科学学的研究反复表明，科学知识具有明显的累积性、继承性；任何新的学科或新的技术，都是在原有学科或技术的基础上分化、衍生出来的，都是对原有学科或技术的发展；也就是说，科学技术的发展是连续的。同时，由于科学的统一性原则，现有的各个学科之间都是彼此联系、相互交叉、相互渗透的。因此，任何一项科学研究都必须在前人研究成果的基础上，吸取他人的经验来进行，因此，作为科学知识记录和科研成果反映的科学文献也必须是相互联系的。在创作科学论文时作者不可避免地要引用其他有关的文献，为论证自己的观点寻找依据，查考资料，在发表科学论著时，作者列出其所引用过的参考文献一方面是为了说明引用资料的出处以强调其可靠性，同时也便于读者查考、核对或在此基础上进行更深入的研究；另一方面也说明作者讲究科学道德、尊重他人劳动，这都是科学活动中必须遵循的行为准则。尊重科学文献的相互引用是科学发展规律的表现，也是科学活动中普遍存在的一种必然现象。

1.3.2 共被引分析理论与方法

用被共引分析 (co-citation analysis) 研究科学演进及变化的理论开始于 20 世纪 70 年代，由斯莫 (Small) 和格里菲斯 (Griffith) 及伊琳娜·玛莎科娃 (Irina Marshakova) 分别提出 (Small and Griffith, 1974)。当两篇文献共同出现在第三篇文献的参考文献目录中时，这两篇文献就成为被共引的关系。共引频率定义为这两篇文献一起被引用的频率。经常一起被引用的文献 (作者)，则表示它们在研究的主题、理论或方法上是相关的。为此，共被引分析认为文献 (作者) 共被引的次数越多，它们之间的关系就越密切，“距离”也就越近。利用现代的多元