

教育政策与法律可视化研究丛书

陕西师范大学优秀著作出版基金资助

中国教育法学的知识图谱研究 1985—2015

祁占勇 著



科学出版社

中央高校科研基金重点资助项目“中国教育政策与法律研究热点的
知识图谱分析”（15SZZD01）的研究成果

中国教育法学的知识图谱研究 1985—2015

祁占勇 著



科学出版社
北京

内 容 简 介

教育法学是以法学和教育学作为理论指导,以教育法律问题为研究对象,通过对教育法律现象的提炼和分析,揭示教育法律规律,指导教育法律实践的一门新兴的社会学科。它是法学的一个分支学科,具有应用性、规范性、描述性、实践性等特征。本书利用知识图谱分析的方法,对 1985—2015 年我国教育法学发展中的研究热点、核心作者、研究范式、前沿演进、合作网络等问题进行知识图谱分析,从而全面、系统地梳理教育法学的研究热点领域、主流学术群体、学科理论结构、经典文献、科研合作等情况,为教育法学研究的拓展与深化提供基础性材料。

本书可供普通高等学校教育学及相关专业的本科生、研究生使用,也可供广大一线教师、教育管理工作或教育学爱好者参考与学习。

图书在版编目(CIP)数据

中国教育法学的知识图谱研究:1985—2015/祁占勇著. —北京:科学出版社,2019.6

(教育政策与法律可视化研究丛书)

ISBN 978-7-03-060238-1

I. ①中… II. ①祁… III. ①教育法-法的理论-研究-中国-图谱
IV. ①D922.161-64

中国版本图书馆 CIP 数据核字(2018)第 293116 号

责任编辑:杜长清/责任校对:王晓茜

责任印制:徐晓晨/封面设计:正典设计

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>

北京虎彩文化传播有限公司印刷

科学出版社发行 各地新华书店经销

*

2019 年 6 月第 一 版 开本:720×1000 B5

2019 年 6 月第一次印刷 印张:9 1/2

字数:163 000

定价:89.00 元

(如有印装质量问题,我社负责调换)

目 录



第一章 知识图谱理论及基本方法	1
一、知识图谱理论概述	1
二、知识图谱的基本方法	7
第二章 中国教育法学研究热点的知识图谱	11
一、数据选择和研究方法	11
二、研究过程与分析	13
三、研究结论与展望	49
第三章 中国教育法学学科范式的知识图谱	57
一、核心作者影响力分析	58
二、核心作者共被引知识图谱分析	69
三、结论与反思	74
第四章 中国教育法学研究前沿的知识图谱	79
一、数据选择与研究方法	80
二、研究前沿的知识图谱分析	86

三、渐强型研究前沿、渐弱型研究前沿与最新研究前沿分析	89
四、研究结论	114
第五章 中国教育法学合作网络的知识图谱	117
一、作者合作网络知识图谱分析	118
二、机构合作网络知识图谱分析	137
后记	145

第一章

知识图谱理论及基本方法

过去，人们对一个学科研究进行资料综述的时候，更多的是基于个体主观经验对资料进行加工，较少采用科学计量的方法对资料进行综合分析。面对今天浩如烟海的文献，我们如果没有科学的计量分析方法，仅凭个人的主观经验判断难免会产生错误或者不当的归类和总结。随着互联网技术的普及和数字化时代的到来，采用信息可视化（information visualization, InfoVis）技术，对已有信息进行整理，使新的知识的科学计量学逐渐发展并且成熟起来。知识图谱（mapping knowledge domain）作为当前国际科学计量学领域热门的研究方法之一，越来越受到研究者的重视和青睐。2003年，美国国家科学院组织了“Mapping Knowledge Domains”讨论会，预示着世界科学计量学中知识图谱和可视化研究的春天已经到来。

要认识和把握知识图谱的准确含义，需要对其概念、发展历程、应用原理及意义进行全面的了解。

一、知识图谱理论概述

知识图谱理论研究是一个以科学为基础，应用数学、信息科学及计算机科学诸学科交叉的领域，是科学计量学和信息计量学的新发展，其基本原理是分析科学文献、科学家、关键词、引文等的相似性及测度。根据具体的技术软件与操作方法的不同可以绘制不同类型的知识图谱。

（一）知识图谱的概念

了解地图、知识地图及图谱的概念有助于准确把握知识图谱的概念。

1. 地图与知识地图

地图（map）是以二维或三维空间形式显示地形和人类活动及相关特征的地理学概念。地图能够科学地反映出自然和社会经济现象的分布特征及其相互关系。在电子和数字时代，地图已经由传统的纸质地图演变成了数字地图和电子地图，但不论其形式如何演变，地图的主要特征依然不变。第一，由特殊的数学法则产生的可测量性。特殊的数学法则包含地图投影、地图比例尺和地图定向三个方面。第二，由使用地图语言表示事物所产生的直观性。地图语言包括地图符号和地图注记两部分。第三，由实施制图综合产生的一览性。第四，必须遵循一定的数学法则。地图是绘制在平面上的，必须准确地反映它与客观实体在位置、属性等要素方面的关系。第五，必须经过科学概括，缩小了的地图不可能容纳地面上所有的现象。第六，具有完整的符号系统。

知识地图（knowledge map）也称为知识分布图或知识映射图，最初源于美国捷运公司绘制的充满知识资源的美国地图。此后，知识地图表示带有索引号或用其他方式表示层次关系的表格和文件，以及用来表示信息资源与各部门或人员之间关系的信息资源管理表和信息资源分布图。知识地图描述了一个组织在知识转化周期过程中的知识资源具体分布及变化情况，有助于组织成员把握本系统内部知识配置，为其进行知识寻求与创造提供准确的可用信息。

知识地图的绘制包括三个步骤：首先，通过知识收集和知识预处理获得知识元；其次，通过对知识的甄别、筛选和分类，将无序的知识元间信息进行重组，以构建知识之间的相互关系；最后，各个知识之间建立索引，通过索引来连接和标识它们之间的位置和关系，以此来呈现知识地图。绘制知识地图的目的就是对组织知识资源总体分布情况进行可视化描述，包括组织知识资源的存在情况及其载体和资源之间可能存在的联系。知识地图实质上就是利用现代信息技术制作的组织知识资源的总目录和各知识条目之间关系的综合体，以及组织专家的导航系统、构造地图的方法将各类知识资源中的知识关联起来，使之成为一个网络，提高知识的利用率。

2. 图谱与知识图谱

图谱(atlas)指系统地编辑起来的、根据实物描绘和摄制的图,是研究某一学科所用的资料。后来,泛指按类编制的图集。

知识图谱也被称为科学知识图谱、知识域可视化或知识域映射地图,可视化地描述人类随时间的推移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知识的合作和深入。它以科学学为基础,是应用数学、信息科学及计算机科学诸学科交叉的领域,是科学计量学和信息计量学的新发展。知识图谱能够用直观图像展现出最前沿领域和学科知识的信息会聚点,从宏观、中观、微观等不同层面来揭示一个领域或学科的发展概貌,便于人们全面审视一个学科的结构和研究热点、重点等信息,生成新的知识。借助于知识图谱,人们可以将知识和信息中引人注目的最前沿领域或学科制高点以可视化的图像直观地展现出来,挖掘、分析和显示知识及其之间的联系,进而判定学科前沿的历史演进路径。

值得注意的是,虽然可以将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。知识图谱与知识地图的区别在于前者一般不能帮助使用者快速获取与知识资源关联的相关信息。

(二) 知识图谱的发展历程

1. 知识图谱的产生

人类经历了五次信息革命:语言的使用,文字的创造,印刷术的发明,电报、电话、广播、电视的发明和普及,计算机技术及现代通信技术的普及与应用。现在,人类正在经历第六次信息革命——云计算与互联网的发展与应用。在互联网和数字化时代没有到来之前,学者们为了解一个学科领域发展的整体状况,必须查阅该领域的几乎所有文献,然后经过自己加工,从大量文献中筛选出相对重要的文献。这样的工作不仅耗费时间,也非常困难。不同学者选取文献时,因为站立角度和主观判断的差异,就如盲人摸象中的各个盲人,选取的材料往往有很大出入,结论也难以得到重复验证。

随着知识大爆炸和信息化时代的到来,海量信息时代也随之到来。这个时候,通过传统方法来捕捉学科发展的脉动越来越困难。在对多学科领域进行研究

时，对文献的动态发展做一个综述性的回顾尤其困难。因此，迫切需要具有客观性、科学性、高效性的方法来研究科学学科的结构与发展。1955年，尤金·加菲尔德（Eugene Garfield）在 *Science* 上发表了关于引文索引的文献，奠定了引文分析的基础，不仅推动了代表学术共同体的多学科数据库——科学引文索引（science citation index, SCI）数据库的发展，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关注。引文分析概念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。经过多年发展，特别是美国科学信息研究所（Institute for Scientific Information, ISI）提供的SCI数据库诞生后，引文结构的大样本统计分析越来越便利，知识图谱已成为科学共同体结构与发展实证研究的主流方法，广泛用于很多学科领域。

2. 知识图谱的发展

知识图谱的发展经历了三个阶段：

第一阶段，引文分析技术阶段。1999年，亨利·斯莫尔（Henry Small）明确提出借助引文图谱实现科学可视化的途径。从普赖斯（D. Price）、加菲尔德到斯莫尔，逐步确立起日臻完备的引文分析理论与方法，其构成科学计量学的基础与主流，在一定意义上形成了科学计量学中一门成熟的分支学科——引文分析学。20世纪90年代以来，科学计量学运用统计分析、引文分析和网络分析的方法，以及计算机图形学、图像处理与可视化技术，在科学知识图谱和知识可视化方面得到了迅猛的发展。

第二阶段，社会网络分析（social network analysis, SNA）技术阶段。在引文网络研究中，研究者引入复杂网络和社会网络的基本概念与最新成果，把引文分析、复杂网络和社会网络三种理论与方法统一起来，以尽可能把科学知识图谱理论与方法提高到一个新的水平。这种变化不仅可以使学者们深化对引文网络知识分布、知识流动、知识演化等的特有规律的认识，还可以促使学者们探索普遍存在于自然、社会和人文的复杂网络的一般规律，具有重大的学术价值。

第三阶段，可视化知识图谱阶段。1987年，美国国家科学基金会（National Science Foundation, NSF）发表《科学计算中的可视化》，标志着科学可视化的诞生。“信息可视化”最早由乔治·罗伯逊（George Robertson）等在1989年提

出,指在计算机、网络通信技术支持下,以认知为目的,对非空间的、非数值型的高维信息进行交互式视觉表现的理论、方法与技术。计算机可视化信息处理软件,是通过直观的动态图像信息处理的方式,揭示出专业领域中出现的交叉学科的复杂现象,从而获得详尽的前沿科学信息分析结果。它不仅有助于科学家在最短的时间内了解和预测前沿科技研究动态,还有助于科学家在复杂的科研信息中开辟新的未知领域,为其提供快速、独立的科学判断的客观依据。2003年,美国国家科学院组织了“Mapping Knowledge Domains”讨论会,预示着世界科学计量学中知识图谱和可视化研究的春天的到来。大连理工大学的刘则渊教授以此次会议为契机,展开了对知识图谱的研究。他于2005年在国内开展知识图谱研究,于2008年出版了《科学知识图谱:方法与应用》一书。此后,知识图谱的应用研究在国内不断涌现,取得了丰硕的成果。

(三) 知识图谱的种类

知识图谱分为传统科学计量图谱(two-dimensional scientometric map, 2DSM)、三维构型图谱(three-dimensional configuration map, 3DSM)、多维尺度图谱(multi-dimensional scaling map, MDSM)、社会网络分析图谱(social network analysis map, SNAM)、自组织映射图谱(self-organizing map, SOM)、寻径网络图谱(pathfinder network, PFNET)等多种类型。

1. 传统科学计量图谱

通过线性函数产生二维或三维统计图形(柱状图、线形图、扇形图、点分布图等),可以以最直观、简单的方式来展示对知识的统计结果。传统科学计量知识图谱以静态嵌合(mosaic)为主要特征,构成部分之间按照外在确定的标准安排,孤立呈现,彼此间几乎无交流和沟通存在。

2. 三维构型图谱

三维构型图谱是基于三维空间模型所产生的知识图谱。三维空间模型由国际著名科学计量学家克雷奇默(H. Kretschmer)教授于1987年创立。她受梅茨格(W. Metzger)1986年创立的心理学中的“构型”(又称“格式塔”)理论启发,将新的数学方法引入科学计量学领域,借助非线性函数形象地描述了科学家合作网络构型的三维空间模型。三维构型知识图谱中的各组成部分和点处于动态

交互中，有着较密切的交流和相互作用。一旦形成某种有序的格局，不仅其他组成部分和点的位置会被确定，构成自身的组成部分和点的位置也会被其他部分所确定。

3. 多维尺度图谱

多维尺度图谱是基于多维尺度分析（multi-dimensional scaling, MDS）绘制出的知识图谱。多维尺度分析是通过非线性方法，把高维空间的数据转换到低维空间，转换后的数据仍可以较好地反映出原数据间的关系。多维尺度知识图谱中，每一个点表示一个事物或物体。点的位置凭借事物或物体间的相似关系安排。越相似的事物或物体，其所代表的两点间的距离越近；反之，其所代表的两点间的距离越远。多维尺度知识图谱中点处于欧几里得几何空间，可以采用二维、三维或者多维图形来展示它们之间的关系。

4. 社会网络分析图谱

社会网络分析图谱是基于社会网络分析绘制出的知识图谱。社会网络分析开始于20世纪二三十年代的英国人类学研究。在社会劳动中，每个劳动者与其他劳动者间存在或多或少的关系。社会网络分析就是通过构建上述关系的模型，描述和揭示群体间关系的结构对群体功能或者群体内部个体带来的影响。社会网络分析的计量法源于美国社会心理学家莫雷诺（J. L. Moreno）创立的社会测量法。如今，社会网络分析法广泛应用于网络社会关系发掘、支配类型发现，以及信息流跟踪，以判断和解释信息行为和信息态度。

5. 自组织映射图谱

自组织映射图谱是基于自组织特征映射模型（self-organizing feature mapping model）理论绘制出的知识图谱。自组织特征映射模型理论于1981年由芬兰赫尔辛基理工大学卡汉（T. Kohone）教授提出。他针对传统图像分割技术存在的缺陷，即无法模仿人对颜色进行主观区分，提取出森林火灾图像的火焰区域，提出了人造神经中枢网络信息可视化及其重要的自组织特征映射模型。它非常适合对图像进行自适应分割，能够模拟人脑中处于不同区域的神经细胞分工不同的特点，采用无监督的自适应分类方法，按照有序的拓扑映射结构，将任意维的输入信号自动转换到一维或者二维的离散网格上。

6. 寻径网络图谱

1990年,美国心理学家 R. W. Schvaneveldt 提出寻径网络图谱。该方法模拟人脑的记忆和联想方式,形成不同概念或实体间的语义网络。首先,对不同概念或实体间联系的相似性或差异性进行经验性评估;其次,应用图论中的基本概念和原理生成特殊的网状模型。

(四) 知识图谱的特征

知识图谱具备动态性、空间性、知识依赖性、关联性四方面的特征。

(1) 动态性。这是知识图谱区别于传统知识地图的本质特征,它以静态的图谱中点与点及连线的关系形式,揭示知识间隐含的动态结构变化信息。

(2) 空间性。知识图谱以二维或者三维图形展示知识之间的空间结构,通过坐落空间位置和所占据领域大小来标示知识所处战略位置的重要程度。

(3) 知识依赖性。知识图谱与一般图像的区别在于,一般图像是对原材料进行直接加工所生成的,而知识图谱是在对知识进行数据统计或者文本分析后,在所构建的客观知识综合体的基础上再次加工处理所生成的。知识图谱一般无法直接从原始知识材料的加工中产生,它依赖于对原材料的再次加工。

(4) 关联性。知识图谱可以揭示各知识点之间的相互关系,点与点在图谱中所处的距离远近表示知识间关系的亲疏,点与点之间连线的粗细表示知识间关系连接的力度大小。

二、知识图谱的基本方法

1973年,斯莫尔发表《科学文献共被引分析:两个文献间关系的一种新方法》,提出了共被引分析(co-citation analysis)。此方法是在原有引文分析方法的基础上,借助国际上新兴的科学知识图谱绘制方法与信息可视化技术,绘制科学计量学期刊共引知识地图,实现了引证分析的可视化。目前,国内外相关学者正致力于对引文分析这种重要工具的研究和完善。学术界推行最广的《中文核心期刊要目总览》和中国科学技术信息研究所的《中国期刊引证研究报告》就是借助引文分析方法对中国期刊情况进行统计分析的,并根据中国期刊具体情况不断

地开发和完善，以期找寻最适合的分析方法。

（一）引文分析法

最早关注引文分析的学者是美国人谢泼德（Shepard），他于 1873 年创办了《谢泼德引文》（*Shepard's Citation*），供律师或法学家查阅法律判例及其引用情况。1948 年，英国学者布兰德福特（S. C. Bradford）在其专著《文献工作》（*Document*）中提出可定量描述文献序性结构的经验定律，即某学科大量的文献相对地集中在一定数量的杂志上，而剩余部分的文献则分散在其他大量相关杂志上，奠定了核心期刊与非核心期刊的思想。1955 年，美国著名情报学家尤金·加菲尔德提出利用引用文献追踪科学进展的概念，引文分析法正式产生。他于 1963 年创办了 SCI，用于探讨科学的结构、评价与选择情况，考察科学著作及科学家的社会影响等，产生了重要的影响。

引文分析法就是利用各种数学、统计学的方法，以及比较、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和被引用现象进行分析，以便揭示其数量特征和内在规律，实现评价、预测科学发展趋势的一种信息计量研究方法。引文分析的出发点是正文和引文，即引用的文献和被引用的文献。引文分析法中多次被引用的文献足以说明它们涉及的主题或内容受到更多的关注，能够反映出学科领域普遍关注的热点问题。普赖斯认为，在科学论文之间形成的引文网络结构中，只有极少数论文被新发表的论文较多引用，被引频次高的这小部分论文可视为学科的新的生长点，成为热门的学科前沿，为利用引文分析探测科学前沿的可行性奠定了理论基础。

（二）共被引分析法

用共被引分析研究学科演进与变化的理论开始于 20 世纪 70 年代，由斯莫尔和格里菲斯（Griffith）及玛莎科娃（Marshakova）分别提出。共被引分析是指当 2 篇文献共同出现在第 3 篇文献的参考文献目录中，这两篇文献就具有共被引关系。共被引分析大体上分为文献共被引分析、期刊共被引分析、作者共被引分析和学科共被引分析等，文献共被引分析和作者共被引分析是其中最为常用的共被引分析法。

文献共被引关系是最基本的一类共被引关系，主要体现了被一篇文献同时引

用的参考文献之间的关系,进而可以对这种关系进行文献学方面的理论研究。从分析共被引文献的类型、内容等角度出发,可以研究文献体系的特征结构及分布等方面的规律。引文网络的疏密反映引文之间的相互关系、联系特征和发展状况及趋势等。

作者共被引关系是指 N ($N \geq 2$) 个作者发表的文献同时被别的文献引用,则称这 N 个作者具有共被引关系。作者共被引是以文献的作者为基本单元而建立的关系,当 N 位作者被某一主题文献的作者同时引用,被引频率越高,则二者之间专业联系越紧密,因此,通过作者共被引分析可以获得作者数量、构成、活动规律等情况。

(三) 多元统计分析法

多元统计分析是对若干相关的随机变量观测值的分析,主要包括因子分析(factor analysis)(主成分分析)、多维尺度分析和聚类分析。在几何学中,这一过程通过矩阵关系体现。例如,在固定的学科范围内选择一定数量的文献,并研究这些文献群组的共被引矩阵。

因子分析的基本目的是用少数几个因子来描述许多指标或因素之间的联系,即将比较密切的几个变量归在同一类中,每一类变量就成为一个因子,以较少的几个因子来反映原资料的大部分信息。其中的主成分分析法把给定的一组变量通过先行变换,转换为一组不相关的变量,在变换的过程中,保持变量的总方差不变,第一主成分具有最大方差,以此类推。

多维尺度分析通过二维空间来展示作者或文献之间的联系,并利用平面距离远近来体现各个作者或文献之间的相似程度。在知识图谱中,作者或文献的位置体现了其与其他作者或文献之间的相似性,并且处于中心位置的作者或文献与其他作者的联系越多,就越体现了其在这一学科中的中心地位,反之,则处于边缘地带,不是研究关注的重点内容。

聚类分析是最常用的多元统计分析方法,其研究起点是通过原始文献相关数据的矩阵来获得点的二维图。通过聚类分析,可以掌握一门学科研究主要关注的主题和研究热点,把握学科焦点所在。

（四）词频分析法

词频分析法是最常见的文献计量学的统计方法之一，其依据是齐普夫定律（Zipf's law）。在海量文献中，不同词汇或词语的使用及出现频率是有一定规律的，有关词频分布的规律通过齐普夫的研究正式成为一种统计方法。它主要是指通过分析词汇或词语出现频次的多少，确定学科研究热点与发展趋势。

词频分析一般主要包括数据检索、清洗加工、词汇提取、统计分析等阶段，整个过程较为烦琐复杂，结束之后对结果进行解释并得到一些结论。学者们通常是将词频分析法与共被引分析法、多元统计分析法相结合来确定学科前沿的知识图谱。

（五）社会网络分析法

社会网络分析是一门交叉科学技术，其主要技术来源于数学与计算机科学。社会网络是一个人的集合，其中的每一个人都与网络中某个子个体的人相互认识，这样的网络可以用点（或矢量）的集合来代表人，用线的连接表示相识，而关系的强弱通常用线的粗细表示。用社会网络进行分析时，主要用到的指标有网络的最短路径、直径、平均路径长度、网络的密度、度中心性、接近中心等。

社会网络分析可以成功地用于研究科学合作网络和互联网中的可视化网络，可以较好地用于展现一个学科范围的合作网络与发展状况。

第二章

中国教育法学研究热点的知识图谱

如何把握一个知识域的结构与演变,对置身于当今信息社会的我们来说,是一项前所未有的、广泛而迫切的重大课题。^①研究热点是某学科领域多数学者共同关注的研究对象,是随着时代需要与学科自身发展而变迁的动态变量,反映了某一段时间内该学科的研究人员对一些特定问题的持续关注度。对研究热点的判断有两类基本方法:一类是经验判断方法,依据本学科领域研究者的经验,借之以文献检索,总结归纳出本领域某一时期的研究热点;另一类是科学计量学的方法,主要以文献的“关键词”为线索,通过知识图谱和多元统计分析,依据关键词之间关联关系的强弱来挖掘学科知识结构之间的亲疏关系程度,将海量的文献信息转换成可视化的知识图谱,直观地展现不同时期的研究热点及变化。

本书研究以科学计量方法为主,通过对1985—2015年中国教育法学研究热点的知识图谱进行分析,直观展示30年来本学科热点问题。同时,为了弥补知识图谱方法的不足,笔者通过阅读和经验判断,对知识图谱体现的研究热点问题进行了主观分析和必要的修正。

一、数据选择和研究方法

关键词是学术论文的重要组成部分,虽然往往只有2~5个词,却涵盖了文章的主要思想。人们不仅可以利用关键词检索到文献,还可以通过关键词了解到文

^① 刘则渊,陈悦,侯海燕,等. 科学知识图谱:方法与应用. 北京:人民出版社,2008:序言4.

献涉及的领域和内容。^①两篇文章中所论述的内容和范围越相似，那么相同的关键词将越多；这些相关的关键词在同一篇文献中出现的频率越高，说明它们之间的关系越密切，该关键词也就代表该学科研究的一个方向，是学科内研究的热点问题。“一个学科较长时间内的数量学术论文的关键词集合，可以很好地揭示出该学科领域的总体内容特征及其发展趋势。”^②

（一）资料来源

教育法学研究热点的资料来源于中国学术期刊网络出版总库，采用标准检索，将期刊年限设定为“1985—2015年”，指定“教育法”为检索内容，共获得相关文献 8 279 篇，为了保证研究的可靠性与有效性，采取去除会议纪要、人物专访、报纸评论、刊物征稿要求、征订启事、刊物总目录信息等非研究型文献的方法，得到 7 617 篇有效文献，形成研究资料。

（二）分析依据与关键词的选择

在科学计量学中，经常以文献的关键词作为研究热点的依据，通过关键词之间的联系掌握文献所研究的主要内容和领域范围。本书利用 BICOMB（bibliographic item co-occurrence matrix builder，书目共现分析系统）共词分析软件对 1985—2015 年中国教育法学 7 617 篇学术论文的关键词进行了处理和分析。初步统计结果表明，一方面，教育法学论文在关键词标注上存在一定的随意性，使得 7 878 个即 65.2% 的关键词只出现过 1 次，关键词共现的概率较低。另一方面，关键词标注存在大量近义词，且分布分散，无法实现有效的共现分析，需做进一步规范化处理。因此，本书需要对这些关键词进行标准化处理，因为同一个意思的关键词在不同论文中的表述会有所差别。例如，将“受教育权利”“受教育权”合并为“受教育权”；“教育法律”“教育法规”“教育法律法规”合并为“教育法律法规”；“《职业教育法》”“职业教育法”合并为“职业教育法”；“高校”“高等学校”合并为“高校”；“自主办学”“办学自主权”合并为“办学自主权”。经过标准化处理后，共提取 12 067 个关键词，总出现频次

① 储节旺，郭春侠. 共词分析法的基本原理及 EXCEL 实现. 情报科学, 2011, (6): 931-934.

② 张婷. 科学传播研究热点的演进——基于科学知识图谱的可视化研究. 科普研究, 2010, (1): 17-27, 47.