

第1章 导论

20 世纪 90 年代以来,各种科学知识网的普及创新了人们获取知识的方式,基于因特网的新型学术交流方式正在走进普通人的生活。在知识经济社会,知识和人力资本在经济发展中的作用越来越重要,已成为创新活动的主要推动力,这就要求人们更多地关注学术研究及其可能的应用,专业学者需要突破研究疆域在跨界视野中实现对复杂研究对象的思考,以打开学术信息交流的新渠道;普通民众和媒体在捕捉大量与自身工作和发展相关的专业信息时,更渴望快捷、便利地获取相关领域的知识和潜藏于标题背后的知识结构信息。知识的消费层次也由传统的以获取单一型的事实和机理为主的“无机温饱型”模式为主向深度发展出以搜索复合型知识的产生、脉络和联系为主的“有机营养型”模式,从而呈现出多层需求共存的状态。作为图书馆学、情报学与档案管理学科的一个重要分支,知识模型和文献计量学理论的核心是要揭示知识资源运动变化的规律,旨在为文献检索和科研文献的交流分析提供有效的信息。为此,需要研究知识传播的特征以及影响知识传播发展的各种因素,并且要对这些差异和形成原因进行分析,于是知识模型和文献计量学不可避免地受到计算机、统计学、应用数学、信息处理等多个学科的影响,在不断吸纳最新理论和技术的过程中获得发展,整个文献计量学也正是在寻找知识形成规律答案的过程中建立起来的。

科学史的研究工作表明,知识的增长和科学的发展具有明显的继承性和积累性,任何的知识创新与技术进步,都是在原有科学或技术基础上发展、演化和衍生出来的,所以对一个学科知识体系规律发展变化的认识是从事学术研究的基础。另一方面,科学的统一性原则也表明,不同的学科、领域之间存在着广泛的合作、交叉、关联和渗透,学术研究和其源源不断创造的跨领域实践是学术研究的生命力的轴心,所以学术研究不仅仅是从文献中发展新文献,不仅仅代表某个学科或某个人的思想,而是通过学者从诸多文献的学术根基中发展新概念并将其按一定的逻辑结构呈现出来的知识流,文献是知识的载体,文献的内容于是作为一个更高层次理论的远景资源,实现一个学术共同体的主要思想脉络的传承和延续。而这一脉络可能横跨不同学科和不同的领域,代表着一个个独特的学术社会。理解这些学术社会的存在需要用一个复杂而庞大的学术系统以及这个系统中的一组能够辨认出这些社会的性质来刻画,所以,一篇文献的价值不仅仅是一项新理论提出的知识根基,而更应该被视为一个学术复杂社会结构中不可或缺的一元,这个学术社会可能是一个复杂的科学问题或社会问题,也可能是某个复杂的分工系统或任务系统。所以,首要的任务是对学术系统这个复杂系统本身的认识,作为由人类智慧积累而建立起来的这样一个复杂而庞大的学术系统,它在解决复杂问题中的分工特性、知识体系构成、迁移互动规律及演化趋势是怎样的呢?各个领域之间是如何彼此影响,互相促进的呢?回答这些问题都离不开文献这个知识的载体以及由这个

载体所分散出来的各种活性知识,这些有再生功能的活性知识形态各异,规模大小不等,彼此相互作用,构成了知识元,这些新型的知识元的生命周期不同,吸引力、黏合度各不相同,它们在大数据背景下层见迭出,令人应接不暇,其价值品质需要甄别,其相互作用需要深入分析,其市场需求需要追踪探测。而要度量这些新的知识元,解读知识元的相互作用,揭示其作为知识传播的客观存在,就必须借助文献计量学及其发展出的复杂的知识模型。

文献计量学的历史接近 100 年,其理论的发展主要得益于两次出版技术的革命^[1],第一次是印刷过程的数字化和标准化,比如:论文的提交规范、查阅和出版等,大量出版物低成本快速推出。这促使人们不得不关注更深层次引起传播和交流的原因,分析科学成果的源变机理。在探寻源头的过程中,计量学家收获了一项发现:科学文献之间的相互引证关系提供了科学知识传播最原始的交流证据,这一交流证据的发现和确立激发了对文献、作者和科学主题演变三者关系的研究,由此形成了文献计量学早期的三大经典定律,奠定了文献计量学的发展基础,文献计量学随即迎来了应用工程时代,这也推动了学术出版领域里的第一项重大工程引文数据库的建设。开发引文索引的初衷是将其作为一种追踪考据的检索工具,但随着这一关联数据的增加,一些科学的发展脉络也逐渐清晰起来,引证关系随之逐渐演变成知识价值的衡量依据。许多研究指出文献的质量与其引文有高度相关性,于是基于引文数据库的学术作品评价方法获得推崇,并首先被用于自然科学领域。目前学术界比较流行的评价指标如美国科学情报研究所的期刊引证报告(JCR)中的 SCI 影响因子 IF(Impact Factor),它是一种期刊的年度影响因子,用于比较研究成果的影响力,也反映刊物的学术水平。引文分析和评价自形成以来的 60 年里,在指导编制各种新型检索工具,为科学管理提供量化依据,探讨科学结构,评价与选择期刊,考察科学著作及科学家的社会影响等方面产生了积极作用。然而,当评价工具获得大范围使用的同时,对其争议也褒贬不一。比如作为理工科领域成果质量评价标准的 SCI、EI 等,近年来因其对学术能力的过度干涉所引起的各种学术不端也招致这些评价本身受到争议,即便如此,这个标准在专业的根基部分、同一类型、同一研究领域内还是获得共识。其实,影响因子在人文社会科学文献价值评价领域的应用一直存有较大争议,首先是 ISI 数据库对人文社会科学领域的文献覆盖率不足 40%,而相比于文献覆盖率达到 80% 以上的生物化学、物理学、化学、临床医学等学科而言,其评价的可靠性受到质疑。第二,人文学科文献的本土化特征强于自然科学文献,虽然用于人文社会科学文献评价的三大数据库 Web of Science、Google Scholar、Elsevier 文献质量很高,但其收录的主要为英文文献。除了出版语言和引文规模等主要因素以外,人文学科文献特征还在研究对象、研究方法、观点逻辑、研究团队、研究周期、传播速度、出版类型等多种因素上与自

然科学文献存在较大差异,如不加区分地使用影响因子值来衡量不同学科期刊的学术水平,势必造成评价结果的偏颇,不能反映真实的学术水平,这些争论也从一定程度上反映了对人文社会科学文献资源整体认识上的局限性和评价方法稀缺的问题。

互联网信息时代推动了第二次出版业的技术革命,整个出版周期向互联网转变,更快、更廉价的全程全方位的知识通信出现,跨领域的学习和融合研究趋势明显,从专业化走向专业融合时仅仅利用知识管理所建立起来的权威引导期刊已不能满足多样化的知识再加工的需要,今天,如何协调知识消费者在面对庞大的学术资源思维割裂时的迷茫,依循怎样的知识分布规律进行协同创新已成为知识管理的新问题。另一方面,互联网中文献的内容资源被结构化,展现出其作为复合型知识构成中活性元块的作用,这使得设计伴随着多样化的知识需求开展基于文献语料的知识模型研究成为可能,其研究价值体现在三方面:第一方面,文献的内容分析是决定文献传播的内生因素,这是因为文献的内部语料结构决定了创新知识所必需的素材供应和研究主题的保持能力,这是决定文献提供的对问题的认知发展所需要的知识基础以及研究领域之间知识交换的重要因素;第二方面,文献的大范围调遣使用是研究主题得以传播的外生因素,这为严整地把握知识的体系特征并控制其在不同情境中的恰当运用提供了契机;第三方面,对于跨学科文献而言,只有通过对其跨领域知识元结构的分析,才能将由社会背景所驱动的研究主题与构成学术的形式和逻辑因素相连,从而揭示出研究主题背后的学科环境与主题发展脉络相互结合的生成语境,而被揭示出的跨领域互联场景又是新的复合型知识的生产土壤。

导论部分首先对知识模型的理论进行综述,接下来重点对人文社会科学学术评价理论进行简要回顾,最后反思现有人文社会科学学术成果评价理论在提取学术成果有价值信息中的局限性。

1.1 知识模型理论的演进

文献计量学(Bibliometrics)是以科技文献的各种外部特征作为研究对象,用数学和统计学的方法定量分析研究科学技术发展现状与规律的交叉学科,其目的在于全面认识文献,能够在浩如烟海的文献中为读者提供全面而有效的所需资料的搜索,并有能力对资料的价值进行鉴别。知识模型是文献计量学和情报学的核心概念,用于知识的形式化和抽象化。学术资源知识模型的核心就是要回答在既定的学术资源背景下,学术资源中有价值的知识的形式及其提取的过程,旨在揭示学术资源中知识体系的形成过程以及原因。为此,需要研究有价值的知识即知识的载体学术成果的形成过程,并且要对学术成果的形成因素及其差异进行分析。整个文献计量学都可以视为是在寻找学术领域形成的知识脉络过程中建立起来的,

这一过程可以分为三个阶段。

1. 文献经验定律的形成阶段

文献计量方法的研究最早追溯到1917年由科尔(F. J. Cole)和伊尔斯(N. B. Eales)对解剖学期刊的统计并用作国别的分析工作^[2]；1923年休姆(E. W. Hulmn)首次提出“统计书目学”^[3]；1969年,英国目录学家阿伦·普理查德(A. Pritchard)第一次使用文献计量学(Bibliometrics)这一术语取代“统计目录学”,并给出定义:“文献计量学是把数学和统计学应用于图书和其他交流媒介的一门学科。”这一定义标志着文献计量学的正式诞生,由于它具有一定的普遍性和代表性,被情报理论界广泛接受和认可^[4]。文献计量学的重要基础理论由布拉德福分散定律、洛特卡定律、齐普夫定律三大定律构成^[5]。

布拉德福定律是由英国著名文献学家布拉德福(S. C. Bradford)于20世纪30年代提出,用于刻画文献分散规律的经验定律。如果将科技期刊按其刊载某学科专业论文的数量多少以递减顺序排列,那么可以将期刊分为专门面对这个学科的核心区、相关区和非相关区。各个区的文章数量相等,核心区、相关区、非相关区期刊数量成 $1:n:n^2$ 的关系^[6]。

洛特卡定律是由美国学者洛特卡(A. J. Lotka)在20世纪20年代提出,用于描述科学生产率的经验规律,又称“倒数平方定律”^[7]。它描述的是科学工作者人数与其所著论文之间的关系:写 n 篇论文的作者数量约为写一篇论文作者数量的 $1/n^2$ ……而写一篇论文作者的数量约占所有作者数量的60%。该定律被认为是第一次揭示了作者与论文数量之间的关系。

齐普夫定律是美国学者齐普夫(G. K. Zipf)于1949年首次出版的《人类行为与最省力法则——人类生态学专著》中提出的词频分布定律^[8]。这部专著提出了“最省力法则”,表述为:如果把一篇较长文章中每个词出现的频次统计起来,按照高频词在前、低频词在后的递减顺序排列,并用自然数给这些词编上等级序号,即频次最高的词等级为1,频次次之的等级为2,……,频次最小的词等级为D。若用 f 表示频次, r 表示等级序号,则有 $fr=C$ (C 为常数)。该式称为齐普夫定律。

早期的文献计量学一直围绕着这几个定律沿着验证与完善这些经验定律和扩大并推广这些经验定律的实际应用这两个方向发展,由这些方法体系所衍生的词频分析法、引文分析法、共引分析法等在研究实践中获得广泛应用。

范并思1986年指出^[5],经验定律在文献计量学的形成和发展中起着理论引导作用,为学者在进行科学研究时提供了理论背景,这些理论不仅规范了科学认识的逻辑,而且还为科学工作者提供了思维的准则和框架,指引着寻求科学问题的大致方向。文献计量学中的经验定律作为理论导引的存在,其作用是把人们定量地研究文献情报的工作纳入经验定律所涉及的范畴,并在确定的范畴中展开新的研

究,这大大减少了探索文献情报流规律的盲目性。

文献计量学的三大经典规律所揭示出的信息资源在传播交流中的规律,在其他网络链接环境中也同样适用。比如 2008 年康柏公司研究实验室、IBM 研究中心以及 Alta-Vista 公司的科学家们通力合作,成功绘制出世界上第一张互联网结构全景图。其基于对 6 亿多个不同网页的研究结果表明:WWW 网站的链接指向,实际上集中构成四个部分,这四部分之间的相互位置关系,恰似一个领结,这就是互联网世界著名的“领结”理论^[9]。按照“领结”理论,约 75% 的 WWW 网站位于领结上,余下的 25% 远离领结并且完全隔离。领结可再划分为三部分:核心部分、左端点部分和右端点部分,各网站紧密联系用户可通过超链接在其中双向随意来往。这三部分的形象正如领结,形式上与布拉德福定律中依照专题论文相关性强弱分区很类似^[10]。

2. 引文评价文献计量学的发展

经典定律所提供的理论框架虽然对文献计量学的贡献巨大,但是从 20 世纪 50 年代起,文献计量学理论研究的核心问题开始由经验定律转变为对趋势规律的发现研究,这主要归功于美国情报学家尤金·加菲尔德创建的科学信息机构(ISI)的科学引文索引数据库的发展。

引文分析的思想可追溯到 1873 年美国出版的可供律师查阅法律判例的检索工具《谢泼德引文》的尝试^[11]。但当时由于缺乏现代化手段,数据量小,不可能进行多方位的引文分析。1958 年,美国情报学家加菲尔德(E. Garfield)与他人合作,为美国卫生研究所的基因研究部门编制《遗传学引文索引》;随后加菲尔德于 1955 年和 1972 年发表了论文^[12,13],首次建立了系统的引文分析理论体系,用来评估学术文章、科学收藏(杂志、会议)被继承和利用的情况。

论文数据库的建立体现了文献计量学科的新定位——提供文献交流的新证据,为实现这一目标,文献计量学这一阶段的研究较侧重于文献计量学定律的应用和评价指标的研究。加菲尔德首次提出了利用论文的相互引用关系来组织和发现科学文献,这一概念在自然科学领域首先发挥作用。这一理论的出现,引发了科学研究的一场革命,其系统地提出了用引文索引检索科技文献的新方法,打破了传统分类法和主题法在检索方法中的垄断地位,确立了参考文献作为学术传承与科学发展轨迹的科学地位。这是因为学术研究必然建立在前人的研究成果之上,参考文献揭示了哪些文献对后续研究产生重要影响,并将该信息传递给学术共同体。以文献引用关系为支撑,科学计量法以文献对后续研究的影响数量作为文献的计量标签,研究不同文献传播深度和广度的差异,结合统计分析、矩阵分析、网络分析、聚类分析等数学工具来定量地研究文献的学术能力,预测科学前沿发展态势,评价科学活动、科学机构的水平及科学论文的质量。以加菲尔德首创的引文分析为方法论的依据,美国科技信息研究所于 1963 年正式出版了《科学引文索引》

(SCD),《科学引文索引》的问世极大地刺激和促进了引文分析方法的发展。随后又诞生了多种科学文献计量学的评价工具,如《科学评论索引》(ISR)、《科技会议录索引》(ISTP)及《工程索引》(EI)等。这四种科技索引刊物一直都是各个国家评价其科学技术发展的重要标准和科技影响力的衡量指标,也是高校科技管理的基本参考依据,成为科学项目立项、科研成果评奖及教学质量评估等业务工作的重要参考。以引文关系为依据的评价系统很快扩展到人文社会科学研究领域^[14]。基于引文评价的常用文献科学评价指标如下:

(1) 引用文献次数(Citation Times),是计算某一期刊当年度被 ISI 引用文献索引资料库所收录期刊引用的全部次数;

(2) 影响因子(Impact Factor),是计算某一期刊的论文在某一年度被引用频率的方法。其计算方式是以该刊前两年中所出版的论文在当年被引用的次数和,除以该刊前两年中出版论文总数。影响因子越高,表示被引用的次数越高,说明该刊的影响力也越大或关注程度越高。

(3) 即时指数(Immediacy Index),是计算某一期刊的论文在该年度多久被引用的方法,计算方式是以某一期刊当年论文被引用的次数,除以该期刊当年出版的论文总数,越早刊登出版的论文,越有机会在当年中被引用,其立即引用率也越高。立即引用率与期刊出刊是否延迟、出刊资料是否陈旧有密切关系。一般而言,当年出刊的论文在当年被引用的机会似乎不太高,除非是极度顶尖发展或预先发布的研究信息(发表的论文很热门)。

(4) 自我引用率(Self-Citing Rate)和被自我引用率(Self-Cited Rate),前者是指某一期刊引用自身期刊的百分比,后者是某一期刊论文被该期刊出版的其他论文所自我引用的百分比。

(5) 引用半衰期(Citing Half-Life)与被引用半衰期(Cited Half-Life),前者是计算某一期刊平均引用多久以前论文的方法(生命力周期),后者是计算某一期刊平均多久以前的论文被引用的方法。从计算当年往前累计,当引用或被引用的论文篇数达全部 50% 时,该篇论文的出版年到当年的间隔年数,即为此期刊的引用半衰期或被引用半衰期。半衰期的值越低,表示该论文引用参考文献或被引用作为参考文献的时效性越高。

引文分析方法能够有效检索出主题法无法查出的相关文献,从而更准确地评价了文章的传播能力,进而更好地评价了作者的影响力^[15]。引文分析方法主要采用文章与其所引参考文献之间的关系来建立引文耦合,突破了分工很细的学科分类,从总体上对文献进行计量和统计分析,充分发挥了被引文的重要性和价值,也使其成为文献计量学的指导性方法。引文分析法的定量分析具有客观反馈性,弥补了专家定性评价的缺陷。

科学引文索引的发展不仅促成了引文评价的合理化,更重要的是科学引文索引与布拉德福的文献离散传播理论一致,即科学的发现就是思想综合体移动更新的结果。至此,文献计量学的研究主体开始摆脱传统以总结所有文献情报流规律的研究,而是通过深入的规律探查揭示背后隐藏着的科学研究演化规律,较侧重于文献计量学的应用及演化指标的设计研究。

3. 基于文献内容分析的科学图谱

所谓“科学知识图谱”,在图书情报界也被称为知识域可视化或知识领域映射地图,是显示知识发展进程与结构关系的一系列各种不同的图形,用可视化技术描述知识资源及其载体,挖掘、分析、构建、绘制和显示知识及它们之间的相互联系,具有“图”和“谱”的双重性质与特征:既是可视化的知识图形,又是序列化的知识谱系,可对知识单元或知识群体之间存在(或形成)的网络结构及其互动、交叉、衍化等诸多复杂关系进行表达和描述^[16]。

事实上,英国著名的情报学家布鲁克斯(B. C. Brooks)在他的经典著作《情报学基础》中最早提出了“知识地图”的概念,布鲁克斯指出情报学的核心是组织、加工和整理人类客观知识,将知识结构绘制成以各个知识单元概念为节点的学科“认识地图”,通过对文献中的逻辑内容进行分析,找到人类思想的相互影响及联系的结合点,从而为用户提供知识之间关系的一种知识组织的理想状态。有关科学知识图谱的早期研究可以追溯到20世纪60年代。1964年,加菲尔德等人就开始尝试应用引文分析方法研制一个精确有用的、导致特定学科取得重大进展的累积性研究的网络图,以便考察它在科学史和学科结构关系分析等方面的有效性,并以Isaac Asimov博士的《遗传密码》一书作为研究基线,手工绘制完成了DNA研究领域的知识演进图谱^[17];1965年,普莱斯(D. Price)运用类似的引文数据和方法发表了其经典论文《科学论文的网络》(*Networks of Scientific Papers*),文中对物理学和它的一个分支领域的结构进行了分析,普莱斯指出在这个网络图上,有密集分布的小条或小块,如果把这些小条小块研究清楚,就可以绘制当代科学的“地形图”^[18]。这些早期具有开创性的研究活动,对科学知识图谱的后续发展起到了非常巨大的影响和推动作用。

1973年美国情报学家亨利·斯莫尔(Henry Small)^[19]发表了具有里程碑意义的研究成果“Co-citation in the scientific literature: A new measure of the relationship between publications”,首先提出了共引分析方法这一新型的引文分析方法。共引分析的核心是文献共引(co-citation),是指当两篇文献被一篇文献同时参考使用时两篇文献的关系。亨利·斯莫尔(Henry Small)等人开发了基于共引理论的单机系统SCI2Map来描绘科学文献间的结构,亨利·斯莫尔和格里菲斯(Griffith)^[20]以图形方式显示了自然科学所有高度活跃的专业领域中该方法在原有引文分析方

法的基础上,借助国际上新兴的科学知识图谱绘制方法(Mapping of Science)与信息可视化技术(Information Visualization),实现了引证分析的可视化。随后,以亨利·斯莫尔为代表的多位学者利用论文间的共引关系进行了大量实践探索,取得了丰硕成果,共引分析的一个直接应用是为情报学学科划分提供了有利证据,1989年怀特(H. White)和麦凯恩(K. W. McCain)通过共引分析,将情报学分为两个主要领域——计量学(包括引文分析)和情报检索。20世纪90年代的主要进展如下^[21]:1997年美国肯塔基大学的 Linxia 首先尝试将自组织映射技术用于共引矩阵,并在2000年生成了一个将情报学家聚集到几个主题域的图谱;1998年,怀特和麦凯恩再次采用ACA技术,通过1972—1995年24年间的一些代表性的文献特征(作者共引数目),归纳总结情报学领域的结构特征和24年来的发展情况,并作出情报学科的知识图谱。1999年Chaomei Chen把寻径网络(Pathfinder Network, PF-NET)技术引入作者同被引分析,在共引基础上发展出作者同被引文献的度量并生成了有关超文本研究的同被引图谱。2003年同被引分析用于对物理学前沿及发展趋势进行的深入研究,绘制出一系列知识图谱。

在知识图谱基础上,雅恩(M. Jahn)^[22]、戈夫曼(W. Goffman)^[23]以及亨利·斯莫尔的研究进一步指出专业对论文的优先依赖特征,即一个专业是由在发展中出现的为数不多且极其重要的论文所定义。基于大量文献的关联分析带动了知识可视化展现技术的发展,多维尺度分析方法是其中一种标准的可视化技术,也是文献计量学中比较热门的一种空间分析法。它为客观事物的相近性(相似性或距离)提供了一种可视化表达方式。在对引文的多维尺度研究中,多数研究人员,包括斯莫尔、怀特、麦凯恩等人,都是通过将原始引文矩阵转化为相关关系矩阵后,再进行主成分分析,并绘制知识图谱。有时学者也会直接借助原始矩阵进行主成分分析,如C. M. Chen就是采用这种分析得到25个代表不同分支学科的主成分,列出了科学计量学三个主要的分支领域,并绘制出了包括403篇文献的核心文献图^[24]。佩尔松(O. Persson)也揭示了采用可视化方法综合引用与被引用信息对于科学发现的广泛可能性,他以《科学计量学》的引用数据作为研究对象,绘制出了这一期刊的发展过程中最有影响力的文献图^[25]。

进入21世纪以来,知识图谱软件发展很快,知识图谱作为一种揭示科学结构的工具得以广泛应用。典型的研究如2005年Loet Leydesdorff等人运用新型知识图谱分析软件Pajek和传统的多元统计分析方法相结合,将同被引研究延伸到网络领域,为同被引分析开辟了一片新天地,近些年他还致力于研究期刊间的引用关系,通过对期刊间引用频次的可视化分析,图示它们的网络关系,他使用从SCI和SSCI套录的数据,通过计算2000年7349种期刊的引用和被引用频次,利用VxInsight作为知识图谱软件,构造了期刊引用和同被引关系的知识图谱。西班牙

的 Felix Moya-Anegon 收集西班牙在 2000 年被 SCI 和 SSCI 以及 AHCI 收录的文献,根据发表这些文献的期刊在 ISI-JCR 中所属的类别,对应到西班牙分类法的 25 个大类中去,从而得到西班牙大科学结构分析的知识图谱^[24]。

2007 年,美国的陈超美博士通过对研究前沿主题的运算分析,发展了基于引文网络和共引轨迹数据描绘科学发展的新动态的算法,他还研制开发了可用于文献动态识别的可视化软件 CiteSpace I 和 CiteSpace II,其中设计了用于新兴主题识别的跳跃检测算法、潜在范式变化的关键文献节点社群分析算法,以及两个互补视图:聚类视图和时区视图。他进一步总结了研究前沿和知识基础理论发展的过程,如表 1.1 所示。

表 1.1 研究前沿和知识基础的定义进展

作 者	年份	研究 前 沿	知 识 基 础	聚 类
普莱斯	1965	对于一篇引文,由频繁引用的近期文章所构成的动态聚类	未定义	引文的最近行为
斯莫尔和格里菲斯	1974	共引聚类		共引
布拉姆	1991		共引文献	词集
加菲尔德	1991	集中关注的一系列相关问题和概念	共引文献	
佩尔松	1994	共引聚类与引文总合	研究前沿的引文映像	共引
莫里斯等	2003	引用相同文献的文章	固定与时间无关的文章群	文献耦合
CiteSpace II	2007	正在兴起的理论趋势和新主题的涌现	共引网络	共被引文章和引用文章

自文献计量学形成以来,知识模型大致经历了经典定律、引文评价再到内容解析三个发展过程,为不同层次科学领域的研究人员提供有效的信息检索服务一直是知识模型发展的根本目标。这些理论的思想基础就是文献计量学中的文献评价理论和知识传播理论。有关文献的评价分析、科学结构与演进分析一直是文献计量与情报学最主要的研究内容。

综上所述,三个时期最主要的变化是知识载体的改变。在文献计量学的早期,经验规律是通过文献这一层知识载体的统计规律进行调查归纳得出的,随着引文评价分析的深入,在宏观应用领域形成了更高一层的应用,例如期刊、学术机构、学科等综合评价。其中基于引文数据库的学术影响因子被视为文献计量学史上的一次革命,已被科学界广泛接受并成为理工科学术成果评价的标准,继而在科研管