

新中国成立70年来数字出版的学术演进

邱立楠

摘要:为系统考察数字出版研究70年的学术场域,文章以国内70年来收录于中国知网(CNKI)中CSSCI期刊的1291篇学术文章为研究样本,运用CiteSpace5.3R8软件版本,对数字出版研究70年来的研究文献进行了共词分析和共现分析,包括作者合作共现图谱分析、核心作者分析、高被引论文分析、数字出版研究的热点呈现、突变词演进分析等。研究发现,我国的数字出版研究在学理层面已经广泛涉及到数字出版产业、数字出版平台、版权保护、数字出版人才、电子书、商业模式、出版转型等诸多学术场域。该研究也作出展望,未来的数字出版研究能够在基金资助项目正态分布、学术研究期刊发布多元和学术研究学科交叉方面有进一步的突破。

关键词:数字出版; CiteSpace; 知识图谱; 学术场域

作者简介:邱立楠,男,讲师,博士研究生。(江苏海洋大学 文学院,江苏 连云港,222005)

中图分类号:G239.29 **文献标识码:**A **文章编号:**2096-8418(2020)03-0048-13

数字出版的有关概念伴随新中国成立而不断发展和演进,其自20世纪50年代萌生以来,历经桌面出版、光盘出版、多媒体出版、在线出版、移动出版、手机出版、跨媒体出版等不同时期阶段,构成了70年来数字出版的发展历史轨迹。当前,国家正在大力推进以数字化和自主知识产权为核心的科技兴业战略,新闻出版业“十三五”时期发展规划也指出,数字出版已成为新闻出版业的第二大产业,并对新闻出版业实现结构调整和提质增效、加快转型升级和促进融合发展起到了重要的促进作用。在第八届中国数字出版博览会上,中共中央宣传部副部长庄荣文在主旨报告中也强调,“数字出版作为文化建设的重要方面,在推进数字中国和网络强国建设中日益发挥着重要作用,正在成为建设社会主义文化强国的新动能。”^[1]

在数字出版的学术演进方面,学者赖茂生发表的《从电子出版到数字出版》一文,被认为是国内最早论及“数字出版”的学术成果,其在思考电子出版业未来发展前景时,提及了“出版数字化和按需出版”,^[2]但学者侯欣洁认为,其“只涉及技术发展方向,并未对数字出版做出界定”,^[3]并非严格意义上的关于数字出版的学术研究。令人欣喜的是,在以“互联互通,共建共享”为主题的首届中国数字出版博览会在北京开幕后,我国数字出版研究在学理层面开始步入快车道。

一、研究设计

知识单元离散与重组理论是由国内科学计量学家赵红州所提出,他认为任何科学都有其自身的创造过程,都是先把结晶的知识单元游离出来,然后再在全新的思维势场上重新结晶的过程。库恩的科学发展模式理论也认为,科学发展是科学革命的历史过程。^[4]本文基于上述理论基础,通过题名搜索,确定研究对象和样本数据,再运用可视化软件,分析数字出版研究70年来的学术场域。

(一) 研究的方法和软件选择

本课题主要采用文献计量法的研究方法。文献计量法是以学术文献为研究对象,从定量的角度分

析文献规律。^[5]文献计量学可通过运用多种软件加以实现，如以基于建构社会关系矩阵分析的 BibExcel、BICOMS、SAT 等软件、基于社会网络聚类分析的 CiteSpace、VOSviewer、Ucinet、NetDraw 等软件。本文采用的是 CiteSpace 软件。其采用共引分析理论和寻径网络算法，可对特定领域文献进行计量，探寻科学领域演化的关键路径及知识转折点，进而完成对学科发展前沿的探测，已经成为科学计量学普遍采用的新工具。^[6]

（二）样本数据的确定和说明

笔者以“数字出版”为主题在中国知网（CNKI）上进行文献检索（文献检索时间为2019年9月2日，下同），得出11797条有效文献；以“数字出版”为关键词搜索，得出5739条有效文献，以“数字出版”为篇名进行精确检索，共得出4545条有效文献。由于CiteSpace软件对报刊文献无法进行有效分析，加之基于文献质量的考量，笔者重新以“数字出版”为篇名，且在期刊来源类别中只选择了CSSCI来源期刊，设置检索条件：（CSSCI期刊=Y）并且（题名=数字出版）（精确匹配），得出1296条有效文献。笔者基于对本课题数字出版研究年度完整考量的需要，故而未统计知网已收录2019年的5篇文献。需要补充说明的是，由于各期刊杂志社录入知网的时间顺序略有差别，笔者以文献实际登载的刊期时间为准。故而，发表在《湖北社会科学》的《中国数字出版产业研究热点及趋势分析》一文，中国知网虽在2019年1月7日才收录，但该文实际登载于2018年12期，故而笔者将该文计入本课题的样本之列。故而本课题最终得到的实际有效文献为1291篇。

（三）样本数据的采集与处理

囿于中国知网文献导出最大篇幅500篇的限制，笔者将在CNKI检索到的1291篇样本分三次导出，导出选择Refworks格式。随后，将导出的文件依次命名为“download_500”“download_1000”“download_1291”。由于CiteSpace软件不能直接识别Refworks格式，所以笔者再次将上述导出的三个文件经过软件完成import/export的数据格式转换。

二、新中国成立70年来数字出版研究的知识图谱分析

（一）作者合作共现图谱分析

对数字出版领域核心作者的关注是本课题的一个中心。笔者运用CiteSpace软件，在Node Types中仅选择Author。由于本文研究的样本数据较大，所以笔者在Pruning剪枝路径选择中选取Pathfinder寻径、Pruning sliced networks修剪切片网络、Pruning the merged network修剪合并网络选项，得到图1。在生成的图谱中，节点的大小表示作者发表论文的数量，之间的连线反映合作关系强度。^{[4](78)}从中，我们可清晰发现，数字出版研究领域网络节点数量 $N=127$ ，连线数量 $E=82$ ，网络密度为0.0102，较为稀疏，表明了作者之间的合作关系较低，作者之间偶尔存在单次合作，还未形成固定的团队化运作模式，也在一定程度上说明了人文社科研究在团队建设方面还有待增强。图1的知识图谱中还显示，有两个合作子网络最为明显，分别为黄先蓉和汤雪梅等为核心的数字出版研究团队。张海新、余庆、杨庆国、杨方铭次之。

首先，黄先蓉和郝婷联系紧密，形成了最明显的合作节点，两人先后合作发表研究5篇。其次，汤雪梅的关系网最为复杂，源自其在2013年发表了《中国数字出版产业发展的规模、态势及趋势分析》一文，该文有强大的作者团队，署名为2012-2013中国数字出版产业年度报告课题组，个人署名依次为郝振省、魏玉山、张立、王飏、毛文思、李广宇、李熙、郝园园、李游、汤雪梅和杨涛等，共计11人。除上述两个团队外，其余作者在知识图谱中零散分布，表明尚未形成成熟的合作关系。

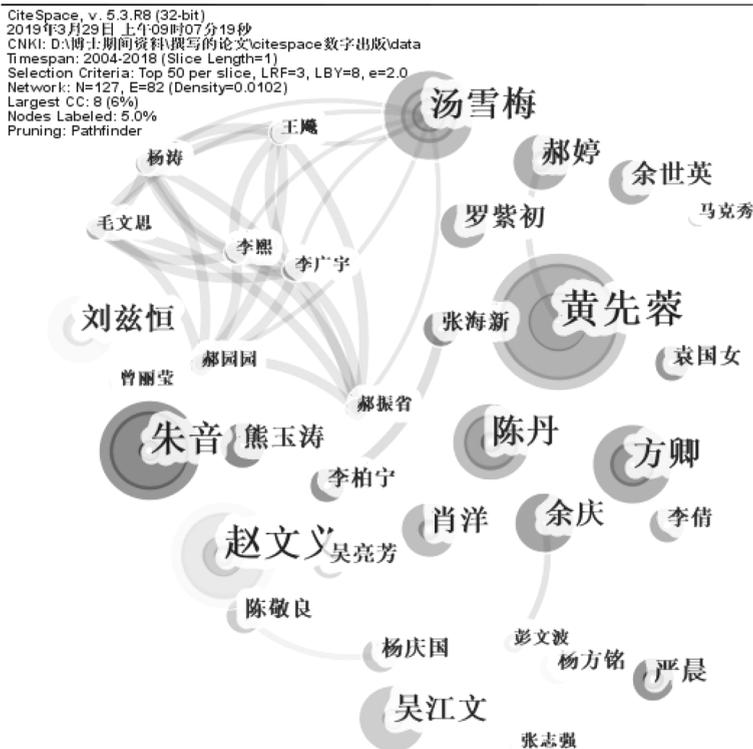


图1 作者合作共现图谱

(二) 核心作者分析

洛特卡和普赖斯根据对科学家的生产率 and 活动规律, 提出了核心作者这一概念, 并认为发表论文数为 N 篇以上的作者为核心作者, 计算公式为 $M=0.749(N_{\max})^{1/27}$, N_{\max} 为样本数据库中最高产作者的论文数, M 即为所求得的核心作者论文篇数, 如若该作者发文数大于 M 值, 则该作者为该领域的核心作者。根据笔者数据统计, 最高产作者为武汉大学信息管理学院的黄先蓉教授, 发文 12 篇, 故 $M=0.749(12)^{1/27}$, 进一步求得 $M=2.5946$, 所以样本数据库中文献研究达到 3 篇或以上的作者即可认定为数字出版领域的核心作者。笔者将 31 位核心作者进行了归纳和整理 (见表 1)。根据 CiteSpace 软件 Export 导出的 Network Summary Table 显示, 在数字出版研究领域, 共有 127 位重要作者, 根据上述分析又得出核心作者 31 位, 核心作者占比 24.4%。其中发文量在 4 篇及 4 篇以上的作者共有 19 位, 占重要作者比例为 15.0%, 发文数量超过 10 篇以上的有 2 位, 占重要作者比例为 1.6%。

经笔者进一步分析, 31 位核心作者共发表文献研究 150 篇, 占样本总数的 11.6%, 与洛特卡和普赖斯理论核心作者要达到样本总数的 50% 有较大的差距, 这也在一定程度上表明了数字出版研究领域是众多学者普遍关注的领域, 但深耕细作高产的学者有待提高, 数字出版研究领域的核心作者群也有待形成, 这也补充说明了图 1 核心作者合作共现图谱所反映出来作者之间研究分散、合作较少等系列问题。

表1 数字出版研究核心作者表

序号	作者	发文量	序号	作者	发文量
1	黄先蓉	12	17	刘成勇	4
2	赵文义	10	18	张立	4
3	朱音	9	19	严晨	4

序号	作者	发文量	序号	作者	发文量
4	汤雪梅	8	20	杨庆国	3
5	陈丹	7	21	杨方铭	3
6	方卿	7	22	石雄	3
7	刘兹恒	6	23	张海新	3
8	吴江文	6	24	张小强	3
9	杨晓芳	6	25	袁国女	3
10	郝婷	5	26	陈敬良	3
11	肖洋	5	27	李倩	3
12	余庆	5	28	吴亮芳	3
13	罗紫初	4	29	李华	3
14	柳斌杰	4	30	王巧林	3
15	熊玉涛	4	31	李柏宁	3
16	余世英	4			

(三) 高被引论文分析

在一个学术场域内，学者论文的被引数量也是衡量该学者研究影响力的重要指标。笔者从样本数据中选择被引用排名前20的文献予以分析，具体情况见表2。根据统计，20篇文章共被引用1554次，平均77.7次。其中只有前7名被引数高于平均值，8-20名被引数低于平均值，说明其在数字出版学术场域中被采纳、认可的程度更高，为数字出版研究的学术演进做出了重要贡献。其中，被引数超过100次的有四篇，共被引529次，占比34.0%。从学者角度来看，学者张立有三篇文章进入榜单，分别位列第3、13、14位，引用数共计245次，占比15.8%；程维红、任胜利、王应宽合作发文2篇，引用数共计222次，占比14.3%。其中，程维红、任胜利、路文如、严谨和王应宽的《我国科技期刊由传统出版向数字出版转型的对策建议》被引数最高，达到154次，该文系统综述数字出版概念的提出及演变、数字出版优势、数字出版的产业链，分析了我国数字出版的发展现状和存在的问题，并提出了8条有针对性的数字出版策略。徐丽芳的《数字出版：概念与形态》一文发表时间是2005年第7期，是入选20名高被引论文中发表最早的一篇，介绍了数字出版的概念与内涵、数字出版物类型与数字出版形态、互动数字服务业的数字出版情况，具有很高的参考价值，也获得了135次的被引次数，位居第二。

表2 数字出版研究20篇高被引论文表

序号	篇名	作者	刊名	刊期	被引数
1	我国科技期刊由传统出版向数字出版转型的对策建议	程维红；任胜利；路文如； 严谨；王应宽	中国科技期刊研究	2011(04)	154
2	数字出版：概念与形态	徐丽芳	出版发行研究	2005(07)	135
3	数字出版相关概念的比较分析	张立	中国出版	2006(12)	130
4	数字出版：距离成熟还有长路要走	聂震宁	出版科学	2009(01)	110
5	论数字出版产业链的整合	刘灿姣；黄立雄	中国出版	2009(01)	96

序号	篇 名	作 者	刊 名	刊 期	被引数
6	大数据时代数字出版产业的发展趋势	孙玉玲	出版发行研究	2013 (04)	90
7	优先数字出版是提高学术期刊出版速度的一种新模式	汪新红	中国科技期刊研究	2011 (01)	87
8	数字出版即全媒体出版论——对“数字出版”概念生成语境的一种分析	张大伟	新闻大学	2010 (01)	75
9	中国科协科技期刊数字出版及传播力建设	程维红; 任胜利; 沈锡宾; 方 梅; 王应宽	中国科技期刊研究	2014 (03)	68
10	传统出版社向数字出版跨越的三条路径	陈超英	出版发行研究	2010 (07)	68
11	数字出版的发展现状及我国科技期刊的应对措施	张维; 邓强庭; 冷怀明	编辑学报	2013 (02)	62
12	国外出版商发展数字出版的特点及给我们的启示	秦绪军	科技与出版	2007 (12)	61
13	我国数字出版产业的发展趋势及对策分析	张立	出版发行研究	2008 (10)	59
14	数字出版的若干问题讨论	张立	出版发行研究	2005 (07)	56
15	传统出版该如何应对数字出版的挑战	祁庭林	编辑之友	2007 (04)	55
16	数字出版的商业模式与传统出版企业的数字出版发展	汪忠	出版发行研究	2008 (08)	52
17	数字出版环境下科技期刊编辑职能、工作方式转变及自我调整策略	顾凯; 邹栩	中国科技期刊研究	2012 (03)	50
18	我国数字出版产业链整合模式分析	周利荣	出版发行研究	2010 (10)	49
19	数字出版: 新的革命	傅强	浙江大学学报 (人文社会科学版)	2008 (04)	49
20	我国科技期刊数字出版的 SWOT 分析	吴月红	编辑学报	2010 (03)	48

从论文发表期刊来看,在这 20 篇高被引论文中,《出版发行研究》有 7 篇文章入选,分别位列 2、6、10、13、14、16、18 位,被引 509 次,占比 32.8%。《中国科技期刊研究》有 4 篇文章入选,分别位列 1、7、9、17 位,被引 359 次,占比 23.1%。两份期刊共有 11 篇文章入选,被引 868 次,占比 55.9%。其次,《中国出版》和《编辑学报》都有 2 篇文章入选,分别被引 226 次和 110 次,占比

14.5%和 7.1%。根据表格不难发现，这前 20 名论文共发表在中国科技期刊研究等 9 份刊物之上。具体权重见图 2。

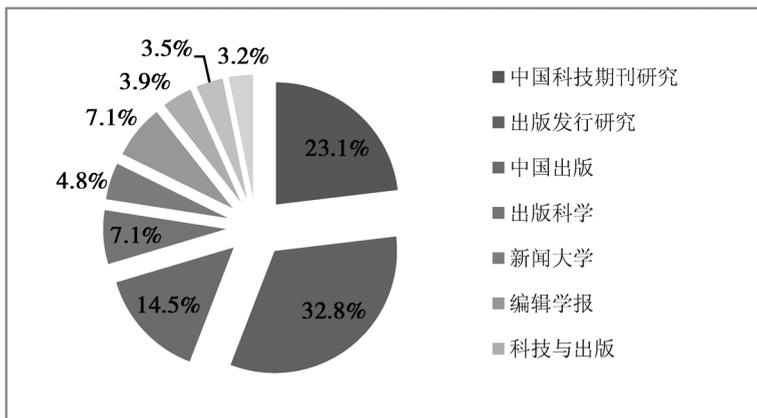


图 2 数字出版研究 20 篇高被引论文期刊分布图

(四) 数字出版研究的热点呈现

1. 关键词共现知识图谱分析

在 CiteSpace 软件中，其他设置保持不变，将 Node Types 选取 Keyword，生成关键词共现知识图谱。设 Threshold=3 后，对各阈值进行了最佳化的调整，形成图 3。从图 3 中得到，关键词共现知识图谱得到 260 个关键节点，362 个连线，密度为 0.0108。



图 3 数字出版研究关键词共现图谱

为了更好的分析，笔者将“Summary Table”中的前 20 位数据制成表 3，即数字出版研究高频中心性词表。关键词中心性在整个共现网络关系中中介者的能力强度。传统认为，如果一个关键词的中心

性指标高于 0.1 的节点, 那表明其在网络结构中扮演着的角色就相对较为重要。从图 3 中不难发现, 处在关键词共现图谱重要位置的是数字出版、中华人民共和国、数字出版产业、出版, 而表 3 中与其对应的关键词中心性分别为 0.57、0.52、0.4、0.45, 恰好也包揽了关键词中心性指标的前四强, 可见上述议题一直是数字出版 70 年来学术所重点探讨的。此外, 数字出版、数字出版产业、出版社等传统强势研究的内容一直保持强势, 关键词出现的频率位居前三, 其中数字出版产业更是获得 947 频次, 其最早的一篇文献于 2005 年发表, 与其同一年出现的还有位居第 19 名的博览会, 因为在 2005 年 7 月 8 日, 我国举办了首届中国数字出版博览会, 自此开启了我国在数字出版领域研究的热潮。

表 3 数字出版研究高频与高中心性词表

序号	关键词	频次	中心性	最早研究时间
1	数字出版	947	0.57	2005
2	数字出版产业	88	0.4	2006
3	出版社	63	0.19	2006
4	文化机构	54	0.03	2006
5	中华人民共和国	49	0.52	2005
6	传统出版	44	0.09	2008
7	学术期刊	44	0.12	2008
8	科技期刊	42	0.12	2009
9	digital publishing	38	0.16	2014
10	数字出版产业链	36	0.31	2006
11	出版	36	0.45	2006
12	商业模式	31	0.02	2009
13	版权保护	26	0.03	2011
14	数字出版企业	25	0.05	2010
15	出版业	24	0.03	2006
16	优先数字出版	24	0.06	2011
17	人才培养	24	0.09	2011
18	出版物	23	0.08	2007
19	博览会	22	0.08	2005
20	数字出版平台	21	0.13	2008

2. 主题词关键词聚类知识图谱分析

CiteSpace 软件的魅力不仅仅体现在上述关键词的共现图谱分析上, 笔者在软件中完成对主题词和关键词的深度聚类后, 采用时区图谱和时序, 以分析数字出版的研究热点和研究前沿, 进而将呈现明显特征的词完成聚类, 寻求数字出版研究 70 年历程中的研究热点学术演进。这里需要说明的是, CiteSpace 软件依据网络结构和聚类的清晰度, 给出了两个特别需要留意的指数, 是我们评判图谱绘制效果的依据。这两个指标其一是 Modularity Q 指标, 指标区间一般在 $[0, 1)$ 内,^{[6](24)}且 Modularity Q 指标 >0.3 才意味着划分出来的社团结构是显著的。另一个指标即为 Mean Silhouette 指数, 其为平均轮廓值, 用来衡量网络同质性, 故而要求其指标应该大于 0.5, 当 Mean Silhouette 指数在 0.7 时, 聚类结果

值得信服。只有满足上述两个指标我们才认为此次聚类是合理的。笔者在完成聚类之后，发现本次聚类得到的 Modularity $Q=0.812$, Mean Silhouette $=0.7413$, 满足上述指标要求，而且 Mean Silhouette 指标大于 0.7, 可见本次聚类的结果有很高的参考价值。具体聚类结果见图 4。

从图 4 中我们可以得出，排名第一的是数字出版，其虽在 2004 年首次被提出，但直至 2011 年才开始进入热烈讨论和研究的阶段，2011 年的研究热点主要集中在内容管理、公共平台、产业创新，2012 年的研究热点是国家数字出版基地、法律制度、知识图谱，2013-2016 的研究热点是系统优化、商业属性、产业融合、技术驱动、全民阅读、众包等，到 2017-2018 年，数字出版研究的热点开始转向内容为王、知乎书店、社群、数字编辑等方向。排名第二的是数字出版权，这一概念的提出稍晚于数字出版，其研究经历了数字出版产业链、电子书、盈利模式、专业出版、内容资源、教育出版、数字化新闻出版单位、“互联网+”数字阅读、实现路径、知识付费等研究热点的演变。排名第三的是版权保护问题，在第一阶段的 2004-2009 年，热词更多是在出版科学、数字出版产业、年度报告等方面演变，但在第二阶段的 2010-2014 年，热词开始聚焦数字版权保护、法律保护、版权保护、数字化转型、按需印刷、产业组织理论等现实困境；在第三阶段的 2015-2018 年，其热词更多聚焦区块链、共享经济、运营模式方面。在图 4 中，处在表格后三位的 digital publishing、优先数字出版、“2010”也需要我们额外加以关注。首先是 digital publishing，其尽管排名未进入前十，但根据表 3 可以发现，从 2014 年左右学术研究开始普遍使用“digital publishing”这一概念，五年的时间里还伴随着学术出版、library、数字图书馆、出版融合、地理学中文核心期刊、高校图书馆、知识服务、knowledge service、学术图书馆等热词演变，并快速跻身关键词聚类知识图谱的前列，并将持续成为未来学术研究的热点之一。而优先数字出版则开始于 2011 年左右，伴随对策、优先数字出版、人才培养、调查、数字产品、publication delay、全产业链的热词演变。关键词聚类“2010”一直拥有顽强的生命，其从 2005 年便一直频频成为研究的热词，其在 2005 年的热词即是博览会和中华人民共和国，这也在图 3 数字出版研究关键词共现图谱明显找到，在表 3 数字出版研究高频与高中心性词表中得到验证，博览会和中华人民共和国分别位列表 3 中的第 19 位和第 5 位，出现频率分别为 22 次和 49 次，中心性为 0.08 和 0.52。

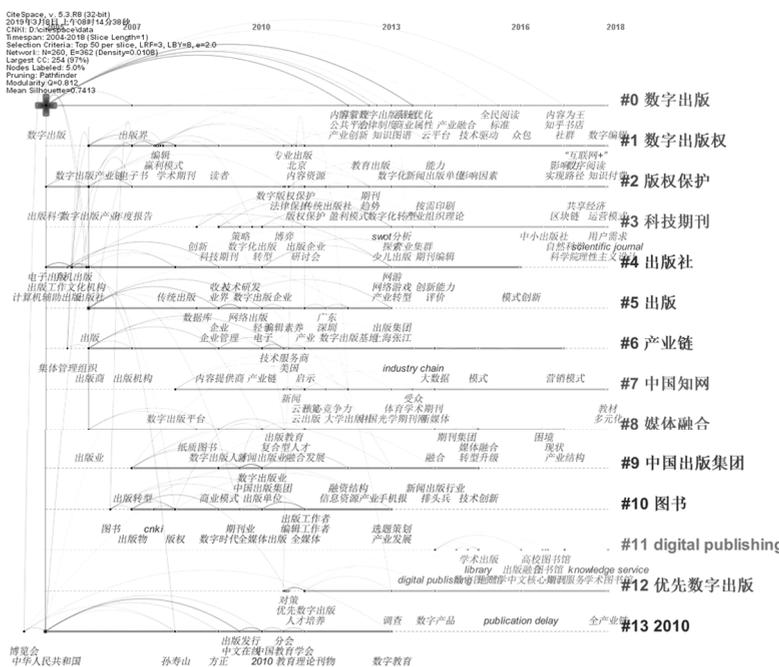


图 4 数字出版研究主题词关键词聚类知识图谱表

3. 研究机构时区视图的知识图谱分析

在 CiteSpace 软件中, 其他设置保持不变, 将 Node Types 选取 Institution, 生成研究机构的合作知识图谱, 然后选择 Timezone View, 生成数字出版机构的时区视图 (见图 5)。时区视图是另一种侧重于从时间维度上来表示知识演进的视图。时区视图可以清晰地展示出文献的更新和相互影响。^{[6](76)} 根据图 5, 数字出版的知识图谱节点数 105, 连线数 24, 网络密度 0.0044, 研究机构之间的合作关系非常低。从时间演进的角度, 北京印刷学院和武汉大学信息管理学院出现的时间最早、年轮最大。从年轮圆环的颜色分析, 这些研究机构近年来在数字出版研究领域仍然成果不断, 表现出较强的研究态势。图 5 还显示, 上海理工大学出版印刷与艺术设计学院、四川大学文学与新闻学院、内蒙古大学文学与新闻传播学院等三所研究机构在 2010 年齐发力, 形成了叠加状态, 数字出版研究开始逐渐走向研究的繁荣。近年来, 地质出版社和长安大学文献资源研究所也开始加入数字出版研究的大军并崭露头角, 形成了一定的科研成果。从合作连线视角分析, 研究机构之间的合作关系并不频繁, 仅有武汉大学、四川大学、南京大学等研究机构有少许的合作关系。此外, 数字研究领域的马太效应也逐渐凸显, 北京印刷学院和武汉大学信息管理学院因为起步较早的缘故, 随着科研成果的不断涌现, 更加凸显了其绝对的研究强势地位。

笔者根据统计, 将结果显示排名前 14 名的机构做成数字出版研究高产研究机构表 (见表 4)。14 家研究机构共发文 206 篇, 占研究样本的 16.0%; 武汉大学信息管理学院发文 56 篇, 占研究样本的 4.3%, 占排名前 14 名机构样本总数的 27.2%, 表明了其在数字出版研究方面拥有绝对的领先地位。北京印刷学院是图 5 中所有机构中最早涉足数字出版领域的机构, 其最早发表该领域的论文是在 2007 年, 经过 12 年的长足发展, 共发表论文 35 篇, 位居榜单第二, 占样本总数的 2.7%, 占排名前 14 名机构样本总数的 17.0%。

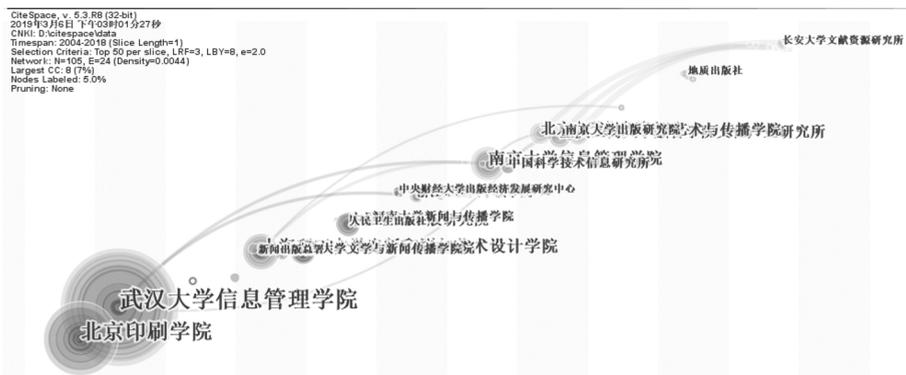


图 5 数字出版研究机构时区视图知识图谱

表 4 数字出版研究高产研究机构表

序号	单位	数量	最早研究时间
1	武汉大学信息管理学院	56	2008
2	北京印刷学院	35	2007
3	上海理工大学出版印刷与艺术设计学院	17	2010
4	南京大学信息管理学院	16	2013
5	四川大学文学与新闻学院	11	2010

序号	单 位	数量	最早研究时间
6	中国新闻出版研究院	11	2011
7	长安大学文学艺术与传播学院	10	2014
8	重庆工商大学融智学院文化产业研究所	10	2014
9	北京大学信息管理系	10	2014
10	浙江工业大学法学院	6	2012
11	河南大学新闻与传播学院	6	2011
12	南京大学出版研究院	6	2014
13	内蒙古大学文学与新闻传播学院	6	2010
14	中国科学技术信息研究所	6	2013

三、新中国成立70年来数字出版研究的突变词演进

数字出版70年来突变词的演进是反映数字出版研究热点的核心关键，而突变词是指在某些年份发表的文献中骤增的专业术语，适合表征研究前沿。而对突变词的分析主要就是为了准确把握在什么时候开始骤然出现大量的专业术语，进而分析本学科研究的热点演变情况。笔者在基于关键词热点分析的基础上，通过分析主题词的突变特征，梳理了数字出版研究的热点（见图6）。突变词的突变特征可以从两个方面得以体现。其一，Strength，即为突变的强度，其二 Begin-End，即为持续的突变时间。根据图6，关于数字出版研究的突变主要分为三个阶段：

第一个阶段主要集中在2006年以前。在此时期，学者们主要对中国数字出版博览会、出版科学、数字出版产业、出版社、出版商、文化机构等开始了起步研究。

根据统计，本时期研究样本只有32篇，占样本总数的2.5%，但该阶段出现的重点突变词数达10个，占遴选重点突变词样本总数的47.6%，这些突变词成为数字出版研究的基底。具体而言，2005年首届中国数字出版博览会的召开，有力推动了关于数字出版的研究，并迅速刮起学界对数字出版研究的热潮，这也使“博览会”“中华人民共和国”双双入选重点突变词之列。该时期关于博览会的样本有6篇，占该时期样本总数的18.9%。如《数字出版产业链急需整合——首届中国数字出版博览会将在京举办》《首届数博会为数字出版产业发展打开新局面》等相关研究都试图为中国数字出版行业面临的问题寻找解决对策，并对全面促进传统出版业的数字化转型进行了多方位的探讨。此外，由于数字出版作为一个新事物，该时期还有大量对其概念、形式、问题的研究，如相继出现了《数字出版的若干问题讨论》《数字出版：概念与形态》《何为数字出版》《数字出版相关概念的比较分析》等学术成果。在该时期，我们发现有两个突变词的指数超过10，即为“出版社”的14.0044和“文化机构”的13.8831。Strength表示强度，数值越大，强度越大。红星电子音像出版社的刘旭英从业界现实困境角度对数字出版的问题进行了思考，其《对数字出版存在问题的思考》中提到“尽管数字出版前景很好，然而目前却存在产业链很不完善、版权体制不健全以及技术上的一些问题。”^[8]上述问题也是出版社、出版业、文化机构等业界面临的实际困境。值得庆幸的是，出版社和文化机构的突变时间只有短暂的四年，即2006-2009年，其突变强度大、持续时间短，说明该类词是某时期突发性的研究热点。突发

性强度大,说明出版社和文化机构由具有影响力的现实因素引发。持续时间短,说明出版社和文化机构是过渡性热词,随着事态的发展和研究的深入,他们就融入、转移到其他研究热点之中。

第二个阶段主要集中在2008-2012年。基于第一阶段研究的积淀,在此时期,学者开始将目标转向盈利模式、数字出版模式、商业模式、全媒体出版、数字化出版、版权保护等热点上来。根据笔者统计,本时期研究样本有493篇,占样本总数的38.2%,较第一阶段有了跨越式的增长。本时期突变词量有9个,占遴选重点突变词样本的42.9%,与上一阶段略低5个百分点。具体而言,本阶段突变词的强度指数普遍偏低,最高的即为“赢利模式”,强度仅为6.828,且持续时间仅为2008-2010年,跨度也只有三年,说明其基于现实的出版困境和数字出版转型路径的双重压力,学界和业界对出版赢利问题进行了多重的思考。赢利模式设计是传统出版社进军数字出版领域需要解决的主要障碍之一,^[9]如学者熊玉芳在《简谈数字出版的赢利模式及发展》中就谈到数字出版的商业运营模式,“应顺应消费者的心理需求和支付意愿,进行赢利模式的创新,并需着力解决加强产品版权保护、协调产业利益分配、提高用户支付意愿等问题”。^[10]赢利和商业模式的问题是传统出版产业转型升级后必须面对的实际问题,这也成为2008-2009年间对数字出版研究的主要学术场域。从图6中还发现,在该阶段后半期的2010-2012年,关于数字出版研究的学术场域开始发生了突变和转向。2010年,中国数字出版年会在北京召开,大会的主题是“2010:推进数字出版跨越式发展”。在图6中也有所体现,“2010”入选了数字出版研究70年的重点突变词,强度为4.0513,持续时间最短,从2010-2011年。以此为节点,数字出版的学术场域开始从单一探究赢利模式和商业模式向全媒体出版、数字化出版、数字化出版基地、盈利模式、版权保护等领域突变扩展。重点突变词的更迭转向,反映了推动数字出版跨越式发展是我国出版业的战略选择,以新一轮技术革命为支撑,以数字出版为代表的新业态已经成为世界出版大国强国的战略选择。在全媒体出版方面,中国青年出版社总社万玉云谈到,“传统出版社应完善基于全媒体的数字出版产品形态,搭建网站服务运营平台,跨界整合营运渠道,实现传统出版向全媒体数字出版的转型,增强自己的核心竞争力”。^[11]在版权保护方面,众多学者从数字出版的版权保护策略、版权保护体系构建与完善、版权保护实践中面临的法律问题等角度进行了新的思考,认为“无论是法律层面、产业层面还是技术层面,都应积极调整和应对,以适应新形势下版权保护的新特点。”^[12]

第三个阶段主要集中在2014-2018年。在此时期,学者们的学术场域突变词主要有两个,即为“大数据”和“digital publishing”,突变强度分别为5.9639和14.9193,突变时间始于2014-2015年,并突变至今。其中digital publishing的裂变强度为21个重点突变词强度之首,也是当下研究的热点和前沿,涉及智能化出版、数字出版信息资源的利用开发、国外出版社数字出版转型、编辑工作转型、人才培养、数字出版与知识服务协同共赢、运作模式等。学林出版社的吴耀根根据实践经验,认为“借助知识服务的专业化,出版集团实施归核化发展战略,是目前转型发展的必由之路”。^[13]学者于成和王林就把国外大学出版社数字出版转型的经验介绍到国内,“依托内容优势、后发技术优势以及深刻的学术出版理念,剑桥社生成了整合式学术数字出版平台Cambridge Core”。^[14]在本阶段,另一个突变词是“大数据”,强度值为5.9636,突变时间从2014年至今。2000年以后,由于“互联网特别是后来移动互联网的出现,数据量不仅剧增,而且开始互相关联,出现了大数据的概念”。^[15]有学者从大数据分析技术与数字出版应用结合角度研究,如中央财经大学出版经济研究中心的祝兴平就对大数据分析技术与数字出版这一战略性新兴产业的关系,大数据技术在数字出版产业中的应用进行了探讨。还有学者探讨了大数据技术与数字出版模式,从“数字出版内容、流程、服务三方面分析了大数据技术给数字

出版模式带来的变革”。^[16]

Top 21 Keywords with the Strongest Citation Bursts

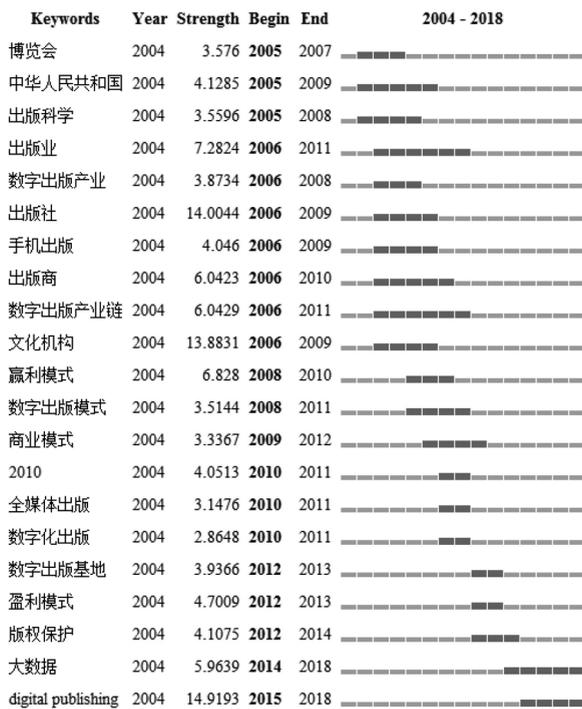


图6 数字出版研究重点突变词演变

四、研究展望

1949年，新中国成立后，我国的出版行业在中国共产党的引领下，从一个出版薄弱之国，发展成为一个出版强国。在新技术的革新与实践运用中，传统的“铅与火”的软质出版开始向“光与电”的数字出版转向，我国的数字出版享受到国家发展的红利，实现了弯道超车。70年来的自强发展，我国的数字出版研究在学理层面已经广泛涉及到数字出版产业、数字出版平台、版权保护、数字出版人才、电子书、商业模式、出版转型等诸多学术场域。但在未来的学术研究中，仍然需要在以下方面予以更多重视。

（一）期待基金资助项目正态分布化

良序的学术研究离不开学术基金的助力和护航，在数字出版研究的70年历程中，学术演进获得基金资助数呈现一种良好的增长态势，从2008年开始实现零的突破，并于2017年达到获得资助28篇的高峰，并实现164篇学术研究得到基金项目的资助支持。但是遗憾的是，根据记者的梳理和分析发现，这164篇获得基金资助的文献分散在29种不同的基金项目，但是“二八规律”明显，没有呈现正态分布。其中，国家社会科学基金项目独占鳌头、一枝独秀，共有82篇文献获得其资助，刚好占该样本总数的50%。其次，国家自然科学基金资助了14篇文献，占比8.5%，随后是湖南省社会科学基金和国家科技支撑计划，分别资助了9篇和8篇文献，占比5.5%和4.9%。总体而言，基金项目类型较为丰富多样，但长尾效应明显，有17种基金只资助了2篇以下（含2篇）的文献，占29种基金类型的58.7%；有10种基金仅资助了1篇文献，占基金类型的34.5%。

（二）注重学术研究期刊发布多元化

总体而言，数字出版研究的发表期刊过于集中，主要分为三个梯队。第一梯队由《出版发行研究》《科技与出版》《中国出版》三家期刊所组成，这三家期刊全部都是出版行业的学术性刊物，以刊载出版实践前沿与实践经验、出版学理论研究、出版活动规律等方面见长，因而该梯队十五年来分别发布文献 219 篇、218 篇、216 篇，共计发文 653 篇，占样本总数的 50.6%。第二梯队由《编辑之友》《出版科学》《现代出版》等 9 家期刊组成，共计发文 457 篇，占样本总数的 35.4%。第三梯队由《重庆社会科学》《新闻界》《图书馆理论与实践》等期刊组成，共计发文 181 篇，占样本总数的 14.0%。学术研究期刊发布多元化，有利于学术研究成果的丰富和多元，更有助于数字出版研究成为新闻传播学研究领域的一个核心领域，助益数字出版学术研究的突变和不断发展。

（三）强化学术研究的学科交叉化

根据样本数据分析结果而言，数字出版研究涉及到新闻传播学、图书情报档案、法学、教育、计算机等 20 多个学科领域，但以新闻传播学领域研究最为火热，样本数达 1035，占样本总数的 80.2%。其次分别为图书情报档案和法学，样本数分别为 48 和 46，分别占样本总数的 3.7% 和 3.6%。但是，尽管“数字出版”这个词本意义上与印刷相关，但是随着社会经济、文化和技术的不断革新和发展，学者们也逐渐意识到“出版方式也是一种文明形式，是传播人类社会文明的载体”。^[17]从研究的样本和学术演进的变迁中也不难发现，数字出版的学术研究也经历了数字化、碎片化和体系化的发展历程，在研究和探讨上逐渐呈现出多学科的研究视角，体现了数字出版学术研究的多元性，表明数字出版研究并不是某一个学科的问题，还需要强化多学科交叉研究。

参考文献：

- [1] 庄荣文. 奋进新时代 展现新作为 [J]. 出版参考, 2018 (8).
- [2] 赖茂生. 从电子出版到数字出版 [J]. 中国电子出版, 2000 (2).
- [3] 侯欣洁. 数字出版概念界定的再认识 [J]. 现代出版, 2014 (5).
- [4] 李杰, 陈超美. CiteSpace: 科技文本挖掘及可视化 [M]. 北京: 首都经济贸易大学出版社, 2016: 2.
- [5] 谢仁恩等. 国内教育领域 PBL 研究的定量分析 [J]. 现代教育技术, 2009 (2).
- [6] 陈悦, 陈超美等. 引文空间分析原理与应用 [M]. 北京: 科技出版社, 2014: 6.
- [7] 史书侠, 李海亮, 杨华. 2005-2010 年《情报科学》刊出论文作者分析 [J]. 情报科学, 2012 (2).
- [8] 刘旭英. 对数字出版存在问题的思考 [J]. 江西社会科学, 2006 (9).
- [9] 方卿, 许洁. 数字出版赢利模式设计的五要素 [J]. 出版发行研究, 2009 (11).
- [10] 熊玉芳. 简谈数字出版的赢利模式及发展 [J]. 出版发行研究, 2010 (6).
- [11] 万玉云. 浅谈传统出版社的全媒体数字出版 [J]. 中国编辑, 2012 (2).
- [12] 黄先蓉, 李晶晶. 浅析数字出版的版权保护策略 [J]. 科技与出版, 2012 (12).
- [13] 吴耀根. 数字出版的知识服务专业化路径探讨 [J]. 编辑学刊, 2018 (6).
- [14] 于成, 王琳. 剑桥大学出版社数字出版转型之路 [J]. 出版发行研究, 2018 (11).
- [15] 吴军. 智能时代: 大数据与智能革命重新定义未来 [M]. 北京: 中信出版社, 2016: 40.
- [16] 洪云. 大数据技术与数字出版模式研究 [J]. 传媒, 2015 (12).
- [17] 米淑惠. 图书出版与国家软实力 [M]. 北京: 光明日报出版社, 2015: 7.