

基于知识图谱的教育大数据的研究热点和未来趋势*

石东平 张孝强

(湘潭大学环境与资源学院 湖南湘潭 410083)

摘要: 基于CiteSpace和Vosviewer对中国知网2014~2021年间的1926篇出版物数量、机构、高频词、突显词等进行计量与可视化分析,通过对教育大数据研究领域的出版物计量统计数据和热点主题、历史发展趋势分析发现,教育大数据教学研究从快速发展阶段进入稳定发展阶段,机构间合作发文较少,缺乏有影响力的科研机构群体。从高频关键词和突发关键词可以看出,研究热点从大数据教育产物的运用过渡到大数据技术的研究,目前对教育改革和教育模式的研究是热点前沿,仍具有很大的研究潜力。

关键词: 文献计量分析 教育大数据 可视化 教育改革

中图分类号: G434 **文献标识码:** A

DOI: 10.12218/j.issn.2095-4743.2023.10.070

引言

大数据时代,技术的快速发展和升级,让人们的工作生活、学习方式都发生了很大的变化。传统的课堂教学方式由于受到时间和空间的局限,对于学生的碎片化学习、多样化学习等各种学习需求的缺失已经愈发明显,而大数据教育打破了时间和空间的限制,具有开放性、自主性、灵活性、个性化、交互性等优势特点,可以很好地弥补传统教育的不足之处。大数据技术为教育变革、教育模式创新的有效实现提供了技术支持。近年来,由于国家对教育信息化的号召,我国教育大数据领域的理论基础和实践应用研究开始快速发展。对一个学科的文统计和可视化分析是理解研究内容、寻找数据间联系和最大化数据价值的重要方法,也能够向读者清晰地展现、快速了解该研究学科的历史发展概貌。本文利用文献计量学统计与可视化分析方法,对2014~2021年间在中国知网上发表的教育大数据领域的出版物进行搜集、整理与可视化分析,阐明了我国教育大数据领域的研究现状与历史发展进程,并对近年来我国教育大数据领域的热点和研究趋势进行了深入分析,对其他学者开展后续研究和实践探索具有一定的参考价值。

一、研究方法

1. 样本选取

中国知网汇聚了数量庞大的学术期刊、专利、学位论文等资源,是目前中国最具影响力的数据库。目前,知网支持简单词汇或句子检索,能使用时间、论文类型等进行论文的

分类。本研究使用中国知网对关键词“教育&大数据”对论文标题、摘要、关键词进行精确检索。选取文献类型为学术期刊和学位论文,同时筛选掉相关度不大的文献,最后得到1927篇文献。

2. 研究工具与研究方法

本文使用CiteSpace联合VOSviewer对相关文献进行文献计量与科学知识图谱分析,这也是目前主流的计量分析软件。CiteSpace和VOSviewer作为学术界常用的文献分析可视化工具,都可以对关键词、作者、杂志等信息进行深入挖掘,并基于引文、共引或者共同作者关系来构建和可视化共现关系网络。可视化呈现,其中CiteSpace可以更好地展现某一学科的研究热点的历史进程。VOSviewer具有更好的可视化呈现效果,有利于挖掘文献中包含的底层知识结构^[1]。

二、文献分析与讨论

1. 发文量时间分布

2014年,教育大数据领域研究文献逐渐出现。从图1可以看到,2014年已有少量的教育大数据领域的论文被发表。2015年,“互联网+”、大数据等关键词被提上历史的舞台。源于国家政策的导向和引领,大数据教育领域受到了广泛关注,2015年开始,国内教育大数据领域的研究文献呈现爆发式增长的趋势。2019年,教育大数据领域研究发文量达到峰值,2020、2021年,发文量有所回落,但也保持着一定的数量,疫情的流行导致学生或一些学术团体的研究受阻可能是造成这一现象的一大原因。

*基金项目:湖南省普通高等教育教学改革重点项目:教育大数据驱动下安全工程专业个性化教学模式研究(HNJG-2022-0080);湖南省普通高等教育教学改革一般项目:基于应急管理形势下安全类专业课程群的课程思政链建设研究(HNJG-2022-0591);湘潭大学校级课程思政示范课程建设项目(安全系统工程)。



图1 发文量时间分布图谱

2. 发文机构时序图分析

高产机构研究时序图谱可以呈现国内教育大数据研究的机构或核心科研群体及其出现历史路径。为探究不同机构在不同时间内对教育大数据领域的研究情况，设置citespace可视化分析的时间跨度从2014~2021，以1年为间隔，生成了高产机构时序图谱。发文机构时序图中，共有210个节点数量，产生了50条相关连线，网络整体密度为0.0023。本文收集的数据结果说明在教育大数据领域开展了研究、有所成果的机构共有210个，机构间的合作次数为50次，机构的中心线很低，研究团体分散，缺乏合作^[2]。

中南大学马克思学院最早在2015年开展大数据教育研究，也与其他机构有合作，但该机构后期在教育大数据方向的发文量较少，关注度有所下降，整体上影响力一般。华中师范大学马克思学院是发文量最多的机构，作为高产机构时序图谱的一个核心节点，可以说是大数据教育领域的核心研究发文机构。华中师范大学马克思学院与镇江船艇学院信息技术教研室和蚌埠医学院公共课程部在2015年合作密切。以浙江师范大学教师教育学院、东北师范大学、重庆邮电大学等机构形成了以浙江、江苏、重庆、东北地区为主体，涉及全国多地机构的比较紧密的合作群体，此后研究群体较为稳定，研究大多集中于2017~2018。2019~2021年开展教育大数据领域研究的机构也很稳定，但是机构间缺乏合作，大多是独立开展的研究。2015年后，我国教育大数据研究进入爆发性增长的状态，很多机构开始进行教育大数据应用的研究。此后，我国对于教育大数据研究层出不穷，广受关注，并在2017~2018形成了有一定影响的研究群体。2019年后，我国教育大数据领域的研究机构大多独立分散，只有少量的机构间有所合作，也因此没有强影响力和凝聚力的科研机构 and 团体出现。

3. 关键词网络共现聚类分析

关键词，特别是高频关键词，是文献主要研究内容和核

心的体现。对关键词进行网络共现分析，可以全面认识该研究领域的核心热点。聚类是将频繁组合在一起的关键词集合，聚类分析可以很清晰地展现某一领域的不同研究课题。将CNKI中检索的文献导入VOSviewer，设置关键词频次最小值为12，剔除无意义关键词、合并同义关键词后，得到共有58个关键词，节点间的连线共438条，相互之间的联系强度为2101。得到关键词网络共现聚类图，中心度和频次高的关键词一定是某一阶段被学者们高度关注的主题或研究方向。关键词占据该研究领域网络结构的中心，其重要性越大。图中的节点和关键词字号的大小与词频高低成正比，出现频次和中心性越高的关键词在这一研究领域越重要。为了更直观地显示高频关键词，使用FineReport对前10个有代表性的高频关键词制作了词频云图，如图2所示。

由图2可见，国内教育大数据领域的出版物中的高频关键词有大数据、高校、大数据时代、教学改革、思想政治教育、创新等。对关键词聚类分析后可以得到三大聚类。一是大数据时代下的教育应用（红色）包括关键词大数据时代、思想政治教育、高校、大学生等。二是大数据教育技术研究（蓝色）包括关键词教学改革、教学模式、教育大数据、数据挖掘等。三是以大数据为基础的教育改革（绿色）包括关键词大数据、教育、改革等。

从高频关键词和聚类分组情况可以看出，利用大数据技术实现对传统教育的突破是时代发展的趋势，将大数据技术应用到真实的课堂教育是最受学者们关注的热门研究方向。目前，大数据教育方式主要运用在思想政治教育上，各类高校的大学生是主要目标人群。对大数据技术的基础研究是实现大数据教育的基石，是受到持续关注的研究方向。对教学模式的改革和创新是实现大数据教育非常关键的一环，作为热点研究方向，但其总体研究热度不高^[3]。



图2 高频关键词云图

4. 研究前沿和趋势分析

突现词是指在一段时间内出现频率很高的关键词，突现词的词频变化映射着该研究领域的热点前沿。利用CiteSpace可以清晰地得到教育大数据领域的关键词突现频次曲线。

从图3可以清晰地知道,“高等教育”“微课”“慕课”“云计算”“思政教育”都是教育大数据领域的阶段性研究热点。其中,“高等教育”“微课”主要体现在2015~2016年,“慕课”体现在2016~2017年,“云计算”体现在2017~2018年,“思政教育”“模式”体现在2019~2021年。教育大数据领域的研究热点从高等教育、在线教育平台转变为对云计算的研究,再转变为教育模式研究,研究趋势表现为逐年上升。根据研究趋势变化,我们可以清晰地看到,研究热点已经不仅仅是在在线教育平台的建设和教育大数据核心技术领域研究,更多是对于教育大数据背景下的教育改革、教育模式的研究。可以看出,现在我国教育大数据研究前沿和趋势是思政教育和教育改革领域。

Top 10 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	2014 - 2021
高等教育	2014	3.34	2015	2016	
微课	2014	2.5	2015	2016	
挑战	2014	3.51	2016	2017	
慕课	2014	3.1	2016	2017	
影响	2014	2.22	2016	2017	
发展	2014	2.78	2017	2018	
云计算	2014	2.69	2017	2018	
改革	2014	2.86	2018	2019	
思政教育	2014	2.64	2019	2021	
模式	2014	2.35	2019	2021	

图3 关键词突显词图谱

结语

2014年,教育大数据领域研究的研究文献开始出现。2015年开始,国内教育大数据领域的研究文献呈现爆发式增长的趋势。2019年,教育大数据领域研究发文量达到峰值,2020、2021年,发文量有所回落,但也保持着一定的数量。从总体上看,我国在教育大数据领域的研究成果很多,已经进入稳定缓慢发展阶段。对教育大数据研究的机构层出不穷,但发文数量大多不超过两篇,仍缺少具有足够影响力的科研机构和合作密切的科研群体。大多研究机构都是独立开展研究,缺乏相互合作。提高科研机构间的合作联系,形成有影响力和领导力的科研群体是保障教育大数据领域主流研究热度的主要待解决问题。教育大数据研究的高频关键词有大数据、高校、大数据时代、思想政治教育、教学改革、创新等。可分为三大聚类。一是大数据时代下的教育应用,对“微课”“慕课”等在线教育方式和平台的研究突显,是2015~2017年的阶段性研究热点,也是最受学者们关注的热门研究方向,占据共现网络的大部分。二是大数据教育技术

研究,技术是实现大数据教育的基石,对“云计算”的研究在2017~2018年凸显,并受到持续关注和研究。三是以大数据为基础的教育改革,对“思政教育”“模式”等关键词的研究在2019~2021年开始凸显,可以看出对大数据时代的教育改革和教育模式的研究越来越受到学者们的关注,已经成为最新的研究热点和趋势。目前,教育改革主要运用在思想政治教育上,各类高校的大学生是主要目标人群。但该聚类组成较少,网络结构相对松散,总体研究热度相对不高,仍具有很大的研究潜力^[4]。

教育大数据领域在近年来广受研究,但理论和实践研究还处于初步发展、探索应用阶段。教育变革是一个渐进的过程。随着技术和研究的不断发展,对其认识也会趋于全面、系统。针对教育大数据研究热点和前沿演变,这里也有了一些想法和建议。目前,教育大数据研究的主题较单一且存在局限性。例如,利用大数据开展思政教育的研究非常多,但对于其他基础知识教育、文化艺术类教育、医疗运动等的研究明显缺乏。同时,对高职院校和大学生群体的研究火热,但对于初等教育的学生群体研究少见。建议可以扩大教育大数据领域的研究范围,除了对教育大数据核心领域的研究外,也应该多关注各类边缘化的和新兴的研究主题,同时也需要大力开展多学科、跨领域的合作研究,积极开拓新的研究方向。

参考文献

- [1]王战军,乔伟峰,李江波.数据密集型评估:高等教育监测评估的内涵、方法与展望[J].教育研究,2015(6):29-37.
- [2]张笑予,冯东.“双一流”研究的热点领域与主题演进——基于CNKI(1992—2017年)的文献计量与知识图谱分析[J].重庆高教研究,2018(2).
- [3]叶小敏.“双一流”研究的前沿演进与热点领域——基于CiteSpace的知识图谱分析[J].现代教育论丛,2018(3).
- [4]刘浩,饶武元.我国高教领域学科建设研究的现状、热点及前沿——基于2005-2016年高等教育类核心期刊的知识图谱可视化分析[J].评价与管理,2017,5(1):90-99.

作者简介

石东平(1988—),女,汉族,湘潭大学环境与资源学院,讲师。