

基于CiteSpace的多模态学习分析研究评述

路 冉¹, 郭 煜^{2*}, 徐世东^{2*}

¹北京工业大学高等教育研究院, 北京

²北京工业大学信息化处, 北京

收稿日期: 2025年7月20日; 录用日期: 2025年8月18日; 发布日期: 2025年8月27日

摘 要

为厘清多模态学习分析领域的历史演进、研究现状与前沿动态,本研究运用CiteSpace工具,对中国知网(CNKI)与Web of Science (WOS)数据库中2014年~2024年的核心期刊文献开展可视化分析,分析内容涵盖年度发文量、机构共现、关键词共现、关键词聚类以及关键词时间线等维度,研究发现近年该主题下在技术驱动下快速发展,总体呈现出多学科交叉、技术驱动创新和多角度实际应用的趋势。未来相关研究将伴随人工智能技术发展,突破跨学科理论融合与异构数据整合瓶颈,提升教师技术应用与教学设计能力,推动复杂场景的适应性落地,我国亟待构建以领军学者为核心的跨机构合作共同体,驱动本土化理论创新与实践应用,助力教育数字化转型发展。

关键词

多模态学习分析, 多模态数据

A Review of Multimodal Learning Analytics Research Based on CiteSpace

Ran Lu¹, Yu Guo^{2*}, Shidong Xu^{2*}

¹Institute of Higher Education, Beijing University of Technology, Beijing

²Information Office, Beijing University of Technology, Beijing

Received: Jul. 20th, 2025; accepted: Aug. 18th, 2025; published: Aug. 27th, 2025

Abstract

To clarify the historical evolution, research status and cutting-edge trends in the field of multimodal learning analysis, this study utilized the CiteSpace tool to conduct a visual analysis of the core journal literature from 2014 to 2024 in the databases of China National Knowledge Infrastructure (CNKI)

*通讯作者。

文章引用: 路冉, 郭煜, 徐世东. 基于 CiteSpace 的多模态学习分析研究评述[J]. 教育进展, 2025, 15(8): 1524-1531.
DOI: 10.12677/ae.2025.1581608

and Web of Science (WOS). The analysis covers dimensions such as the annual number of published articles, co-occurrence of institutions, co-occurrence of keywords, keyword clustering, and keyword timelines. The research finds that in recent years, this theme has developed rapidly driven by technology, generally showing a trend of multi-disciplinary intersection, technology-driven innovation, and multi-angle practical application. In the future, related research will follow the development of artificial intelligence technology, break through the bottlenecks of cross-disciplinary theoretical integration and heterogeneous data integration, enhance teachers' ability to apply technology and design teaching, promote the adaptive implementation in complex scenarios. China urgently needs to build a cross-institutional cooperative community with leading scholars at its core to drive localized theoretical innovation and practical application, and facilitate the digital transformation and development of education.

Keywords

Multimodal Learning, Analysis Multimodal Data

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

我国高度重视教育数字化,于2022年启动国家教育数字化战略行动,并在《教育强国建设规划纲要(2024-2035年)》中明确了国家教育数字化战略行动的发展方向[1]。随着大数据、物联网、移动互联网、人工智能等新兴信息技术的持续演进与广泛应用,教育数据的内在价值不断显现,这一趋势促使教育领域研究从“假设驱动”向“数据驱动”转变,并推动“教育大数据研究范式”的逐步形成[2]。在此背景下,多模态学习分析作为一个显著的研究热点逐渐受到学术界的广泛关注,相关研究在理论、方法和应用各方面不断完善。2012年多模态交互国际会议上正式提出的概念受到较多认可,其核心在于通过采集、融合与分析学习场景产生的语音、动作、神情、生理信号等多源异构信息,揭示与预测学习者的行为模式、认知规律、心理特征及情感变化,进而对改善学习体验与提升学习成效具有显著意义。本研究借助CiteSpace工具对近十年相关文献展开可视化分析,通过对比国内外研究在文献产出量变化、机构合作网络、关键词等,系统梳理研究热点与发展脉络,得出结论并为后续研究提供相应启示。

2. 基本概念

2.1. 多模态

多模态是指人类通过多种感官与外部环境进行互动的方式,该研究始于20世纪90年代的西方,Kress和Van Leeuwen将其定义为综合运用多种符号学模态来加强意义表达,或用于补充功能,或进行层次排序[3]。国内对多模态的研究起步较晚。顾跃国提出的概念已得到广泛认可,多模态需要人与环境通过多种感官进行互动[4]。从传播学视角来看,多模态是信息交流的重要媒介,其核心在于通过不同符号之间的交互作用传递信息。功能语言学进一步拓展了话语分析的范畴,将言语、表情、手势和肢体语言等模态纳入分析框架,因此催生了多模态话语分析这一新领域[5]。

2.2. 多模态学习分析

多模态学习分析是一门典型的交叉学科,旨在通过多元技术手段捕获、处理并分析复杂学习情境中

的多源行为数据,揭示学习机制并对学习轨迹、学习成效及参与度等关键指标进行精确预测。该领域由 Scherer、Morency 与 Worsley 于 2012 年首次提出[6],其核心在于整合自然交互过程中生成的多元异构数据,如语音、手势、姿势、眼球运动、面部表情、脑电图、心率等,对学习者的行为、认知、动机和情感状态进行客观刻画与深刻洞察[7]。Blikstein 将其界定为“借助多源数据检验复杂学习情境的一系列技术集合[8]”。Drachsler 等进一步指出,相较于单一模态,多模态数据能够更全面地还原学习过程,从而精确归因学习结果的影响机制[9]。王一岩等学者强调该领域具有多元素互补、分层分解、智能感知与时空融合等特点[10],可有效支撑多样化学习环境的过程建模和机理分析。随着情感计算、语义分析、表情识别等关键技术的持续突破,多模态学习分析的理论深度与应用广度正不断拓展。

3. 研究方法

为深入了解多模态学习分析领域的研究现状,本研究以中国知网(CNKI)和美国科学情报研究所(ISI)的 Web of Science (WOS)的核心收录数据库作为数据源。在 CNKI 中,选择时间跨度为 2014 年至 2024 年,采用了高级检索,主题词为“多模态”并含“学习分析”。在人工剔除无关文献后,最终获得 136 篇有效样本。在 WOS 数据库中,检索式为 TS=(“多模态学习分析”),语言限于“English”,文献类型为“Article”。经过筛选、对比和去重,共获得 166 篇有效文献。研究利用数据库内置的计量分析功能进行统计分析,并选用陈超美教授开发的 CiteSpace 6.3.R1 软件进行可视化分析,总结出该领域的概况、热点、挑战及趋势。

4. 研究可视化分析

4.1. CNKI 和 WOS 计量可视化分析

各年度发文量

通过分析 CNKI 检索到的文献的发表时间,可以得到发表量分布的折线图,如图 1 所示。研究发现,2014 年之前,各类核心期刊均未出现相关主题的文献。从 2015 年到 2022 年,我国多模态学习分析研究呈逐渐上升趋势。发表论文数量在 2022 年最多,从 2023 年到 2024 年开始逐渐减少。如图 2 所示,分析 WOS 检索到的文献可以看出,2015 年至 2017 年,WOS 发表的相关论文相对较少,2018 年至 2021 年发表的论文数的多模态学习分析研究呈逐渐上升趋势,2022 年与 2023 年开始逐渐减少,但发表论文数量在 2024 年达到顶峰。总体来看,国内外文献发表数量趋势较为相似,研究整体呈现快速发展趋势,近两年仍是研究者关注的热点话题,尤其是国外对于该领域的研究数量在 2024 年达到顶峰。

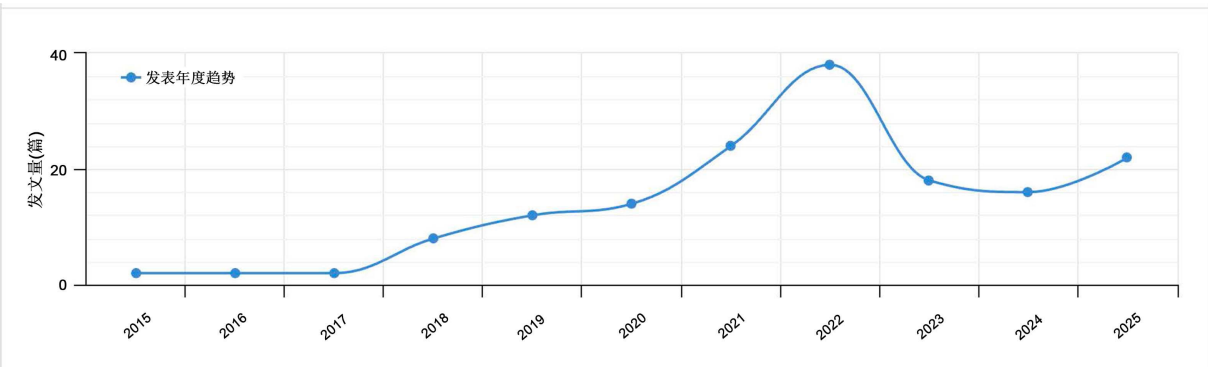


Figure 1. Shows the annual changes in the number of published papers in the field of domestic multimodal learning analysis research from 2015 to 2024

图 1. 2015~2024 年国内多模态学习分析研究领域发文量年度变化

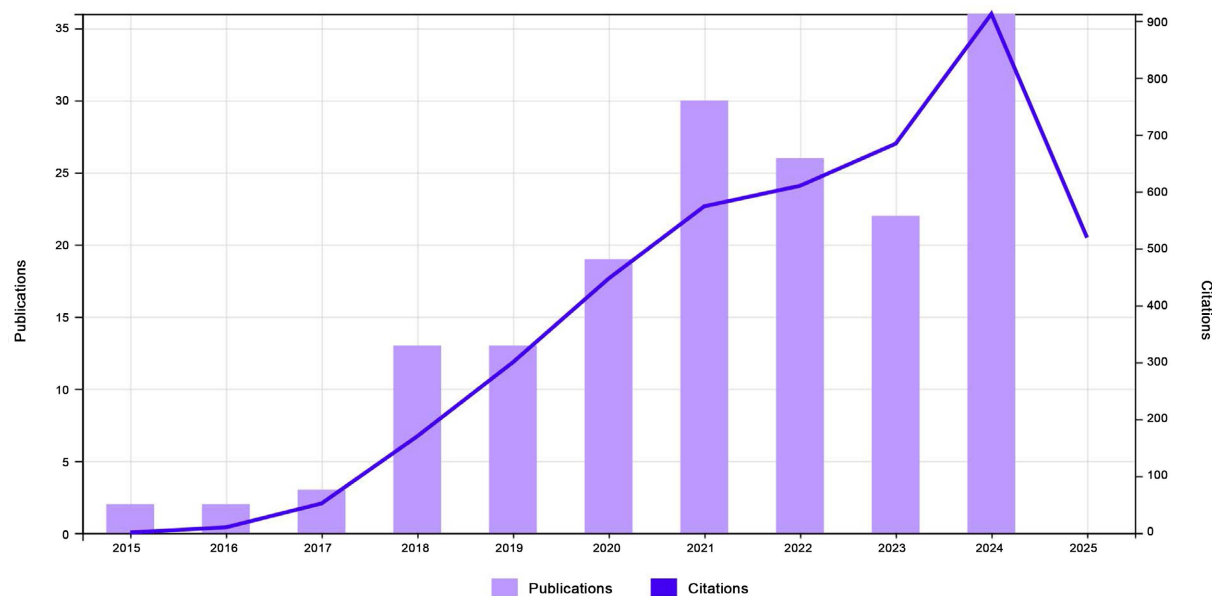


Figure 2. Shows the annual changes in the number of published papers in the field of multimodal learning analysis research abroad from 2015 to 2024
图 2. 2015~2024 年国外多模态学习分析研究领域的发文量年度变化

4.2. CiteSpace 可视化分析

4.2.1. 机构合作网络

为了解国内外研究机构在相关领域的合作情况，本研究运用 CiteSpace 软件对机构合作进行分析，国内外排名前 5 位的研究机构如从表 1 所示知。国内对该领域的研究主要以以师范院校为主导。北京师范大学、华东师范大学、华南师范大学与华中师范大学处于领先地位，在该领域具有重要的学术影响力，其他如江南大学、北京大学等也表现出一定的活跃度。CNKI 数据库的机构贡献图谱中共有 142 个结点，106 条连线，网络密度为 0.0106。这表明机构间的合作练习相对稀疏且有限，合作模式主要以内部合作为主。其中北京师范大学发文量最多，该机构包括了教育学部、教育技术学院、互联网教育智能技术及应用国家工程研究中心。WOS 平台上发文量居前三的机构分别为：Pontificia Universidad Catolica de Valparaiso、Universidad de Valparaiso、Harvard University，其他机构也有部分程度的参与。WOS 平台上机构共现图谱中有 155 个节点数，229 条联系线，网络密度为 0.0192，机构之间的合作联系较为紧密，形成以几所核心高校为首的多个研究机构团体，跨机构合作模式较为普遍。综合来看国内研究机构间的合作模式表现出较强的内部性，跨校际交流与合作相对较少。而国际研究机构则展现出更强的外向协作性特征。

Table 1. The top 5 institutions in terms of the number of published papers in the field of multimodal learning analysis research at home and abroad from 2014 to 2024
表 1. 2014~2024 年国内外多模态学习分析研究领域发文量前 5 名的机构

CNKI 数据库		WOS 数据库	
机构	发文量	机构	发文量
北京师范大学	29	Pontificia Universidad Catolica de Valparaiso	8
华东师范大学	23	Universidad de Valparaiso	8
华南师范大学	18	Harvard University	8
华中师范大学	17	Norwegian University of Science & Technology (NTUN)	5
江南大学	5	Universidade Federal de Santa Catarina (UFSC)	5

4.2.2. 关键词共现分析

对高频关键词进行共现分析，节点周围圆圈越大，表示该关键词是被关注的热点关键词，关键词中心性大于 0.1 为重要关键词，中心性越大，影响力越高。表 2 为国内外频率排名前 15 位的核心关键词。

Table 2. Keyword distribution in the field of multimodal learning analysis research at home and abroad from 2014 to 2024
表 2. 2014~2024 年国内外多模态学习分析研究领域关键词分布

中文				外文			
序号	关键词	频率	中介性	序号	关键词	频率	中介性
1	学习分析	31	0.40	1	multimodal learning analytics	31	0.31
2	多模态	19	0.42	2	machine learning	28	0.16
3	人工智能	16	0.37	3	artificial intelligence	13	0.14
4	学习投入	12	0.23	4	analytics	12	0.09
5	深度学习	12	0.34	5	deep learning	12	0.10
6	协作学习	8	0.18	6	learning analytics	9	0.13
7	数据融合	7	0.11	7	feature extraction	8	0.17
8	情感计算	6	0.15	8	big data	8	0.23
9	情感分析	6	0.17	9	model	7	0.11
10	数据驱动	5	0.20	10	education	7	0.17
11	智慧课堂	5	0.22	11	performance	6	0.04
12	发展趋势	4	0.1	12	classificaation	6	0.07
13	人机协同	4	0.1	13	framework	6	0.10
14	机器学习	4	0.14	14	multimodal data	5	0.00
15	数据挖掘	4	0.25	15	visual analytics	5	0.00

中文文献中的关键词分布中，“学习分析”和“多模态”关键词中心性较高，构成研究核心范畴。人工智能、深度学习、机器学习等关键词凸显技术对该领域发展的驱动作用；学习投入、协作学习属于该领域的研究分支；多模态、数据融合、数据驱动突出了多源数据融合分析的核心特征。在线学习、智慧课堂则可以看出中文研究聚焦于教育场景的应用，强调技术赋能教育，关注学习过程中的“人机协同”与“人机交互”，致力于推动“人工智能 + 教育”的实践落地。

外文文献中，multimodal learning analytics (多模态学习分析)这一关键词中心性突出；machine learning (机器学习)是支撑该领域的核心技术与方法。learning analytics (学习分析)和 performance (表现)、collaborative learning (协作学习)构成重要基础研究主题；visual analytics (视觉分析)、feature extraction (特征提取)涉及具体技术方法。与中文文献不同，外文研究除关注教育领域外，还包含了如 science (科学)、biomedical monitoring (生物医学监测)、atmospheric measurements (大气测量)等关键词，显示出更强的跨学科拓展性，注重基础技术的创新应用。

这些关键词揭示了该领域近 10 年的研究热点，“人工智能”、“机器学习”、“深度学习”是其重要关注点，教育为核心应用场景，通过场景中数据收集、融合与分析以理解学习过程。中文研究在教育应用场景的深化、集成化以及特定教学环节的细分研究上表现突出；外文研究则更突出“机器学习”的基础性作用，并展现出更强的多领域交叉应用拓展趋势。

4.2.3. 关键词聚类分析

基于关键词聚类分析,获取多模态学习分析的研究热点主题。根据 CiteSpace 软件进行分析得到聚类分析结果。中文文献图谱参数模块值 $Q = 0.8355$, 大于 0.3, 平均轮廓值 $S = 0.9476$, 大于 0.7; 外文文献图谱参数模块值 $Q = 0.848$, 大于 0.3, 平均轮廓值 $S = 0.9555$, 大于 0.7。这说明聚类图谱结构显著, 结果可信度高。中文文献图谱主要聚类为人工智能、深度学习、学习评价、学习分析、在线学习、大数据、发展趋势、教学视频、数据融合、数理实验、影响因素、教师画像。外文文献图谱主要聚类为眼部追踪、机器学习、深度学习、脑电图、大数据、活动认知、特征提取、适应性学习、远程学习等。聚类结果可以概括出当前研究热点集中在以下几方面: 第一、机器学习与深度学习等技术的发展, 研究者能够系统获取、精细处理并深度分析学习过程中的多源异构数据。通过结合数据挖掘与可视化工具, 将原始数据转化为可解释的学习关键指标, 进而精准预测学习者知识状态, 实现个性化资源推送。这显著推动了学习分析领域的发展。第二、研究以学习者为中心, 借助于大数据的信息采集和处理技术, 研究者可以对学习过程中的各种感官进行分析, 通过对语音、表情、姿态、生理信号等多元异构数据的采集与分析, 能够实现对学习者认知、情感、动机与行为等学习过程的立体化分析, 为学习者良好学习体验打下基础。第三、研究凝聚于学习场域, 研究者可以根据在线学习或者混合学习场景中的视频、语音等进行分析, 不仅能够实现场景间的无缝衔接, 还能揭示个体学习规律的动态演化, 努力实践精准教学和适应性学习的方式。此外, 除教育领域外, 相关方法已被迁移至医疗、环境等场景, 显示出显著的跨学科迁移潜力。

4.2.4. 关键词时间线图

关键词时间线图能够反映某一研究主题随时间变化的研究趋势, 对中文文献每个聚类主题关键词沿时间轴分析可知, “多模态”、“学习分析”的研究在 2015 年成为整个研究领域内最为突出的研究热点后, 关于深度学习、协作学习、人工智能在该领域的探讨陆续被研究者关注, 且上述聚类的相关研究一直延续至近两年。对外文文献每个聚类主题关键词沿时间轴分析可知, 2015 年以后, 多模态学习研究在 machine learning (机器学习) 聚类与 deep learning (深度学习) 聚类下形成多个研究热点, 被大量学者关注, 对中外关于多模态学习分析研究热点的演进情况具体可分为三个阶段:

2015~2018 年, 这一阶段随着混合式学习、学习分析和大数据技术的逐渐成熟, 大量线上、线下学习数据涌现, 可供研究者采集和分析, 研究者通过挖掘分析多模态特征数据识别行为、情感等关键要素, 并尝试构建数据模型进行实际场景应用。

2018~2021 年, 在技术与教育深度融合的背景下, 人工智能技术在多模态学习分析领域得到广泛运用, 数据呈现出复杂性、情景性、层级性等特点, 国内外学者更加注重实践应用场景, 在课堂中开展了一系列多模态数据交互行为相关研究。随着在线会议系统、网络平台、虚拟现实等技术在课堂中的应用, 多模态数据应用于在线学习研究, 注重收集与观察手势、目光接触、语言交流、脸部头像和平台记录等模态数据。在智能教育环境的支持下, 更加关注学习者的个性化培养, 努力实践精准教学和适应性学习的方式。

2021~2024 年, 大模型时代的到来, 多模态学习分析研究进入全新发展阶段, 信息技术向更高级别的智能化和自动化迈进, 研究视角更加关注教育过程中的“人机协同”、“情感计算”等; 教育应用场景也不断拓宽到智慧课堂、STEM 等领域; 同时信任机制与人机伦理等问题也进入了学者的研究视野中。

5. 研究结论

5.1. 多模态学习分析研究现状

对比国内外多模态学习分析发文量可知, 该领域研究内容逐步丰富, 研究广度与深度逐步拓展。特

别是 2020 年以来,研究成果呈骤增态势。根据国内外研究机构和作者的分布情况可知,核心文献均集中于少量高等师范院校,形成显著的“核心-边缘”结构。机构间的合作网络稀疏,跨校际、跨地域的协同研究成果数量有限。当前亟须构建研究共同体,通过常态化跨机构合作、共享数据库平台及学科交叉基金等机制,驱动研究向纵深发展。

5.2. 多模态学习分析研究热点分析

国内外研究始终以“深度学习”“大数据”“人工智能”为驱动力,但技术迭代持续演变下研究路径具有明显差异:国内研究从早期聚焦于学习行为与场景应用探索逐步转向并深入学习内在机制与复杂环境的构建,当下更关注虚拟现实应用及教育伦理规范;国外研究从早期侧重于学习机制解析与基础模型构建到注重深入挖掘学习复杂认知过程与自动化评估等,当下更加强调技术效能提升与跨学科实践整合。总体而言,国际研究注重底层技术突破与理论创新,国内研究则更侧重于应用场景的拓展。与国际研究相比,我国多模态学习分析研究热点更为丰富,且研究持续性较强,但主题间关联较为松散,缺乏深入研究。

5.3. 多模态学习分析面临挑战分析

综合分析中外研究进展,多模态学习分析在展现巨大潜力的同时,也面临着多维度挑战:在理论构建层面,不同学科固有的知识体系和研究范式存在差异,如何实现跨学科的有效融合,是当前面临的重要挑战;在伦理方面,多模态数据的广泛收集易引发对学生隐私权、“全景监控”风险等的担忧,如何确保数据收集的知情同意、最小化原则,以及数据存储、使用、共享的透明度和安全性,是亟待解决的核心伦理问题;在技术方面,模型存在跨场景迁移性差、普适性不足的问题,多模态异构数据的深度整合与协同分析仍是领域发展的核心瓶颈;在实践方面,人机协同时代下教师技术素养不足影响分析结果向教学决策的有效转化,同时复杂动态的教育场景缺乏成效评估标准与实证支持。

5.4. 多模态学习分析未来趋势分析

未来的中外文研究将展开深入探索:首先,在理论构建层面,追踪教育神经科学、脑科学和学习科学的前沿进展,强化教育学、计算机科学、心理学等学科的深度协作,实现跨学科的有效融合。其次,针对模型迁移性差、异构数据整合困难等问题,利用深度学习等先进技术挖掘多模态数据的潜力,提升分析的精准度和洞察力。重点关注全方面数据的汇聚与融合、不同模态的分析与建模,并建立隐私计算驱动的数据安全防护体系以应对伦理风险;为克服实践障碍,需要将人工智能、云计算等关键技术嵌入真实教育场景,建立起线上、线下跨时空教学场景数据的关联,实现对学习者发展的动态分析^[11];通过分层培训与专家指导,提升教师技术应用与教学能力,结合学科特点梳理差异化实践路径,鼓励教师开展人工智能赋能的实践探索;对我国而言,亟需以领军学者为核心、跨机构团队为载体,构建常态化合作机制,以多样化的团队合作促进该领域理论概念统一与集成方法拓展,从而有效驱动多模态学习分析向更高水平的知识生产与实践转化。

6. 研究小结

综合近十年文献可见,多模态学习分析领域的发文量持续增加,研究不断深入。总体而言该领域呈现出多学科交叉、跨学科参与、技术驱动创新和应用场景深化的特征:在理论层面融合教育学、计算机科学及认知神经科学等多学科范式;技术层面依托深度学习推动多模态数据整合与模型优化;应用层面从单一环境向真实复杂教学场景迁移。然而当前研究仍面临跨学科理论融通不足、数据隐私伦理风险、模型跨场景迁移性差、教师技术素养缺失等核心挑战。未来研究需攻关异构数据融合与隐私计算技术,

建立人机协同发展机制,提升教师角色转变与能力提升,同时以本土化跨机构合作共同体推动理论创新与场景适配,为教育数字化转型提供方法支撑与实践路径。

基金项目

本项目由北京高教学会 2024 年研究课题资助,基于服务蓝图的高校 AI 门户建设研究,课题编号 MS2024405。

参考文献

- [1] 中共中央国务院印发《教育强国建设规划纲要(2024-2035 年)》[N]. 人民日报, 2025-01-20(006).
- [2] 李政涛, 文娟. 计算教育学: 是否可能, 如何可能? [J]. 远程教育杂志, 2019, 37(6): 12-18.
- [3] Faigley, L., Kress, G. and van Leeuwen, T. (2002) Multimodal Discourse: The Modes and Media of Contemporary Communication. *College Composition and Communication*, **54**, 318-320. <https://doi.org/10.2307/1512155>
- [4] 顾曰国. 多媒体、多模态学习剖析[J]. 外语电化教学, 2007(2): 3-12.
- [5] Veltman, R. (1996) Gunther Kress and Theo Van Leeuwen. *Reading Images: The Grammar of Visual Design. Functions of Language*, **3**, 275-277. <https://doi.org/10.1075/fol.3.2.15vel>
- [6] Scherer, S., Worsley, M. and Morency, L.P. (2012) ICMI'12: International Conference on Multimodal Interaction. *Proceedings of the 1st International Workshop on Multimodal Learning Analytics*. Santa Monica, 26 October 2012. <https://dl.acm.org/doi/10.1145/2389268>
- [7] Worsley, M. (2012) Multimodal Learning Analytics: Enabling the Future of Learning through Multimodal Data Analysis and Interfaces. *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, Santa, 22-26 October 2012, 353-356. <https://doi.org/10.1145/2388676.2388755>
- [8] Blikstein, P. and Worsley, M. (2016) Multimodal Learning Analytics and Education Data Mining: Using Computational Technologies to Measure Complex Learning Tasks. *Journal of Learning Analytics*, **3**, 220-238. <https://doi.org/10.18608/jla.2016.32.11>
- [9] Drachsler, H. and Schneider, J. (2018) JCAL Special Issue on Multimodal Learning Analytics. *Journal of Computer Assisted Learning*, **34**, 335-337. <https://doi.org/10.1111/jcal.12291>
- [10] 王一岩, 王杨春晓, 郑永和. 多模态学习分析: “多模态”驱动的智能教育研究新趋向[J]. 中国电化教育, 2021(3): 88-96.
- [11] 尹睿, 何淑茵. 基于系统性文献综述的多模态学习分析研究进展与前瞻[J]. 现代远程教育研究, 2022, 34(6): 54-63.