arXiv:2603.28802v1 [cs.DL] 27 Mar 2026

用于可视化和理解系统综述的交互式证据地图

Aditi Mallavarapu*[1]、Rohan Khandare[1]、Mokshagna Kadiyala[1]、Neelesh Yaddanapudi[1]、、Noah L. Schroeder[2]、、Shan Zhang[2]和 Jessica R. Gladstone[3]

[1]北卡罗来纳州立大学，美国北卡罗来纳州罗利，邮编 27695[2]佛罗里达大学，美国佛罗里达州盖恩斯维尔，邮编 32611[3]伊利诺伊大学厄巴纳-香槟分校，美国伊利诺伊州尚佩恩，邮编 61820

2026年4月1日

摘要

系统综述对研究领域进行全面综合。因此，系统综述往往侧重于综合大量文献，而非仅仅描述得出结论所依据的研究。这可能导致样本描述不完整，进而造成对研究结果的过度概括、掩盖现有研究之间的联系，或忽视文献中的空白。为应对这一挑战，我们提出交互式证据地图——一种易于使用的可视化工具，使研究人员能够动态地探索、筛选和分析综述数据。我们的方法利用大语言模型提取主题模型，将异质的综述数据组织成可交互、可探索的知识图谱，从而支持深入考察，超越静态表格和图表的局限。我们使用一项已发表的、关于 K-12 教育中教学代理的范围综述数据，展示交互式证据地图的实用价值，并将证据地图的结果与该范围综述报告的结果进行比较。结果表明，交互式证据地图能够补充传统综合方法，提升透明度、支持探索性分析，并揭示仅凭叙述性摘要可能难以发现的模式和空白。

1 引言

系统综述和荟萃分析是教育研究中的基石性方法论。它们一直在提供全面的综合

*通讯作者：amallav@ncsu.edu

为循证实践和政策决策提供依据 [1]。与其他领域一样，在教育人工智能（AIED）领域，系统综述在梳理学习技术、教学干预及其在不同教育情境中的有效性方面发挥了重要作用 [2]。然而，正是系统综述的全面性使其具有价值，同时也带来了重大挑战。综述作者综合大量文献中的研究结果时，重点往往转向高层次的综合结论，以牺牲详细的研究层面描述为代价，概括出“大致轮廓”。这种以综合为重点的方法虽然对于得出可操作的见解必不可少，却可能导致研究结果被过度概括、研究之间的细微联系被掩盖，并使文献中亟待进一步调查的关键空白黯然失色。

随着研究数量持续呈指数级增长，如何在综合与透明度之间取得平衡这一挑战日益严峻。无论研究领域如何，传统的呈现形式——静态表格、叙述性摘要和固定可视化——都难以应对现代系统综述所固有的复杂性与异质性 [3] 。嵌入综述文档中的预设条形图、森林图和汇总图等固定可视化，虽然能清晰呈现信息，却缺乏支持读者驱动探索的灵活性 [4, 5]。例如，一篇系统综述可能会用条形图展示小学、初中和高中各教育阶段开展的研究数量，从而概览证据分布。然而，这种可视化是固定的：读者无法与之交互，因而不能进一步询问诸如“这些小学研究中，哪些聚焦于数学？”或“高中阶段开展了多少项随机对照试验？”之类的问题 [6, 7]。同样，展示发表趋势随时间变化的散点图也无法进行筛选，以考察特定干预类型的趋势，或单独查看采用特定结果测量指标的研究。这些静态呈现反映了综述作者的分析重点，却无法适应不同读者对证据基础抱有的各种问题和兴趣 [8]。多位作者指出了静态可视化在系统综述中的局限，并强调需要采用更灵活的交互式方法，让用户能够根据自身的分析需求探索证据 [9, 10, 7, 11, 12, 5]。读者往往无法深入探查底层证据基础，无法探索与自身具体兴趣相关的研究子集，也无法发现可能不符合作者主要综合叙述的模式。在教育人工智能等跨学科领域，这一局限尤其突出，因为相关研究涵盖多种方法论、理论框架和应用情境。尽管 Open Knowledge Maps、VOSviewer 和 CmapTools 等现有可视化工具已提升文献计量分析与知识图谱绘制能力，但它们往往仍难以支持特定的分析

需求，尤其难以动态探索研究特征、结果和情境因素，而这些因素正是教育技术研究的界定依据 [16]。

为弥补这一空白，我们提出了交互式证据地图——一种新颖的方法，利用基于大语言模型（LLM）的主题建模，将系统综述数据转化为动态、可探索的知识结构。我们的方法将自动化主题提取与交互式可视化相结合，创建证据地图，使研究人员、从业者和政策制定者能够从多个粒度层次了解系统综述结果。与静态综合文档不同，交互式证据地图支持对底层研究语料库进行筛选、聚类和探索性分析，从而揭示传统叙述性摘要中可能被掩盖的模式与空白。通过保持高层次综合与研究层面细节之间的联系，我们的方法提升了透明度，满足了不同用户的需求，并对传统系统综述方法形成补充，而非取而代之。

2 相关研究

2.1 研究综述中的可视化与知识图谱绘制

研究图景的可视化在多个学科中有着悠久的历史，已有各种工具和方法论用于支持文献计量分析 [17, 18] 和知识图谱绘制 [19, 15]。这些研究成果确立了有效的科学出版物衍生可视化的重要原则，包括交互性、多尺度表示和语义清晰度的重要性。然而，尽管这些框架擅长大规模绘制整个科学学科和引文网络，但它们通常应用于综合性文献计量数据库，而非针对特定研究问题、经过筛选并系统综述的证据集；后者正是系统综述的特点。

VOSviewer [14], 及其扩展工具 *CitNetExplorer* [20], 一直广泛用于构建和可视化文献计量网络。这些工具侧重于基于引文的关系和关键词共现，但对纳入系统综述核心内容的研究层面特征、方法学特征或结果数据的支持有限。像*Open Knowledge Maps* [13], LitMaps [21] 和 OpenAlex [22] 这样的网络制图工具，将文本挖掘与可视化相结合，以直观呈现研究主题。然而，这类工具采用自动化的黑箱方法提取信息，缺乏系统综述所要求的严谨性和全面的纳入标准。首先，这些工具难以纳入系统综述者针对纳入研究开展的详细编码和质量评估。此外，它们只能呈现文献综述的概貌

，若要提取全面的综合信息，则需要具备专业技术知识，因此在我们的使用场景中不易普及。

2.2 用于综合与知识图谱绘制的主题建模和计算方法

随着自然语言处理和机器学习的发展，用于理解研究图景的计算方法取得了显著进展。传统的主题建模方法，尤其是潜在狄利克雷分配（LDA），已应用于大型科学文献语料库，以识别潜在的主题结构 [23]。尽管这些方法能够揭示数千份文档中的主题驱动模式，但它们生成的主题往往缺乏语义连贯性，或无法与系统综述情境中有意义的区分相契合。

BERTopic 及其科学领域变体（例如 SciBERT [24], Specter2 [25]）代表了主题建模领域的一项较新进展，晚于其语义能力较弱的前身 LSA（潜在语义分析） [26]。这些基于Transformer的方法能够生成语义连贯性更强、情境相关性更高的主题簇 [27]。简而言之，它们将文档嵌入与基于类别的 TF-IDF（词频-逆文档频率）或其他相似度度量表示相结合，通过捕捉语义细微差别并区分关键术语来生成主题。众所周知，与前代方法相比，这些方法能够生成更准确的表示和更清晰的聚类区分，更贴合人类对主题结构的理解。迄今为止，BERTopic 及其后续发展在系统综述情境中的应用仍不多见（例如， [28]）。这些有限的应用有力地展示了将结构化编码方案与主题建模相结合、用于系统分析的价值。然而，仍需探索如何以便于使用的方式呈现这些结果，从而促进此类应用。

一种颇受欢迎的方法是将主题建模方法与聚类算法和降维技术相结合，把研究文献整理成知识图谱 [29, 30]。虽然主题模型可能产生难以解读的信息，但知识图谱可以呈现为可视化图表，便于对证据中的多个信息层进行映射和导航。除了揭示主题/子主题之外，这些技术还可以揭示研究之间的邻近关系；若与人工编码信息相结合，还能揭示有助于开展系统综述多层次综合的具体方面。

2.3 系统综述的可视化方法

虽然可视化工具在文献计量学中很常见，但其在系统综述综合中的具体应用仍相对不成熟。传统系统综述主要依靠表格和森林图呈现荟萃分析结果 [31], ，并通过叙述性文本展示研究发现。森林图能够有效传达效应量和置信区间，但仅适用于定量研究，也无法展现证据基础更广泛的特征 [31]。证据与空白地图是可视化系统综述中证据范围及分布的一种方法 [32]。这类地图通常采用矩阵结构，展示哪些干预措施在哪些人群或情境中得到研究，并用单元格表示现有证据的数量，有时也表示其质量。传统的证据与空白地图虽有助于识别空白，但属于静态地图，无法支持基于多种研究特征同时进行动态探索或筛选。

将网络可视化应用于系统综述并非新概念。Radianti 等人 [33] 开展了一项系统综述，采用网络可视化技术绘制研究、主题和应用领域之间的关系。他们的方法展示了可视化如何增强综述结果的呈现效果，使证据基础中的模式和聚类更加清晰。其他研究则采用可视分析，以更有效地呈现研究结果，例如使用热图展示不同维度上的研究活动模式，使用桑基图说明参与者流向或干预路径[34, 35], ，以及使用时间线可视化展示时间趋势。然而，这些实现方式主要停留在描述层面，未能充分整合交互式探索功能，无法让用户根据自身兴趣动态筛选和分析证据基础。

2.4 贡献

尽管可视化工具、主题建模和可视化系统综述方法论取得了这些进展，但在理解如何将系统综述的严谨性与交互式知识图谱的探索能力相结合的有效策略方面，仍存在显著空白。本文提供了一种概念验证性交互式证据映射工具，用于探索结合以下要素的策略：(1) 传统综述的系统严谨性；(2) 由大语言模型驱动的主题建模所具备的模式发现能力；以及 (3) 交互式可视化的探索灵活性。我们的方法将经过系统编码的综述数据转化为动态知识图谱，支持多个层次的分析——从高层次的主题概览，到对单项研究的详细检查。通过保持自动化主题模型与底层研究特征之间的明确联系，我们的

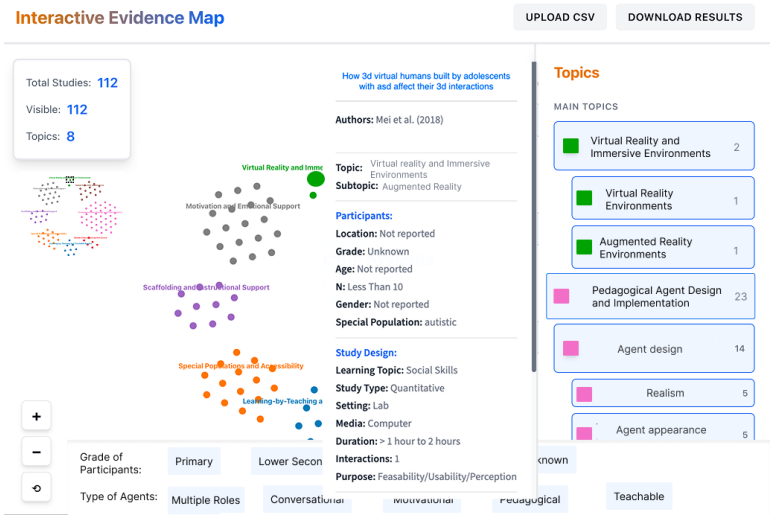


图1：证据地图工具的概念验证实现，包含一个信息与导航磁贴（左上）。该工具将论文可视化为节点，并用颜色表示建模主题；用户可以浏览建模主题和子主题（右侧面板）、选择编码条目（底部面板），还可以通过详情面板（中央）浏览单篇论文。

证据地图增强了透明度，满足了多样化的用户需求，并揭示出仅凭静态综合叙述可能无法呈现的模式。我们使用一项已发表的关于 K-12 教育中教学代理的范围综述数据来展示这一点， [36]将交互式证据图谱中的洞见与原综述中报告的洞见进行比较，以说明交互式证据映射所带来的互补价值和新的分析可能性。

3 试点实施

为了生成交互式证据地图，我们的集成大语言模型的流水线将经过系统编码的综述数据转化为可探索的可视化知识结构（图1）。流程始于数据准备：已根据预定义方案完成系统综述、提取并编码的研究，以结构化 CSV 格式上传至流水线。文件中的每一行代表一项研究，各列记录其关键特征，例如书目信息，以及系统综述特有的因素，如研究设计详情、参与者人口统计特征、干预、结果，以及原始综述者识别出的任何定性发现或主题。除书目信息外，所有内容均可根据综述者的发现和领域实践灵活定制。

在主题建模阶段，我们采用基于提示的方法，通过应用程序接口使用可免费访问的 Claude（Anthropic 的大语言模型） [37] （截至2026年1月），从综述数据中提取主题聚类。我们向 Claude 提供结构化的研究数据，并在提示中要求其根据研究摘要和标题识别连贯的主题。提示内容很简单：“分析这个数据集中的研究，并识别 *6–8* 个主要主题，以及每个主题下的 *2–3* 个子主题。”此处的“数据集”一词由编码数据替代。Claude 会返回主题标签和描述，并根据语义相似度将每项研究归入一个主题。这些归类结果和描述可供下载。我们在这一初步概念验证中采用基于提示的方法，因为它支持快速原型开发和迭代优化；不过，后续版本将采用预训练且可微调的基于API的模型，以便从更大型的科学数据集中自动、可扩展且透明地（尽可能）提取主题。

在数据结构化 阶段，我们将主题分配结果与原始编码研究特征相结合，创建统一的数据结构，将每项研究与其主题关联起来，同时保留所有系统提取的元数据。随后，将这些结构化数据转换为兼容交互式可视化库的格式。在可视化生成阶段，我们采用 D3.js 和基于 React 的组件，创建交互式证据地图界面。可视化包含多个相互协调的视图（见图1）：(1)（中央）空间布局，展示主题聚类，并以节点表示归入相应主题的各项研究；聚类的大小与研究数量成正比；(2) 筛选控件，允许用户按主题筛选研究（右侧），以及另一个面板，允许用户按编码特征筛选研究，例如发表年份、方法学、教育阶段或结果类型（底部）；(4) 详情面板，用户将鼠标悬停在单项研究上或选中单项研究时，显示该研究层面的信息（中央偏右）；(5) 动态统计数据，根据应用的筛选条件更新（左上角）；以及 (6) 导航地图，帮助用户在大型地图中进行空间探索。

我们最初的试点实施代码将在 GitHub 上公开。

4 概念验证：用于指导系统综述综合的证据地图

4.1 数据集

我们将 [36] 的范围综述中纳入的 112 项研究作为数据集。该综述对多种因素进行了编码，包括学习主题（例如科学、数学、社交技能、语言）、智能体类型（教学型、对话型、多角色）、年级（小学、初中、高中）、智能体角色（指导/支架、信息来源）、研究目的（实验、可行性/可用性）等。

尽管该数据集包含 112 项研究，且作者提供了完整的编码表，其中包含每项研究的数据，但要识别哪些交叉因素较为常见，以及存在哪些研究空白，仍是一项挑战。用户需要处理 CSV 中多个列的数据。诸如“对话代理在哪些领域被用于评估？”或“与用作信息来源的智能体相比，有多少研究涉及用于指导的教学型智能体？”之类的问题，以目前这种形式很难进行探究。

4.2 主题模型

如图1所示，交互式证据图谱从最外层视图开始，展示了为该数据集定义的八个主题。尽管提示非常笼统，但值得注意的是，生成的主题反映了证据基础中有意义的区别，例如方法学方法（如以教促学范式）、技术实现（如虚拟现实和沉浸式环境、教学代理设计与实现）、教学和理论框架（如支架和教学支持、学习成果和成就、动机和情感支持）、应用领域（如社会感知和人类代理交互）以及情境因素（如个体差异和学习者特征）。虽然这些建模主题与系统综述团队编码的维度并非完全正交，但与部分编码信息存在重叠。

4.3 数据探索

在图1右侧，主题筛选器可帮助缩小样本范围，底部面板则显示编码特征。该工具根据主题模型生成主题和子主题标签，并显示各类别中的研究数量，以便对不同主题和子主题进行定量比较。底部面板使用研究编码详情中的列标题填充。

用户可以放大任意聚类，直到查看聚类中的具体论文（图1中目前未显示），还可以点击节点，阅读单项研究的详情（图1右侧中央）。这些操作流程使用户能够探索提取出的主题与编码信息如何相互交织，从而补充人工审查流程。此外，这也让读者看到研究在哪些维度上聚集在一起，并根据主题和编码因素，帮助他们找出感兴趣的具体论文。提出“哪些研究讨论了通过教学代理来学习小学年级学生？”这类问题，可以让用户获取四层整合信息。首先，用户可以在主题筛选器（右侧面板）中选择通过教学学习，并在特征

筛选面板（底部）中选择“小学”选项。面板按钮会根据研究的编码情况变为启用或停用。用户可以统计研究数量，并了解其他交叉信息，例如这些研究的发表时间，以及研究采用了哪些环境（例如课堂环境）。此外，用户还可以放大视图，查看子主题，并了解每项突出显示的研究的详情。如果证据地图规模较大，导航面板可以帮助用户浏览。另外，如果某些论文被标记了多个主题，详情面板就会显示这些信息。

4.4 综合 K-12 教育教学代理的研究发现

交互式证据图谱支持发现多层次的模式，揭示了主题聚类或人工编码单独分析时都无法呈现的内容。

揭示隐藏的设计原则。探索“特殊人群与包容性设计”聚类时，研究识别出三个子主题：感知研究、社会外观研究和人机交互研究。前两个子主题包含大量比较设计变体的研究，其模式可通过人工编码观察到；相比之下，研究数量较少的人机交互子主题则揭示出一种连编码数据中或单独考察该子主题时都无法发现、却连贯一致的设计理念：针对自闭症谱系障碍学生的研究，无论学科内容如何，都始终采用以社交技能发展为重点的非人类代理形式。只有将孤立的研究按内容而非仅按预定义类别进行聚类时，这一跨领域原则才得以显现。

揭示时间演变。将主题与人工编码的时间和方法学维度交叉筛选后，显现出三个截然不同的研究时期。第一个时期（2003-2010年）侧重通过实验室实验验证效能，探究“它是否有效？”。第二个时期（2011-2018年）转向设计优化，系统地调整代理形式和角色，尤其是用于社交技能培养的对话代理，并以早期概念验证成果为基础。较近的时期（2019-2023年）则强调生态效度，通过在课堂中针对特定人群开展纵向部署，追踪学习者与长期充当教练的代理互动时产生的动机结果。从受控实验到自然情境实施的这一进程，反映了该研究领域更广泛的技术成熟过程。

揭示关键空白。双重筛选系统凸显了明显的缺失：仅有三项研究考察了面向高中高年级学习者的对话代理；针对青少年（10—19岁）的语言学习干预完全缺席；非人类代理设计则只出现在针对自闭症的社交技能干预中，这表明人们错失了将这类设计模式应用于神经典型人群或

其他学习领域的机会。相反，这张图谱显示，小学科学教育中，在课堂环境下采用教学代理的研究已趋于饱和；这意味着，在这一配置下，逐步增加贡献的回报正在递减，而尚未充分探索的组合（例如中学阶段语言学习、面向青少年的对话系统）仍有很大的研究空间。该工具让研究人员能够在内容驱动的主题集群与方法学筛选条件之间灵活切换，从而将系统综述从静态摘要转变为分析工作空间；借助交互式探索，跨维度模式、时间轨迹和战略性研究空白得以立即呈现，无需费力进行人工综合。

5 讨论与后续步骤

我们实现的双重筛选架构中，大语言模型根据标题和摘要生成主题，在揭示新兴主题模式的同时，人工编码类别则支持按方法学维度进行筛选，从而解决了传统系统综述的一项关键局限。不过，还有若干尚待解决的问题值得进一步研究：

1. 可扩展性与性能当证据库中的研究超过 500–1000 项时，主题模型连贯性和可视化响应速度都需要通过实证验证。2. 文本粒度需要权衡基于摘要的聚类（上下文有限，但普遍可用）与全文分析（语义信息更丰富，但需要额外的数据准备和计算资源）。

3. 概率分配透明度目前的实现会将每项研究分配给单个主题—子主题对，却不展示置信度分数、其他候选主题或共现模式。后续版本应将分配不确定性可视化，并支持多标签视图，以便合理呈现跨越多个主题聚类的研究。

4. 大语言模型输出的稳定性与可重复性这是系统综述尤为关注的问题，因为系统综述要求方法可复现。基于提示词的大语言模型的黑箱行为意味着，聚类分配和标签在不同运行中可能存在显著差异；生成过程中的随机抽样也可能损害证据综合所要求的可重复性标准。

5. 跨领域普适性，测试该方法能否有效迁移到采用不同编码惯例的领域中的系统综述，

例如医疗干预、政策评估和工程解决方案。

展望未来，我们设想构建一个采用基于 API 的架构的开源平台，支持可插拔的主题模型（包括带温度控制和种子参数、可实现部分稳定性的大语言模型，以及 SciBERT 和 Specter2 等确定性替代方案）、用于量化多次运行间聚类稳定性的模型比较指标，以及可自定义的可视化布局。研究人员可通过图形界面上传编码数据、配置分析参数、将大语言模型的输出与传统主题建模基线进行比较，并生成可供发表的交互式证据地图。总之，我们希望帮助研究人员将静态的 CSV 表格转化为可探索的研究图景，在人工智能驱动的聚类所带来的发现潜力与系统综述方法学所要求的可靠性之间取得平衡。

6 结论

系统综述通常旨在提供高层次的综合分析，以得出可推广的结论。为此，人们很容易无意间过度概括结果，或未能充分说明样本的描述性特征；已有研究指出，需要探索新的方式，根据读者的需求可视化数据 [9, 10, 7, 11, 12, 5]。我们基于语言模型和知识图谱绘制技术构建的交互式证据地图概念验证，凸显了其在探索样本特征方面的明显优势。我们认为，交互式证据地图有望作为传统系统综述报告规范的补充一并报告。

资金来源

作者 Mokshagna Kadiyala 和 Neelesh Yaddanapudi 获得了北卡罗来纳州立大学“工程你的体验”计划本科生研究经历的资助。本材料基于美国国家科学基金会资助项目 DUE-2518159 支持的工作，以及美国国家科学基金会和教育科学研究院资助项目 DRL-2229612 支持的工作。本材料中表达的任何意见、发现、结论或建议均属于作者，不一定反映美国国家科学基金会或美国教育部的观点。

参考文献

[1] Mark Newman 和 David Gough。《教育研究中的系统综述：方法论、视角与应用》。载于 Olaf Zawacki- Richter、Michael Kerres、Svenja Bedenlier、Melissa Bond 和 Katja Buntins 编，《《教育研究中的系统综述：方法论、视角与应用》》，第 3–22 页。Springer Fachmedien，威斯巴登，2020年。[2] Lijia Chen、Pingping Chen 和 Zhijian Lin。《教育人工智能：综述》。*IEEE Access*，8:75264–75278，2020年。[3] Florence Martin、Vanessa P. Dennen 和 Curtis J. Bonk。新兴学习环境与技术系统综述研究综述。*Educational Technology Research and Development*，68(4):1613–1633，2020年8月。[4] Daniel Keim、Gennady Andrienko、Jean-Daniel Fekete、Carsten Görg、Jörn Kohlhammer 和 Guy Melançon。可视分析：定义、流程与挑战。信息可视化：以人为中心的议题与视角，第 154–175 页，2008年。[5] Camila Vieira、Paul Parsons 和 Vetria Byrd。教育数据的可视化学习分析：系统性文献综述与研究议程。*Computers & Education*，122:119–135，2018年。[6] Michael Sedlmair、Miriah Meyer 和 Tamara Munzner。设计研究方法论：来自一线与文献的反思。*IEEETransactions on VisualizationandComputerGraphics*，18(12):2431–2440，2012年。[7] Michael Kossmeier、Ulrich S. Tran 和 Martin Voracek。荟萃分析和系统综述图形展示领域概览：全面综述、分类体系与特征分析。*BMCMedical Research Methodology*，20(1):26，2020年2月。[8] Stuart Card。信息可视化。载于人机交互，第 199–234 页。CRC press，2009年。[9] James Thomas 和 Angela Harden。定性研究综合方法：批判性综述，第 8 卷。BioMed Central，2008年。

[10] Elaine Beller、Paul Glasziou、Douglas G Altman、Sally Hopewell、Hilda Bastian、Iain Chalmers、Peter C Gøtzsche、Toby Lasserson 和 David Tovey。荟萃分析的交互式可视化。系统综述，7 (1):1–10, 2018年。

[11] Rhiannon Owen、Naomi Bradbury、Yiqiao Xin、Nicola Cooper 和 Alex Sutton。MetaInsight：一款基于网页的交互式工具，利用 R-shiny 和 netmeta 分析、探究和可视化网状荟萃分析。*Research Synthesis Methods*，10(4):569–581，2019年。[12] Beat A Schwendimann、María Jesús Rodríguez-Triana、Andrii Vozniuk、Luis P Prieto、Mina Shirvani Boroujeni、Adrian Holzer、Denis Gillet 和 Pierre Dillenbourg。一览学习情况：学习仪表板研究的系统性文献综述。*IEEE Transactions onLearning Technologies*，10(1):30–41，2017年。[13] Peter Kraker、Christopher Kittel 和 Asura Enkhbayar。开放知识地图：基于自然语言处理创建通往全球科学知识的可视化界面。载于*027.7 Zeitschrift fürBibliothekskultur/Journal for Library Culture*，第4卷，98–103页，2016年。[14] Nees Jan Van Eck 和 Ludo Waltman。软件综述：VOSviewer，一款用于文献计量映射的计算机程序。*Scientometrics*，84(2):523–538，2010年。[15] Alberto J Cañas、Graham Hill、Roger Carff、Niranjan Suri、James Lott、Gloria Gómez、Thomas C Eskridge、Mario Arroyo 和 Rodrigo Carvajal。CmapTools：知识建模与共享环境。载于概念图：理论、方法学、技术。第一届国际概念图会议论文集，第1卷，125–133页，2004年。[16] Bilal Abu-Salih 和 Salihah Alotaibi。教育领域知识图谱构建与应用的系统性文献综述。*Heliyon*，10(3):e25383，2024年2月。[17] Katy Börner、Chaomei Chen 和 Kevin W Boyack。知识领域可视化。*Annual Review of Information Science and Technology*，37(1):179–255，2003年。[18] Katy Börner。科学图谱：可视化我们所知。MIT出版社，2010年。[19] Martin J Eppler 和 Remo A Burkhard。知识可视化。载于知识管理：概念、方法论、工具与应用，781–793页。IGI Global Scientific Publishing，2008年。[20] Nees Jan Van Eck 和 Ludo Waltman。文献计量网络可视化。学术影响力测量：方法与实践，285–320页，2014年。

[21] Li tmaps。 Litmaps (version 2025年9月16) , 2024. [Search tool] on .[22] Jason Priem、Heather Piwowar 和 Richard Orr。《OpenAlex：一个完全开放的学术作品、作者、发表场所、机构及概念索引》。*arXiv* 预印本 *arXiv:2205.01833*，2022年。[23] David M Blei、Andrew Y Ng 和 Michael I Jordan。潜在狄利克雷分配。《机器学习研究杂志》，3(Jan):993–1022，2003年。[24] Iz Beltagy、Kyle Lo 和 Arman Cohan。SciBERT：一种面向科学文本的预训练语言模型，2019年9月。arXiv:1903.10676 [cs]。[25] Amanpreet Singh、Mike D'Arcy、Arman Cohan、Doug Downey 和 Sergey Feldman。SciRepEval：科学文档表示的多格式基准测试，2023年11月。arXiv:2211.13308 [cs]。[26] Scott Deerwester、Susan T Dumais、George W Furnas、Thomas K Landauer 和 Richard Harshman。基于潜在语义分析的索引编制。《美国信息科学学会杂志》，41(6):391–407，1990年。[27] Maarten Grootendorst。BERTopic：基于类别的 tf-idf 过程进行神经主题建模。*arXiv* 预印本 *arXiv:2203.05794*，2022年。[28]Shan Zhang、Chris Palaguachi、Marcin Pitera、Chris Davis Jaldi、Noah L Schroeder、Anthony F Botelho 和 Jessica R Gladstone。范围综述流程的半自动化：是否值得？一项方法学评估。《教育心理学评论》，36(4):131，2024年。[29] Mauricio Marrone 和 Martina K. Linnenluecke。跨学科研究地图：一种可视化研究主题的新技术。*PLOSONE*，15(11):e0242283，2020年11月。[30]Manuel J Cobo、Antonio Gabriel López-Herrera、Enrique Herrera-Viedma 和 Francisco Herrera。科学制图软件工具：综述、分析及工具间的协作研究。《美国信息科学学会与技术学会杂志，62(7):1382–1402，2011年。[31] Sara Lewis 和 Mike Clarke。森林图：试图既见树木，也见森林。《英国医学杂志》，322(7300):1479–1480，2001年。[32] Birte Snilstveit、Martina Vojtkova、Ami Bhavsar、Jennifer Stevenson 和 Marie Gaarder。证据与空白地图：促进循证政策和战略研究议程的工具。《临床流行病学杂志，79:120–129，2016年。

[33]Jaziar Radianti、Tim A Majchrzak、Jennifer Fromm 和 Isabell Wohlgenannt。沉浸式虚拟现实在高等教育中的应用系统综述：设计要素、经验教训与研究议程。*Frontiers in Virtual Reality*，2:721933，2021年。

[34] Rebecca E. Glover、Mustafa Al-Haboubi、Mark P. Petticrew、Elizabeth Eastmure、Sharon J. Peacock 和 Nicholas Mays。桑基图有助于阐明“证据流失”：快速诊断检测对抗微生物药物耐药性的有效性的系统综述与荟萃分析。*Journal of Clinical Epidemiology*，144：173–184，2022年4月。[35] Neal R. Haddaway、Andrew Feierman、Matthew J. Grainger、Charles T. Gray、Ezgi Tanriver-Ayder、Sanita Dhaubanjar 和 Martin J. Westgate。EviAtlas：用于可视化证据综合数据库的工具。环境证据，8(1)：22，2019年6月。[36] Shan Zhang、Chris Davis Jaldi、Noah L Schroeder 和 Jessica R Gladstone。K-12 教育中的教学代理：一项范围综述。*Journalof Research on Technology in Education*，57(6)：1230–1257，2025年。

[37] Anthropic。Claude 3.5 sonnet，2024年。大语言模型。