

中图法分类号: TP391; TP18 文献标识码: A 文章编号: 1006-8961(2026)06-2198-24

论文引用格式: Wang Y H, Cao N, Chen S M, Li C H, Zeng W, Tao J, Zeng Q, Wang C B and Zhang J W. 2026. Survey on large model-empowered visualization and visual analytics. Journal of Image and Graphics, 31(6):2198-2221(汪云海, 曹楠, 陈思明, 李晨辉, 曾伟, 陶钧, 曾琼, 王长波, 张加万. 2026. 大模型赋能的可视化与可视分析研究综述. 中国图象图形学报, 31(6):2198-2221)[DOI:10.11834/jig.260046]

## 大模型赋能的可视化与可视分析研究综述

汪云海<sup>1</sup>, 曹楠<sup>2</sup>, 陈思明<sup>3</sup>, 李晨辉<sup>4</sup>, 曾伟<sup>5</sup>, 陶钧<sup>6</sup>, 曾琼<sup>7</sup>, 王长波<sup>8</sup>, 张加万<sup>9\*</sup>

1. 中国人民大学信息学院, 北京 100872; 2. 同济大学设计创意学院, 上海 200092; 3. 复旦大学大数据学院, 上海 200433;
4. 华东师范大学计算机科学与技术学院, 上海 200062; 5. 香港科技大学(广州)计算媒体与艺术学域, 广州 511453;
6. 中山大学数据科学与计算机学院, 广州 510000; 7. 山东大学计算机科学与技术学院, 青岛 266237;
8. 华东师范大学数据科学与工程学院, 上海 200062; 9. 天津大学软件学院, 天津 300072

**摘要:** 数据可视化作为连接人类认知与数据科学研究的重要技术基础, 随着大模型的发展迎来范式转型。传统数据可视化主要依赖人工设计的视觉编码规则、图形语法与交互机制, 通过显式映射和操作支持数据探索与信息传达。然而, 面对日益增长的数据规模、多样化的数据类型以及复杂的分析与决策场景, 基于静态图形和参数化交互的传统方法在效率、表达能力和语义支持方面逐渐显现局限。近年来, 大规模预训练语言模型、多模态基础模型及数据智能体的兴起, 为数据可视化的自动生成、智能分析和交互优化提供了新的技术动力。本文围绕可视化基础理论、可视分析、可视化叙事与可视化评估4个方面, 对大模型赋能下的数据可视化研究进展进行系统综述, 从基础理论层面看, 大模型通过其强大的语义理解与推理能力, 推动可视化从低层视觉编码向语义驱动的表达与感知建模演进。在可视分析层面, 大模型与数据代理的结合使分析流程从以人为主导的工具操作, 转向人—模型—知识协同的混合智能模式。在叙事可视化方面, 大模型显著降低了数据叙事的创作门槛, 使系统能够自动构建叙事结构、整合图文内容, 并根据受众和情境动态调整表达方式。在可视化评估方面, 大模型在图形质量评估与设计建议生成方面展现出潜力。本文分析当前面临的关键问题与发展趋势, 为大模型时代的数据可视化研究与系统设计提供结构化的理论支撑。

**关键词:** 可视化基础理论; 可视化交互; 可视化分析; 可视化叙事; 大模型

### Survey on large model-empowered visualization and visual analytics

Wang Yunhai<sup>1</sup>, Cao Nan<sup>2</sup>, Chen Siming<sup>3</sup>, Li Chenhui<sup>4</sup>, Zeng Wei<sup>5</sup>, Tao Jun<sup>6</sup>,Zeng Qiong<sup>7</sup>, Wang Changbo<sup>8</sup>, Zhang Jiawan<sup>9\*</sup>

1. School of Information, Renmin University of China, Beijing 100872, China; 2. College of Design and Innovation, Tongji University, Shanghai 200092, China; 3. School of Data Science, Fudan University, Shanghai 200433, China;
4. School of Computer Science and Technology, East China Normal University, Shanghai 200062, China; 5. Thrust of Computational Media and Arts, Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511453, China;
6. School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510000, China; 7. School of Computer Science and Technology, Shandong University, Qingdao 266237, China; 8. School of Data Science and Engineering, East China Normal University, Shanghai 200062, China; 9. Software College, Tianjin University, Tianjin 300072, China

收稿日期: 2026-01-21; 修回日期: 2026-02-03; 预印本日期: 2026-02-12

\* 通信作者: 张加万 jwzhang@tju.edu.cn

基金项目: 国家自然科学基金项目(62132017, 62472099, 62572191, 62572415, 62372271)

Supported by: National Natural Science Foundation of China(62132017, 62472099, 62572191, 62572415, 62372271)

**Abstract:** Data visualization, as a fundamental technology supporting human cognition and data-driven scientific reasoning, is undergoing a profound shift driven by the rapid development of large-scale models and is attracting widespread attention from the academia and industry. For decades, traditional visualization approaches have mainly relied on manually or automatically designed visual encoding rules, graphical grammars, and human-computer interaction techniques. Through explicit visual mappings, graphical composition, and interface operations, these methods enable users to observe patterns, understand structures, and communicate insights from data. However, with the continuous growth of data size, task complexity, and decision-making contexts, statistical visual mappings and parameter-driven interaction are increasingly insufficient to support modern analytical needs. Instead, users now demand more than visual presentation alone; they require systems capable of semantic understanding, task-driven reasoning, and cross-modal information integration. The emergence of large models—particularly large language models, vision-language foundation models, and intelligent agents—has introduced unprecedented technological momentum to the visualization field. These models are not only transforming how visualizations are generated and interacted with but also reshaping the theoretical foundations, analytical workflows, and narrative practices of data visualization. Their strong capabilities in semantic representation, abstraction, and reasoning enable visualization systems to move beyond surface-level depiction toward deeper support for analytical intent and knowledge construction. From a theoretical perspective, visualization research has long been grounded in the Grammar of Graphics and declarative specification languages, which provide highly abstract and formalized descriptions of graphical structures, data mappings, and interaction logic. The introduction of large models does not replace this; instead, it reinforces its importance. Declarative grammars increasingly function as an interpretable “intermediate language” between natural-language intent and executable visual representations, enabling large models to translate high-level analytical goals reliably into controlled, consistent visualization specifications. This mediation is critical for ensuring the interpretability, reproducibility, and controllability of automatically generated visualizations. Moreover, the semantic reasoning capabilities of large models allow them to infer the underlying intent and logic behind visual layouts, color encodings, and spatial structures, and pushes visualization research from low-level perceptual encoding toward semantic-driven visual understanding and intent modeling. At the same time, techniques such as differentiable rendering, neural implicit representations, and Gaussian splatting provide new frameworks for high-fidelity scientific rendering, continuous data representation, and optimization in parameterized spaces. These methods enable visualization systems to become differentiable, learnable, and optimizable, and allows tighter integration between visual representation, computational models, and analytical objectives. As a result, visualization is increasingly situated within richer signal and representation spaces, supporting adaptive, expressive, and controllable visual analysis. Beyond theoretical advances, large models are fundamentally changing how users collaborate with visualization systems. Traditionally, effective visualization required users to master visualization grammars, tool-specific operations, and data processing workflows. By contrast, large-model-driven systems allow users to express analytical goals, design preferences, and data-related questions directly through natural language. The system can then interpret user intent, generate appropriate visualizations, restructure data views, or optimize visual encodings accordingly. This fusion of model-based semantic knowledge with formal visualization representations forms the foundation of a new generation of intelligent visualization systems. At the level of visual analytics, large models and agents are driving a transition from human-centered interaction toward a hybrid collaborative paradigm involving humans, agents, and knowledge. Early machine-learning-for-visualization research primarily focused on addressing isolated subtasks, such as view recommendation, feature detection, or layout optimization. By contrast, visualization frameworks based on large language models (LLMs) leverage unified semantic understanding across tasks and modalities to support end-to-end analytical pipelines. These systems can assist with visualization generation, pattern discovery, and reasoning-based explanation, and enable more holistic support for complex analytical processes. As human-machine collaboration evolves from simple question answering toward knowledge generation, large models increasingly function as proactive cognitive collaborators rather than passive assistants. They can produce trend interpretations, support hypothesis testing, and summarize underlying mechanisms while continuously modeling user behavior, analytical state, and task context. This approach enables a shift from reactive assistance to mixed-initiative collaboration, where systems actively guide users through complex, dynamic, and uncertain analytical environments, augmenting human reasoning and sensemaking capabilities. In the domain of narra-

tive visualization, large models mark the beginning of a new era of intelligent content generation. Traditional narrative visualization requires expertise in design, storytelling, writing, and programming, which makes producing high-quality narratives costly and time consuming. Large models significantly lower this barrier by automatically identifying data themes, extracting salient trends, constructing narrative structures, and integrating visual, textual, and multimodal elements into coherent, stylistically consistent narratives. This capability substantially improves efficiency and accessibility and enables end-to-end automation from data to story in applications such as data journalism, educational visualization, science communication, and business reporting. With the maturation of multimodal generation technologies, narrative processes can further incorporate natural-language interaction, dynamically adapt narrative perspectives, and tailor content to audiences with different backgrounds and goals. In visualization evaluation, large models demonstrate promising potential due to their understanding of visual aesthetics, layout quality, perceptual principles, and readability. They can automatically assess visualization quality, detect misleading encodings, propose design improvements, and generate multiple design alternatives for comparison. However, model-based visual judgment is not inherently equivalent to human perception, a situation raising critical research challenges. Ensuring alignment between model assessments and human visual cognition, mitigating hallucinations and biases, and improving transparency, reliability, and interpretability remain central issues in large-model-driven visualization research. To organize the rapidly evolving landscape of large-model-powered data visualization systematically, this paper presents a comprehensive survey from four perspectives: visualization fundamentals, visual analytics, narrative visualization, and visualization evaluation. This paper reviews representative research advances in each area, analyzes emerging technologies enabled by large models, and discusses key challenges and future research directions. By providing a structured knowledge map and theoretical framework, this survey aims to support future innovation in intelligent visualization systems and contribute to the development of data visualization as a critical bridge between human intelligence and artificial intelligence in the era of large models.

**Key words:** fundamental theories of visualization; visualization interaction; visual analytics; visual storytelling; large language models; agents

## 0 引言

数据可视化作为连接人类认知与数据科学推理的核心技术,正迎来由大模型驱动的深刻范式变革,受到了国内外学者的广泛关注。传统可视化长期依赖人为设定的编码规则、图形语法和人机交互模式,通过设计映射、图形组合和界面操作帮助用户观察、理解并传达数据。然而,随着数据密度、任务复杂性和决策情境的增长,静态图形与参数式操作界面逐渐难以满足分析需求,用户需要的已经不仅是数据呈现本身,而是更高层次的语义理解、任务驱动的推理以及跨模态的信息表达。在这一背景下,大模型,尤其是大语言模型(large language model, LLM)、视觉语言模型(vision-language model, VLM)与数据智能体的崛起,为数据可视化带来前所未有的技术动力,并正在重塑可视化的理论基础、交互方式、分析流程与叙事表达。

从理论层面看,可视化长久以来建立在图形语法与声明式规范语言之上,通过高度抽象的视觉描

述模型实现对图形结构、映射关系与交互逻辑的形式化定义。大模型的引入并未替代这一体系,反而使其变得更加关键:作为自然语言与可视化表达之间的“中间语”,声明式语法使大模型能够可靠地将高层语义映射为可执行的可视化表达,保证自动生成图表的可控性与一致性。同时,大模型带来的语义推理能力,使其能够理解图形布局、色彩关系与空间结构背后的“意图”与“逻辑”,推动可视化从低层编码推理迈向基于语义的视觉感知建模。与此同时,可微分渲染、隐式神经表示、高斯泼溅等新型表示方法不断强化可视化的数学基础,为高保真渲染、连续数据表达和参数空间优化提供新的计算框架,使可视化在更丰富的信号空间与表示空间中获得可微、可控与可学习的能力。另一方面,大模型正在改变人与可视化系统的协作方式。过去,用户必须掌握可视化语法、工具操作与数据处理流程,才能完成有效表达;而如今,用户可以通过自然语言描述分析目标、设计偏好或数据疑问,由模型自动解读意图并生成相应图表、重构数据视图或优化视觉编码。由此,大模型提供的语义知识与可视化的形式化表达

开始融合,构成新一代智能可视化的理论底层。

在可视分析层面,大模型与数据代理(Agent)正在推动从“以人为中心”转向“人一Agent—知识”三方协同演化。早期机器学习赋能可视化(machine learning for visualization, ML4VIS)主要解决可视化中的具体子任务,如视图推荐、特征检测或布局优化,而大语言模型赋能可视化(large language models for visualization, LLM4VIS)则凭借跨任务、跨模态的统一语义理解能力,能够深度赋能数据治理、图表生成、洞察发现及推理解释等完整可视分析流程。随着人机协同从单纯的图表生成延伸至知识生成,模型不仅能实现数据可视化,还可输出趋势解释、假设验证与机制归纳,同时持续感知用户的行为模式、分析状态与任务上下文,推动协同模式从“问答式”辅助升级为具备主动性的混合主动协作,成为助力研究者破解复杂系统的“认知合作者”。

在叙事表达层面,大模型为叙事可视化带来智能生成的新阶段。传统叙事可视化依赖设计、编排、写作与编码多项能力,而大模型能够自动识别数据主题、抽取趋势、构建叙事逻辑,将图表、文本与多模态内容整合为结构清晰、风格一致的叙事作品。这不仅降低了专业门槛,也极大提升了内容创作效率,使数据新闻、教育可视化、科普传播与商业报告等领域出现了从数据到故事的“全流程智能化”。进一步地,随着多模态生成技术的成熟,叙事过程可与用户

自然语言交互,系统可根据语境动态生成图文混排内容、自动调整叙事视角并适配不同受众的知识背景。

在可视化评估层面,大模型对图形美学、布局质量、视觉可读性与感知规律的理解,使其能够自动评估图表质量、检测误导性编码、提出改进建议甚至生成多版本设计供用户选择。然而,模型的感知机制并非完全等同于人类视觉系统,这带来新的研究挑战——如何确保模型的视觉判断与人类一致?如何校准模型的幻觉与偏差,使其建议可靠且可解释?这些问题构成大模型时代可视化交互研究的核心议题。

陶钧等人(2023)从数据管理、可视化创建、交互探索和可视分析的角度探讨人工智能方法在可视化流程中的应用和发展;任洋甫等人(2024)探究了在不同沉浸式环境和多场景中视觉提示信息的研究前景与发展方向。与前人工作不同,本文系统梳理大模型驱动下的数据可视化发展脉络,从可视化基础理论、可视分析、可视化叙事和可视化评估4个方面展开综述,重点讨论大模型赋能可视化与可视分析的国内外研究进展,分析当前存在的核心挑战及未来发展趋势,整体结构如图1所示。本综述旨在为大模型时代的数据可视化研究提供清晰结构化的知识图谱与理论框架,为后续的技术创新与系统设计奠定基础。

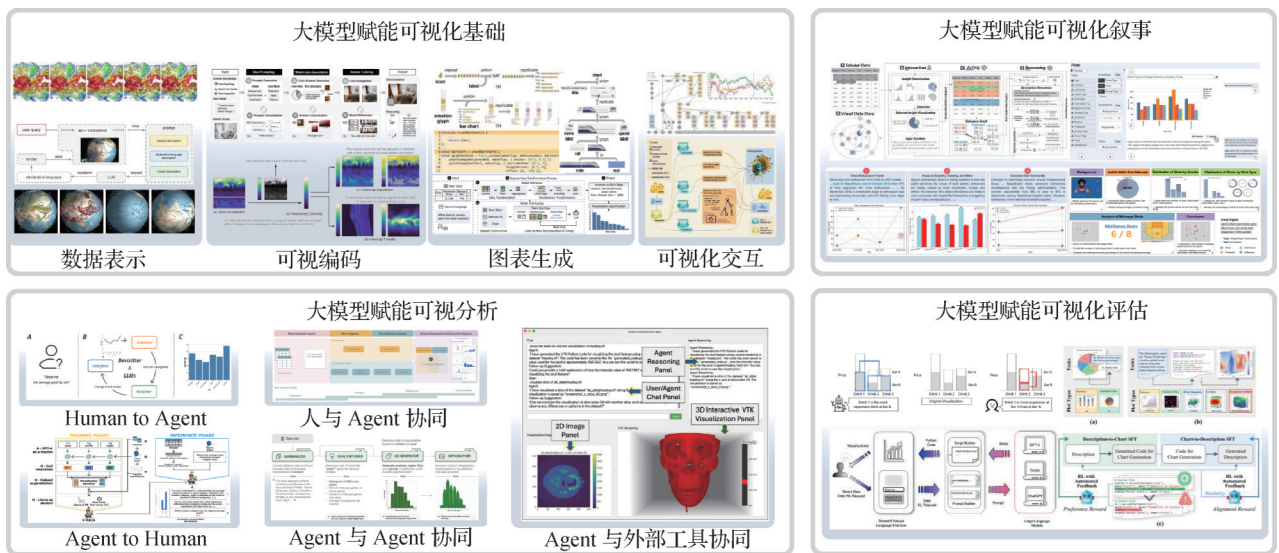


图1 本文框架概览

Fig. 1 Overview of this survey

## 1 国际研究现状

### 1.1 可视化理论基础

本节从可视化基础理论出发,系统梳理国际范围内数据表示、可视编码、语义驱动的图表生成、可视化交互等方向的最新进展、发展趋势以及主要挑战。

在一般的数据可视化任务中,表示学习已广泛应用于抽取数据潜在特征,其核心在于将原始数据映射至更为紧凑或语义相关的空间,以支持高效智能的可视呈现、分析与推理。例如,利用图神经网络为社交网络数据生成节点和子图的隐向量,从而支持社区结构可视化与交互探索(Song等,2022);通过序列建模提取事件流或时间序列中的关键模式,辅助用户探索数据趋势(Li等,2020a);使用语言模型提取文本流的语义表示,为文档可视化和主题演变分析提供辅助信息(Qiu等,2024)。科学可视化的研究任务近年逐渐从规则数据转向非结构化数据、复杂场景与多物理场耦合,隐式神经表示(implicit neural rendering, INR)迅速成为国际研究热点。INR以神经网络直接建模连续场,实现任意分辨率的采样与渲染。美国范德堡大学、德国慕尼黑工业大学和美国圣母大学等团队在标量场压缩、体绘制加速与多任务统一架构方面持续推进INR的边界(Lu等,2021;Weiss等,2022;Han和Wang,2023a);与此同时,美国多所高校通过超网络、哈希编码与分布式训练等技术显著提升INR的训练效率,推动神经表示逐渐成为国际科学可视化研究的主流方向。

与数据表示方面的发展速度相比,国际社会在视觉感知建模与智能视觉编码方面的突破尚处于起步阶段。在视觉感知层面,研究人员提出利用多模态大模型评估人类图表理解、视觉错觉等感知任务(Wang等,2025d;Mao等,2025),发现尽管模型能够表现出类人的感知偏差,但在空间布局、颜色区分等方面仍存在显著差距。部分研究构建了颜色感知评测基准(Salim等,2026),用以系统分析不同模型在颜色语义识别任务中的表现,揭示了模型在基础视觉感知上的局限。在视觉编码与设计优化方面,学者们尝试通过语言与视觉信号的协同学习,实现从语义概念到视觉表现的自动映射;例如,Butt通过语言提示控制配色与形状生成,实现了对视觉元素的

精细化操控(Butt等,2025)。此外,部分观点性研究提出将多模态基础模型视做“视觉感知裁决者”,用于评估并指导可视化设计质量(Berger和Liu,2024)。上述研究为构建具有人类感知特征的智能可视化系统奠定了基础。

语义驱动的可视化生成国际研究从规则驱动以及数据驱动的方法,逐渐走向大模型驱动的图表生成。早期规则驱动方法依赖人工定义的语法或可视化规则构建自然语言到图表结构的映射,如NL2Viz(Wu等,2022)与Charticulator(Ren等,2019),但普遍面临规则维护成本高与泛化能力弱的问题。随着深度学习发展,数据驱动方法(如NL4DV(Narechania等,2021))通过端到端模型学习隐式映射,并推动Lux(Lee等,2021)等系统在可视化推荐方面取得进展,但仍依赖大量标注数据且难以适应开放场景。多模态大语言模型推动了NL2Vis的第3次跃迁,LIDA(Dibia,2023)、ChatGPT代码解释器(OpenAI)、Chat2Vis(Maddigan和Susnjak,2023)能够根据自然语言自动生成可视化代码并保持对话一致性;DirectGPT(Masson等,2024)与DynaVis(Vaithilingam等,2024)通过点击、拖拽等方式实现混合模式交互,降低对提示词的依赖;Charagraph(Masson等,2024)支持从文本文件中动态生成交互式图表与注释,C<sup>2</sup>框架(Koh等,2025)通过自校正增强结果稳定性,多智能体协作方法如METAL(Li等,2025a)将图表生成拆解为协作性子任务以优化最终质量;ChartLlama(Han等,2023b)与Text2Chart31(Zadeh等,2024)展示了大模型在图表问答、描述生成与要素抽取任务中的跨任务泛化能力。总体而言,国际研究正形成以大模型为核心的自动化、多模态协同与迭代优化的新思路,但仍面临上下文感知不足、数据转换逻辑不稳定与缺乏主动反思机制等挑战。

在可视化交互领域,国际研究正从传统的声明式规范语言向更具结构性与语义表达能力的新范式演进。Vega与Vega-Lite(Satyanarayan等,2016,2017)等体系仍是影响力最大的声明式语言,但主要聚焦于“数据到视觉”的映射,难以全面描述复杂图形结构与动态交互语义,因此促使研究者探索以图形结构和可操纵语义组件(manipulable semantic components, MSC)为核心的统一对象模型,以增强声明式系统的交互表达能力。与此同时,深度学习和大模型的引入显著拓展了交互分析的边界:早期方

法利用神经网络进行变量关联评估与集成一致性判断(Han等, 2021; He等, 2020), 或通过代理模型加速参数空间探索(Chen等, 2025c; Shen等, 2025a; Shi等, 2022; Shi等, 2023); 近期则进一步发展出以自然语言为媒介的语义交互, 如多智能体协同定位兴趣区域(Jia等, 2025)、声明式语法结合高斯泼溅实现风格与视角控制(Tang等, 2026)、基于语言—图像对齐的可视化传输函数参数生成(Jeong等, 2024), 以及基于大语言模型多智能体以及可编辑三维高斯

泼溅的体数据可视化交互(Ai等, 2026), 如图2所示。这一系列研究表明, 可视化交互正从传统事件驱动走向语义感知与智能协同, 人机协作方式正在被系统性重塑。

综上, 模型正在推动国际可视化研究从经验驱动转向智能化与可解释的下一代范式, 在数据表示、视觉编码、图表生成与可视化交互方面都取得了重要进展, 但在视觉感知理论、人机协同推理、多模态交互解析与高复杂度场景泛化等方向仍面临诸多关键挑战。

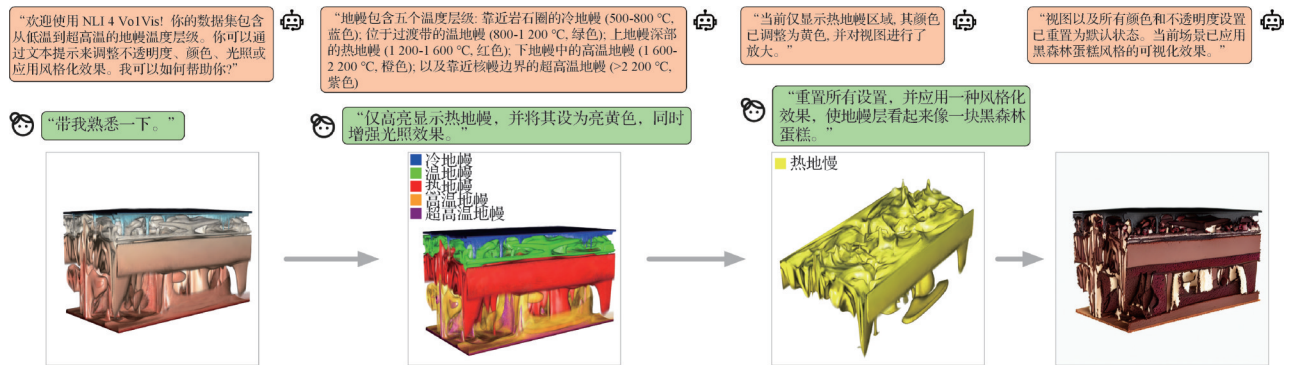


图2 基于自然语言的体数据可视化系统NLI4VolVis在地幔温度场数据集上的交互示例(Ai等, 2026)

Fig. 2 An interaction example of the natural-language-based volume visualization system (NLI4VolVis) using the mantle temperature dataset (Ai et al., 2026)

## 1.2 大模型赋能可视分析

国际上围绕大模型赋能可视化协同分析的研究呈现出高度多元化的态势, 涵盖理论框架构建、人机交互范式探索以及系统实现等多个层面。研究不仅关注可视化工具本身的智能化, 还强调人与智能体、智能体之间以及智能体与外部系统的多层协同, 以提升复杂分析任务中的效率和可解释性。在理论层面, 学者们提出了多种支持协同分析的理论模型, 例如基于协同—自适应的框架(Sperrle等, 2021)以及以人为中心的人工智能(human-centered artificial intelligence, HCAI)理念(Elmqvist和Klokrose, 2025), 前者强调系统与用户在任务执行过程中的持续适应和互动, 后者强调在智能决策过程中保持用户控制权与理解能力。这些理论为设计能够在复杂分析场景下实现高效、可控和个性化交互的系统提供了指导。

在人与Agent协同(human-agent collaboration)方面, 研究可分为3个关键维度: 用户意图的理解(human-to-agent)、智能体对用户的指导与推荐(agent-to-human)以及人与Agent的协同 workflows。在

human-to-agent层面, 早期的研究主要依赖LLM进行自然语言接口(natural language interface, NLI)的解析。Maddigan和Susnjak(2023)展示了如何通过提示工程缓解自然语言表达的歧义性, 实现端到端的图表生成方案。然而, 由于用户请求中常常包含模糊性和不确定性, 系统在理解意图和推断可视化任务方面仍面临挑战。研究者们因此提出了评估LLM处理模糊查询能力的新型基准测试, 例如Bako等人(2024)针对LLM在不确定情境下的语义理解和数据上下文识别进行了系统评测, Luo等人(2026)则发布了nvBench2.0, 专门用于评估系统在处理模糊查询时的性能。为应对模糊请求, 一些方法采用多路径推理和反馈驱动的优化框架, 通过生成多个候选可视化并进行评估, 最终合成最优结果(Seo等, 2026), 但这种全自动化方法的局限在于缺少用户干预的灵活性。

为进一步提升human-to-agent的准确性, 研究逐步引入了迭代交互和多模态输入。迭代协同范式使Agent能够基于上下文主动辅助用户, 例如在对话过程中监控生成的可视化以检测潜在问题(Choi等,

2024)。为突破纯文本表达的限制,研究者开始结合VLM处理图表和界面视觉提示(Chen等,2025b),同时将语音输入作为低延迟、高精度的交互方式与可视化探索结合(Jia等,2025)。在洞察提取和组织阶段,Gathani等人(2024)提出用户可直接操作图表元素来指定和完善数据洞察,系统则基于用户的选择提供辅助建议。更进一步,以用户为中心的自适应系统开始出现,通过评估用户的专业水平、历史分析行为及眼动数据,实现交互策略、可视化设计和推理过程的动态调整(Schnizer和Mayer,2025)。

在agent-to-human方向,研究关注智能体如何以及何时向用户提供指导和推荐。复杂可视化工具的操作复杂性可能阻碍用户有效达成分析目标,Ceneda等人(2017)对操作指导进行了系统化分类,包括系统为支持用户提供的指导以及用户为支持系统优化提供的反馈。在实际应用中,可视化推荐是最主要的指导形式之一。Zeng和Battle(2024)系统性回顾了过去的25年的可视化推荐研究,并对目标和策略进行了分类。为了融合领域知识,Coelho等人(2024)利用词向量、依赖解析器和专家定义的关键字列表,将领域文章中的专家知识编码为可计算形式,用于推荐可视化方案。随着LLM的出现,Podo等人(2024)提出了首个基于LLM的可视化推荐系统,增强了解释、字幕和建议功能;Ceballos等人(2025)则通过LLM根据数据类型自动推荐最适合的可视化图形。在引导用户入门和理解数据方面,Jianu等人(2025)提出“分析图”方法,将分析流程可视化为图结构以捕获关键语义单元,辅助用户形成数据洞察。此外,Alexander等人(2024)利用GPT-4检测可视化中的潜在误导内容,而Stokes等人(2026)构建框架理解文本元素在图表中的功能,为设计优化提供参考。

在人与Agent协同 workflow 层面,研究核心问题是如何分配控制权以及组织协同 workflow,通常以协同—自适应框架(Sperrle等,2021)和以人为中心的人工智能理念(Elmqvist和Klokmoose,2025)为指导原则,实践中,协同的形式从完全自动化的端到端系统(如Dibia,2023)到混合主动模式不等。业界常见的做法是将LLM嵌入编程环境,支持用户对部分流程进行修改,而无需重新分析全部内容(Wolter等,2025)。Agent主动发起交互是提升协同效率的关键,例如Choi等人(2024)的系统在对话中监控可视

化生成以检测问题,Takahira和Ueno(2025)则在实时数据演示的团队环境中分析数据集和受众特征,为用户提供优先推荐的可视化方案。对于可视化产物的精细化调整,Yan等人(2025)将图表图像分解为分层树,用户可通过自然语言或直接交互修改特定组件。未来趋势是个性化和自适应,系统根据用户专业水平和需求定制交互和可视化策略(Schnizer和Mayer,2025)。

在Agent-Agent协同范式中,国际研究聚焦多智能体系统在复杂流程自动化中的应用。LIDA(Dibia,2023)将图表生成流程拆分为4个模块(摘要器、目标探索器、可视化生成器和信息图表器),通过Agent间的反馈和迭代机制提升生成质量。SEED系统(Chen等,2024c)设计了优化器Agent,可从多个辅助LLM模块中选择和组合最佳数据管理管道。Goswami等人(2025)提出多模态协调框架,包括规划Agent与3个反馈Agent(数字、词汇、视觉),通过自我反思不断优化图表生成。Berger和Liu(2024)、Pan等人(2025)进一步将LLM/VLM用于生成流程中的“批评家”角色,提供设计改进建议。然而,目前LLM Agent的可视化素养仍有限,其生成图表在精度和设计偏好上尚不能完全与人类匹配(Ford等,2024;Wang等,2025c)。研究者尝试通过“教师—学生”模式微调Agent能力(Podo等,2024)。在洞察生成和叙事阶段,多Agent架构覆盖了从数据探索到视觉叙事完整流程(Wolter等,2025)。

Agent-External协同主要研究关注外部知识和工具的集成,以克服LLM在常识性偏差和领域知识缺乏上的局限。Shin等人(2025)展示了主动知识获取方法,通过与专家交互形成专门知识库,辅助后续可视化任务。Biswas等人(2025)的VizGenie系统在复杂任务情境下,可动态生成新的可视化脚本(如VTK Python代码)并无缝集成回系统,提高适应性。Borisov等人(2023)通过自回归生成LLM采样高度真实的表格数据,为数据探索提供支持。

总体来看,国际研究显示,大模型在可视化协同分析中正呈现3个趋势:1)人与Agent协同正在向多模态输入、迭代交互和自适应指导演进,实现更高水平的用户意图理解和洞察生成;2)Agent-Agent协同通过多智能体分工、反馈与自我评估机制优化生成流程,提高图表生成质量和可靠性;3)Agent-External协同通过知识库集成、工具拓展及真实数据生成增

强系统适应性,为复杂场景分析提供支撑。未来的发展方向可能包括个性化协同、自适应交互、端到端可解释生成以及多模态知识融合,这将进一步推动数据可视化在科学研究、商业分析和教育场景中的应用深化和智能化发展。

### 1.3 大模型赋能可视化叙事

在学术界,大模型已广泛用于可视化叙事,正成为自动化可视化叙事生成的关键引擎,展现出在构建叙事结构与生成叙事文本上的双重能力:首先,在叙事结构构建方面,大模型能够基于对数据事实与语义关系的深度理解,自动识别并聚合关键证据,从而优化故事的整体框架与叙述顺序。它能够将原本分散的数据洞见组织成因果清晰、层次递进的故事主线(Aodeng等,2025)。例如,香港科技大学与微软研究院将GPT-4作为“数据分析师”,通过链式思维提示技术,引导模型将多条数据事实串联成逻辑严谨、可读性强的叙事内容,而非简单罗列结论。其次,在叙事文本生成方面,模型能够在叙事骨架确立后,高质量地生成各类文本要素,包括标题、图注、批注与脚本等,以增强信息传达的完整性与理解度。大模型擅长从可视化中提炼关键信息,并生成简洁有力的标题与说明文字(Srinivasan等,2025;Islam等,2024)。例如,Tableau及MIT研究团队的Pluto(Srinivasan等,2025)采用混合主动式设计,利用大语言模型为常见图表生成语义丰富的标题、注释与解释性文本。在生成图表总结性文字的基础上,大模型进一步展现出构建完整数据叙事的能力。这些模型不仅能提炼关键信息,更能将分散的数据点有机串联,形成逻辑连贯、语义一致的叙事流。新加坡南洋理工大学研究团队提出了一种多智能体框架(Islam等,2024),通过两个智能体分别模拟人类叙事过程中的“理解与描述”和“验证”环节,前者负责生成大纲与叙述内容,后者在每一步进行质量校验,从而协同完成叙事生成。总体而言,大模型在显著降低叙事设计门槛的同时,既保障叙事逻辑的连贯一致,又显著提升表达效率与创作质量。

在产业界,科技巨头们正将大模型驱动的可视化叙事技术迅速推向实用。例如,OpenAI的GPT-4模型已成为许多数据分析与可视化工具的核心引擎。借助GPT-4的代码生成与分析能力,用户可以用自然语言自动生成数据可视化,并得到相应的分析解释。这一理念也被各大BI(business intelligence)

产品所采纳:微软Power BI的Copilot支持在报表内侧栏直接提问、生成解释性摘要与图形,并面向制作者提供“报告草案”等自动化能力;2024年,Tableau推出Tableau Puls,强调“个性化指标+自然语言洞察”的日常化触达;Qlik将生成式问答与其关联引擎结合,主打“结构化+非结构化”的统一回答与可追溯。这些产品的共同特点是从做图表升级为围绕业务问题自动产出可视化叙事。

### 1.4 可视化评估

可视化评估作为可视化领域的重要组成部分,传统上主要关注图表生成与交互系统的有效性,包括视觉编码质量、推荐效果、任务完成效率以及用户交互体验等指标(Savva等,2011;Lee等,2021)。进入大模型时代后,评估体系呈现多维拓展趋势:自动反馈机制(Maddigan和Susnjak,2023;Koh等,2025)支持生成图表的自我校正,多模态数据对齐(Han等,2024;Zadeh等,2024)确保语言、数据与图表渲染之间的一致性,而多智能体协作框架(Li等,2025a)通过任务分解提升复杂可视化任务的可扩展性。同时,研究者不断构建更系统的评估数据集与基准,如虚拟图表生成(Coelho等,2024)与综合测试框架(Zadeh等,2024),促进了从功能准确性向综合性能与认知契合度的转变。针对视觉分析中的智能引导生成,Ceneda等人(2024)提出了一种结合双重专家与最终用户的评估方法,旨在解决自动生成过程中人机协作的有效性问题。DracoGPT(Wang等,2025c)利用LLM从自然语言中提取可视化设计偏好,将其转化为形式化的软硬约束,实现了对生成结果合规性的自动化检查。

除了客观的评价指标,目前的图表评估开始关注生成内容对用户认知的深层影响。Wall等人(2024)提出的信任垃圾理论揭示了生成界面中的视觉装饰可能导致用户对AI(artificial intelligence)产生盲目信任或过度怀疑,强调了生成系统需校准用户的信任度。在认知偏差层面,Stokes等人(2026)的研究指出,生成图表中伴随的文本注释会显著重塑用户对数据的感知偏差,甚至改变预测判断。为了验证AI生成逻辑是否符合人类直觉,Wang等人(2025d)探讨了人机对齐问题,通过对比人类从图表中获取的结论与LLM的预测结果,评估生成模型是否真正理解数据的视觉表达,而非仅仅是统计概率的拟合。

此外,图表理解逐渐成为可视化评估的核心方向,其目标从传统的图表信息提取扩展到结构建模、跨模态对齐和复杂语义推理。多模态大模型时代的研究重点包括图表—JSON(Yan等,2024)、图表—Python字典(Chen等,2024a)以及融合表格、代码与JSON的联合对齐(Xu等,2026),为提升模型的泛化能力奠定了数据结构基础。同时,研究者提出无需数值标注的轻量级评估框架(Xu等,2024)、扩展图表类型覆盖(Xia等,2025)以及面向真实复杂场景的推理任务(Wang等,2024b),推动图表理解从实验室验证迈向真实应用。

总体来看,国际研究趋势表明,可视化评估正从单一准确性评估向多维度、任务驱动的综合评价体系演进,涵盖交互质量、结构一致性、多模态语义对齐和智能协作能力,标志着可视化系统正向智能化、语义驱动的下一代范式迈进。

## 2 国内研究进展

### 2.1 可视化理论基础

国内研究团队在隐式神经表示领域取得了显著进展,为数据压缩与表示提供了重要技术支撑。研究人员针对时变数据特征,采用知识蒸馏技术将不同时间步的神经表示统一融合到单一网络,利用时变数据中的重复模式提升压缩效率(Han等,2024);提出分治策略,通过对均匀划分的数据块分别进行压缩处理,有效提升了压缩性能(Han和Yang,2025);基于自监督学习机制解耦可视化图像中的颜色与数据信息,以实现可视化图像的重映射(Liu等,2026);设计编码器—多解码器架构的混合专家模型,实现了数据的非均匀划分,在保证质量的同时进一步优化了数据表示效果(Han等,2026)。

在可视编码方面,国内研究多集中于将大语言模型引入语义驱动的颜色与图表设计。在视觉编码设计方面,研究者提出利用语言模型理解用户的抽象设计意图,并自动转化为颜色语言或设计指令。例如,Hou等人(2024)构建了基于大语言模型的色彩设计系统,能够根据用户描述生成符合语义与风格的配色方案,并进一步将颜色映射到具体的空间或图表元素中,从而实现从语义到视觉编码的多层次映射;该团队进一步提出基于生成模型的语义—颜色关联学习框架,能够自动分析不同语义概念对

应的颜色分布(Hou等,2025);此外,Shi等人(2024)提出基于自然语言的图表配色优化方法,使非专业用户能够通过描述性语句对图表进行配色调整,降低了设计门槛。

浙江大学陈为教授团队提出的基于可微渲染的传输函数设计方法(Pan等,2024),通过探索传输函数的潜在设计空间并支持交互式图像编辑,使用户能够更直观地完成体渲染传输函数的定制化设计。尽管目前大模型在可视编码方面具有一定应用,但在视觉感知建模、感知一致性与人类感知对齐等方面仍缺乏系统性研究,尚未形成可支撑图形感知的理论体系。

国内学者在图表生成领域取得了系统性进展,其中既包括声明式语法层面的表达能力提升,也包括自然语言驱动的大模型生成方法的突破。在可视化规范语言方面,GoTree(Li等,2020b)通过解构树图可视化布局过程为可组合的几何关系有效封装复杂算法细节;PiCCL(pictorial chart composition language)(Kuang等,2025)通过引入4类组合操作符,推动了图形图表设计的结构化与模块化发展。在此基础上,随着大模型技术的兴起,自然语言驱动的图表生成能力迅速发展。Tian等人(2025a)提出基于任务分解的生成策略,将图表生成拆解为数据筛选、图表选择与视觉编码等子任务,并基于FLAN-T5实现了高可控性的自然语言到图表规范转换;Han等人(2023b)开发的ChartLlama融入可视化领域知识,在多项图表理解与生成任务中表现优异;Li等人(2025d)进一步针对自然语言查询难以精准映射到可视化规范的瓶颈,提出基于上下文学习的Prompt4Vis框架,在NVBench数据集上达到领先效果。此外,国内学者在财务可视化叠加(Hao等,2024)、图表实例驱动创作(Xie等,2025)、数据叙事可视化(Shao等,2025b)等垂直领域取得了突破,同时也在积极推进大规模基准(Hu等,2024;Niu等,2025)的构建工作。总体而言,从声明式语法到大模型驱动生成的多技术融合,使图表生成正从以规则为中心的“规范构建”向以语义为核心的“智能生成”加速演进,为自然语言到可视化的全流程自动化奠定了重要基础。

在可视化交互方面,国内研究同样取得了重要进展。中国科学院单桂华研究员团队开发的科学数据编辑方法(Li等,2025b)实现了通过拖拽操作探索

特征未知区域,并借助代理模型生成相应的模拟结果与参数配置,大幅提升了科学数据探索的交互体验。中山大学陶钧副教授团队创新性地将大语言模型技术应用于流场可视化领域,提出的基于自然语言查询的流场可视化方法(Li等,2025f),通过将用户指令转化为声明式语法,实现了可视化流程的智能化定制;该团队还进一步利用大语言模型促进科学数据的语义表征,在无需人工标注的情况下,实现了基于语义的流场特征匹配(Zhang和Tao,2026)和视角选择(Zhao和Tao,2026)。在可视化交互语法层面,Nebula(Chen等,2022)为跨视图协调交互提出独立的声明式语法,实现了交互逻辑与视觉规范的彻底解耦,为复杂交互的系统化表达奠定了理论基础。Libra(Zhao等,2025a)从数据转换的角度出发,提出能够同时满足多重交互需求的模型框架,并以组件化架构支持复杂交互过程的分解、组合与可重用化。然而,现有交互语法仍主要依赖人工显式定义,在表达多轮且人机协同的交互意图方面存在局限。

总体而言,国内可视化研究从隐式神经表示、自然语言驱动可视化以及声明式规范语言等多个层面都形成了独具特色的可视化理论,为构建自主可控的可视化技术体系奠定了坚实基础。

## 2.2 大模型赋能可视分析

国内在大模型赋能可视化协同分析领域的研究呈现出蓬勃发展的态势,在构建端到端智能分析框架、探索新型多模态交互以及深化特定领域应用方面展现出强劲的创新活力。国内最早开始探索的大模型赋能可视分析成果包括浙江大学巫英才团队提出的ChartGPT(Tian等,2025a)与复旦大学陈思明团队提出的LEVA(Zhao等,2025c)。

在人与Agent协同研究方面,随着大语言模型的兴起,研究重点已转向融合多模态输入和通过迭代交互理解用户意图的新范式。针对高层次模糊分析任务,复旦大学陈思明团队采用了交互式任务分解策略(Qiu等,2025),而浙江大学巫英才团队和香港科技大学(广州)杨维铠团队则允许用户检查并修改工作流中每个步骤的中间输出,以识别错误并提供针对性调整(Tian等,2025b;Shuai等,2026)。Li等人(2025e)特别关注分析任务的“冷启动”阶段,系统通过主动发起交互帮助用户迭代阐明目标,并支持回滚到历史记录点以探索新的分析路径。为突破

纯文本交互的局限,研究者引入了更丰富的输入模态。Xie等人(2026)支持用户提供SVG(scalable vector graphics)可视化作为“示例”进行调整和修改,Wen等人(2025)和Lin等人(2024)引入了包括示例图、草图、修改操作在内的多模态提示作为文本补充。Wang等人(2025a)将文本提示与精确的GUI(graphical user interface)交互相结合,Zhao等人(2026)通过主动监控用户操作上下文推断真实意图。Shen等人(2026)进一步将这种组合形式化为“交互增强指令”模型,提炼了可组合的原子交互范式。

在Agent向人传达方面,为实现更好的可解释性和确定性,Wang等人(2024a)提供了基于经验驱动和洞察驱动的可解释规则系统。Li等人(2023)使用LSTM(long short-term memory)模型捕捉用户交互和视觉状态,为新手用户提供多样化、实时的交互建议。Xie等人(2024b)采用强化学习技术,通过蒙特卡罗图搜索算法主动获取用户反馈,迭代推荐可视化。在推荐时机和介入方式研究上,复旦大学陈思明团队提出ProactiveVA,通过形成性研究深入探讨了用户需求时机和帮助方式,提出了“感知—推理—行动”三阶段自主代理框架(Zhao等,2026)。Cai等人(2025)在数据治理阶段采用原位可视化帮助用户理解数据转换脚本,直观呈现转换语义及数据沿袭关系。浙江大学陈为团队通过可视分析系统ConceptViz将稀疏自编码器的特征空间与人类概念建立可视关联,使用户能够交互式发现、解释并验证大模型中的概念表示,系统如图3所示(Li等,2026)。

在Agent与人协同工作流程方面,国内研究者致力于构建以人为中心的协同机制。香港科技大学屈

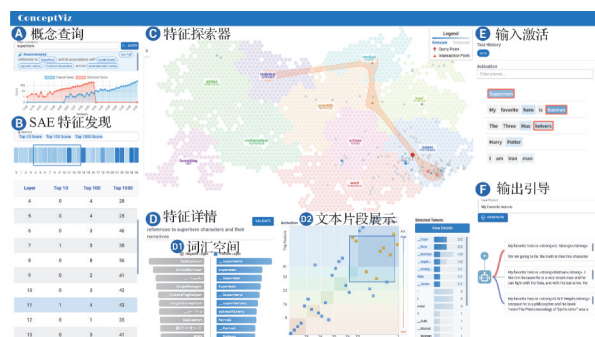


图3 ConceptViz系统界面(Li等,2026)

Fig. 3 The interface of the ConceptViz system  
(Li et al., 2026)

华民团队提出的PyGWalker作为广受欢迎的实时探索性分析库,有效解决了程序分析脱节问题(Yu等,2024)。随着协同深入,Agent分析过程的透明化和可控化成为研究重点。同济大学曹楠团队通过动画方式呈现数据分析管道、操作及数据变化(Guo等,2026),Xie等人(2024a)将LLM生成代码转换为交互式、分步可视化表示,支持用户理解、验证和调整数据操作。在分析流程的最终阶段,洞察组织与报告生成受到广泛关注。Chen等人(2020)指出复杂分析结果需要通过故事合成联系数据分析与结果呈现,Chen等人(2024b)系统总结了6种叙事可视化类型和4种自动化级别的创作工具。Li等人(2025c)观察到了从传统“人类创造者+AI助手”向“AI创造者+人类评论者”等新兴模式的转变,Zhao等人(2025c)展示了高效的协同报告策略,支持用户通过流可视化追溯分析历史并生成多格式报告。

在Agent与Agent协同研究领域,可视化生成阶段的核心挑战被分解为一系列子任务,由不同专业Agent协作完成,这种分解策略显著提升了系统鲁棒性与性能(Tian等,2025a)。针对关系型数据库场景,Yang等人(2024a)利用思维链技术将任务分解为生成SQL(structured query language)、确定图表类型和数据映射3个步骤。Qiu等人(2025)采用类似方法,在与人类协作完成初步任务分解后,每个子任务Agent首先生成详尽的数据处理计划,再根据计划生成代码与图表,最终将多个静态图表链接为统一联动的整体。

为提升生成质量,Agent间的反馈与迭代机制成为研究重点。复旦大学陈思明团队提出的LightVA可视分析系统,结合少样本/思维链提示、自反思调试、规则化捕错与反馈机制,显著降低生成失败风险,同时提出多视图联动的一致性策略(Zhao等,2025b)。Yang等人(2024b)融合代码LLM和多模态LLM,利用后者提供的可视化反馈帮助前者迭代调试代码。Shi等人(2025)利用微调视觉语言模型Agent评估并细化Text2Vis输出,形成“生成—评估—改进”的闭环。Shao等人(2025a)指出Agent模拟只能作为用户研究的补充,而Shuai等人(2026)通过构建专门思维链数据集提升Agent能力。

在洞察提取与组织阶段,Wang等人(2025b)分配多个Agent生成初始图表摘要并进行迭代协作,期间调用外部数据分析模块提取深度见解。Zhao

等人(2025c)设计了两阶段流程来筛选、计算和评估洞察,并最终在视图上进行标注。在叙事与报告生成阶段,Agent协同致力于将分析过程和洞见组装成结构化多模态报告。Zhao等人(2025c)提出选择性报告策略,通过交互/洞见矩阵维护分析历程,利用流可视化追溯历史。Yang等人(2025)针对可视化与文本报告集成难题,提出可视化形式化描述这一图表结构化文本表示方法,使LLM能学习并生成多样化图表。

在Agent与外部协同研究方面,国内学者在外部知识、工具和数据集成方面取得显著进展。为克服LLM图表推荐偏见,Qiu等人(2025)采用检索增强生成技术,向LLM注入来自权威论文和教科书的专业知识,使其能够为特定任务推荐更合适的多样化图表。Gao等人(2025)将领域问题列为LLM集成到可视化系统的三大核心挑战之一,并提出用于微调LLM以增强特定领域任务的指导框架。

在外部工具拓展方面,Wang等人(2025b)系统在执行图表洞察任务时,各Agent主动调用外部数据分析模块提取深层次见解,最终将零散见解编译成连贯完整的摘要。在Agent与数据协同层面,作为以代码为中心的Agent,能够执行数据检索和爬取,为后续分析准备数据基础(Zhang等,2025)。在可视化生成阶段,这种以代码为中心的协同模式尤为突出,Zhang等人(2025)接收自然语言请求后生成相应代码处理数据并生成可视化。学界也在探索更灵活的桥梁,Luo等人(2022)研究指出当时自然语言到可视化方法的局限性,提出基于Transformer的序列到序列模型,将自然语言翻译为简化的中间可视化语法,支持端到端生成。

总体而言,国内在LLM与可视分析交叉领域的研究已形成系统化、多层次的技术体系,在协同机制、交互范式和应用创新等方面均取得重要突破,为构建智能可视分析生态系统奠定了坚实基础。

### 2.3 大模型赋能可视化叙事

国内学术界在大模型赋能可视化叙事领域的研究呈现出系统化发展态势,从早期的自动图表生成逐步拓展至具备深度逻辑推理、多模态融合与动态交互能力的叙事系统构建。在叙事逻辑构建层面,同济大学研究团队提出Erato系统,该系统微调BERT(bidirectional encoder representations from Transformers)模型实现数据事实的逻辑嵌入,使向

量表征能够有效捕捉事实间的因果与约束关系(Sun等,2023)。浙江大学研究团队开发了基于大模型的报告分析系统,能够从现有报告中逆向工程分析逻辑并适配到新数据集,生成逻辑连贯的数据报告(Tian等,2025b);设计了面向体育领域的叙事可视化系统SNIL,将比赛关键事件自动转化为图文并茂的叙事序列(Cheng等,2025)。

在多模态融合方面,可视化叙事研究呈现出超越静态图文的发展趋势。复旦大学与香港科技大学联合团队提出的SceneLoom系统基于视觉语言模型,实现了数据可视化与真实世界图像的智能融合(Gao等,2026)。香港科技大学团队在动态数据视频领域研发了具有叙事—动画交互功能的动态数据视频生成技术(Shen等,2025b;Shao等,2025b;Shen等,2024),如图4所示。香港科技大学(广州)团队创新性地利用文本到图像生成模型,将语义上下文嵌入图表元素,增强了可视化的语义表达能力(Xiao等,2024)。

产业界的发展与学术界形成良好呼应。字节跳动的VisActor专注于叙事可视化与智能可视化的深度实现,支持场景化动画编排和丰富的交互能力。

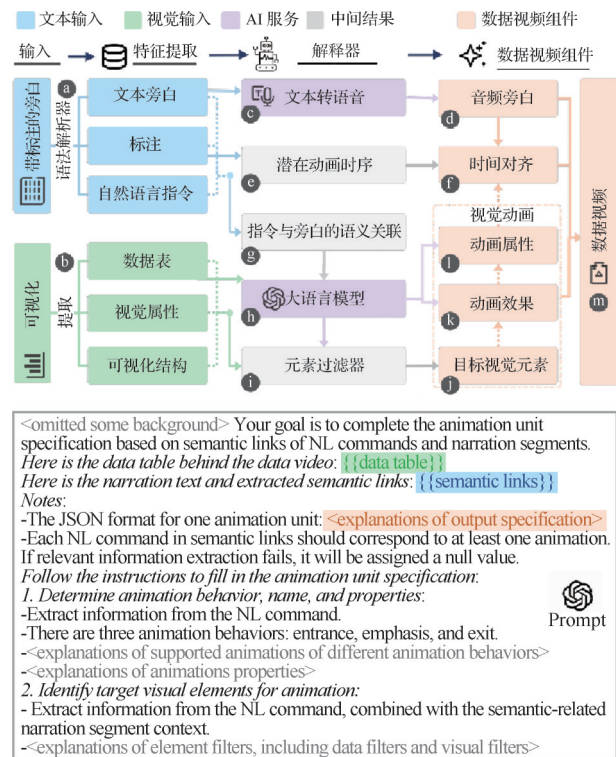


图4 从用户标注中自动合成数据视频(Shen等,2025b)  
Fig. 4 The automatic interpreter to synthesize data video components from users' annotations (Shen et al., 2025b)

AntV 提供系统化的动画与叙事解决方案,阿里云 QuickBI 深度融合 AI 能力,创新性地将自然语言交互、智能归因等功能融入分析流程。可视化叙事正从专业技术转变为普遍技能,形成了从理论创新到产业实践的完整技术生态。

总体而言,国内在该领域已建立起较为完善的技术体系,推动着可视化叙事向智能化、普及化方向发展,为数据驱动的决策支持提供了坚实的技术基础。学术研究与产业应用的深度融合,不仅拓展了可视化叙事的应用边界,也为其未来发展注入了持续动力。

### 2.4 可视化评估

我国在可视化评估领域的研究呈现出体系化、场景化与智能化协同推进的发展态势。在评测体系与基准构建方面,清华大学刘世霞教授团队发布了面向真实业务场景的专用基准数据集,在多轮问答、复杂语义解析等任务中提供了系统化的评估标准,为可视化系统性能测评奠定了基础(Xie等,2025)。Chen 等人(2025b)建立了一个包含多维度任务的评估数据集 VisEval,主要包含有效性检查器、合法性检查器、可读性评估器等3个主要模块,专门用于量化 LLM 在图表代码生成中的准确性与鲁棒性,如图5所示。

针对多模态可视化理解的评估需求,香港科技大学(广州)曾伟助理教授团队提出基于可视化参照的指令微调方法(Zeng等,2025),有效提升了模型在图表语义细粒度理解上的稳定性与可解释性,并进一步构建了图表基础读图能力的评估架构(Wang等,2025e),推动了可视化能力测评向结构化、可量化方向演进。从整体格局看,我国可视化评估研究正从传统的任务性能度量扩展到涵盖数据理解、语义推理、交互行为与任务完成效果的多维度综合评估,并逐步向真实场景驱动的评测框架转变。同时,借助前沿大模型与多模态技术,可视化评估方法正从静态测试迈向动态、人机协同与智能诊断导向的发展路径。

## 3 国内外研究进展比较

在大模型时代的推动下,数据可视化研究呈现出智能化、语义驱动与多模态协同的新趋势。国际与国内在该领域均取得显著进展,但在研究侧重点、

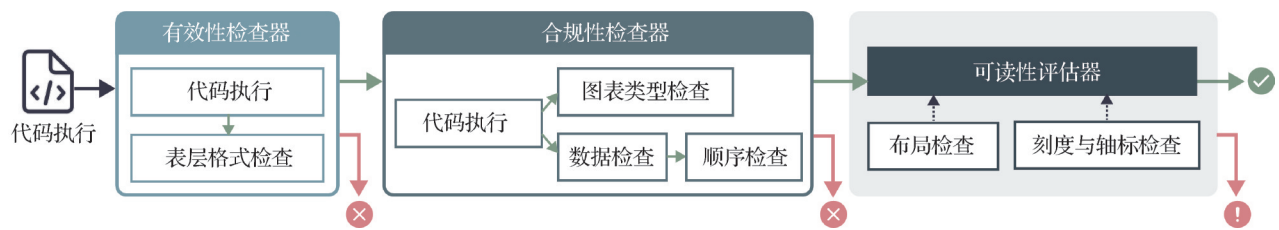


图5 VisEval评估数据集构建过程(Chen等,2025b)

Fig. 5 The pipeline of constructing the benchmark VisEval dataset (Chen et al., 2025b)

方法布局与应用落地方面存在差异。总体来看,国际研究在理论框架、方法通用性及跨领域适用性方面具有领先优势,国内研究在系统落地、工程实现和垂直场景应用上表现突出,两者互补,共同推动数据可视化向智能化、个性化与普及化发展。

在可视化基础理论与数据表示方面,国际研究注重方法全面布局和高通用性。研究人员广泛应用图神经网络、序列建模及大语言模型,对社交网络、时间序列和文本数据进行潜在特征提取,实现高效可视化呈现与智能推理。在科学可视化领域,隐式神经表示已成为主流方法,支持任意分辨率的连续场采样与渲染,并在标量场压缩、多任务统一架构及训练效率优化上取得突破。视觉编码与图表生成方面,国际团队利用多模态大模型实现语义—视觉映射、颜色和布局优化,以及自然语言驱动的图表生成和多智能体协作,使生成过程更自动化、可迭代优化。交互研究从传统声明式操作逐渐演进至语义感知、自然语言交互及智能协同模式,推动人机交互高效化。相比之下,国内研究更强调应用导向和工程落地。在数据表示与压缩方面,国内团队在时变数据压缩、多解码器架构设计和非均匀数据表示上取得创新,推动隐式神经表示在实际场景下的高效应用。大语言模型广泛引入可视编码和语义驱动图表生成,实现自然语言到规范化图表的高可控映射,并在财务可视化、叙事可视化等垂直场景落地中发挥作用。交互方面,国内团队提出独立声明式交互语法、组件化数据转换模型以及自然语言驱动的数据探索方法,使复杂交互能够系统化表达,并支持任务驱动的智能分析流程。

在可视化协同分析与智能化框架构建方面,国际研究注重方法通用性和多层协作体系。多智能体框架(human-agent、agent-agent及agent-external)通过任务分解与反馈优化图表生成与分析流程,同时结合外部知识库和工具,实现复杂场景下的高适应性。

例如,LIDA(Dibia,2023)、SEED(Chen等,2024c)等通过迭代交互和自适应指导,实现图表生成和分析智能化。然而,国际研究在个性化协同与具体领域应用落地方面仍有限。国内研究以应用落地为核心,通过端到端智能分析框架、多模态交互与外部工具整合,实现用户意图精细捕获和洞察生成。国内团队结合SVG示例、草图操作、GUI交互和外部分析模块,构建系统化、多层次的协同分析体系,并在可视化报告生成、叙事整合及垂直应用中取得突破。

在可视化叙事领域,国际研究突出方法深度与通用性,关注大模型在叙事结构构建与文本生成上的双重能力。通过链式思维、多智能体协作和主动式提示,国际团队生成逻辑连贯、语义一致的自动化叙事,并在产业界推动GPT-4、Power BI Copilot和Tableau Plus等工具,实现端到端自动化可视化叙事和业务问题导向分析。国内研究更强调系统化实现和多模态融合,学术界构建具备逻辑推理、动态交互和跨模态能力的叙事系统,如Erato(Sun等,2023)、SNIL(Cheng等,2025)、SceneLoom(Gao等,2025),能够将数据、图表、图像及视频进行智能融合;产业界以VisActor、AntV和QuickBI等产品推动可视化叙事向普及化和场景化应用发展,注重动画交互与智能分析结合,实现可视化叙事的落地与规模化。

在可视化评估方面,国际研究通用评估体系与方法创新,关注自动反馈、多模态对齐、多智能体协作及综合性能基准的建立,使图表理解从实验室验证向复杂场景应用延展。国内研究则侧重体系化落地与场景化评测,通过专用基准数据集、指令微调及图表基础能力评估,强化模型在真实任务中的可解释性与稳定性,并推动评估方法向动态、人机协同及智能诊断方向发展。总体而言,国际在方法广度和理论深度上领先,国内在应用落地、跨模态融合及垂直场景实现上具优势,两者互补,共同推动可视化评

估向智能化、任务驱动的新范式演进。

大模型时代的数据可视化研究呈现理论创新与应用落地协同发展的格局。国际研究在方法通用性、理论框架和多模态智能协作方面引领前沿,而国内研究在系统化实现、工程落地和特定场景应用方面展现优势。随着大模型在语义理解、知识整合和任务分解能力的提升,国内外可视化研究正从“数据到图形”的传统逻辑逐步扩展至“意图—生成—理解—交互—评估”的智能化全流程,推动数据可视化向更高效、智能和个性化方向发展,为科研、产业和决策支持提供坚实支撑。

## 4 发展趋势与展望

随着大数据和多模态大模型技术的发展,数据可视化正在迈向智能化、协同化和普适化的新阶段。大模型的语义理解、推理与表达能力,使用户能够以自然语言描述分析意图,推动交互方式从参数调控向语义驱动探索转变。同时,跨任务、跨场景的多智能体协作、自然语言交互、可解释的视觉生成及沉浸式数据叙事成为研究重点。表1—表4从大模型赋能可视化基础、大模型赋能可视

表1 大模型赋能可视化基础涉及的主要任务以及主流技术

Table 1 Key tasks and technologies for LLM-driven data visualization fundamentals

子类	主要文献	可视化任务	主要技术
数据表示	[CGF 范德比尔特大学](Lu等,2021);[CGF 慕尼黑工业大学](Weiss等,2022); [IEEE TVCG 香港科技大学](Han和Wang,2023a;Han等,2024;Han和Yang,2025);[IEEE VIS 香港科技大学](Han和Yang,2025)	体数据压缩、时变体数据生成、大规模时变体数据压缩	隐式神经表示
可视编码	[IEEE TVCG 浙江大学](Pan等,2024);[ECCV 巴塞罗那自治大学](Butt等,2025);[ACM CHI 香港科技大学(广州)](Hou等,2024;Hou等,2025); [IEEE VIS 山东大学](Liu等,2026); [IEEE VIS 纽约州立大学石溪分校](Salim等,2026); [CVPR 加州大学伯克利分校](Mao等,2025); [BELIV 范德比尔特大学](Berger和Liu,2024)	体绘制传递函数设计、颜色生成、颜色与语义关联、颜色优化、颜色迁移、颜色感知评估	可微渲染、隐式神经表达、扩散模型、多模态大模型、大语言模型
语义驱动的图表生成	[IEEE TVCG 加州大学圣巴巴拉分校](Ren等,2019); [VLDB 加州大学伯克利分校](Lee等,2021); [IEEE TVCG 佐治亚理工学院](Narechania等,2021); [ESEC/FSE 伊利诺伊大学厄巴纳—香槟分校](Wu等,2022); [ACL 微软研究院](Dibia,2023); [ACM CHI 滑铁卢大学](Masson等,2023;Masson等,2024); [arXiv Preprint 南洋理工大学](Han等,2023b); [ACM CHI 哈佛大学](Vaithilingam等,2024); [EMNLP 首尔国立大学](Zadeh等,2024); [VLDB 香港理工大学](Li等,2025d);[NAACL 延世大学](Koh等,2025); [ACL 加州大学洛杉矶分校](Li等,2025a);[IEEE TVCG 浙江大学](Tian等,2025a)	自然语言到可视化生成、语法无关可视化与信息图自动生成、自然语言到数据可视化代码生成、多模态内容编辑(文本、代码、矢量图)、动态UI组件合成、参考引导的图表生成、图表问答、图表数据提取、文本引导的可视化查询生成	大语言模型、视觉语言模型、提示工程、多智能体框架、基准数据集、对比学习、上下文学习
可视化交互	[IEEE TVCG 斯坦福大学](Satyanarayan等,2016;Satyanarayan等,2017); [ACM CHI 北京大学](Li等,2020a); [Visual Informatics 俄亥俄州立大学](He等,2020); [IEEE TVCG 圣母大学](Han等,2021;Tang等,2026); [IEEE TVCG 浙江大学](Chen等,2022); [IEEE TVCG 俄亥俄州立大学](Shi等,2022;Shi等,2023;Chen等,2025c;Shen等,2025a);[IEEE VIS 范德堡大学](Jeong等,2024); [IEEE TVCG 阿卜杜拉国王科技大学](Jia等,2025); [IEEE VIS 圣母大学](Ai等,2026);[IEEE TVCG 中国科学院](Li等,2025b); [ACM CHI 山东大学](Zhao等,2025a);[Visual Informatics 中山大学](Li等,2025f); [PLoS ONE 中国人民解放军总医院](Kuang等,2025); [IEEE TVCG 中山大学](Zhao和Tao,2026;Zhang和Tao,2026)	声明式交互语法、多变量时变数据选择与转换、多轮对话式可视交互、多视角可视化、多视图联动交互、树可视化交互、多集合比较与差异定位、集合空间查询、参数空间探索、语义引导的传输函数交互、体数据视角选择、流场模式交互	表示学习、去噪自编码器、对比学习、生成对抗式网络、响应式数据流架构、归一化流基代理模型、大语言模型、提示工程、多智能体、高斯泼溅

表2 大模型赋能可视分析涉及的主要任务以及主流技术

Table 2 Key tasks and technologies for LLM-driven visual analytics

子类	主要文献	可视化任务	主要技术
human-to-agent	[Visual Informatics 俄亥俄州立大学](He等,2020); [C&G 康斯坦茨大学](Sperrle等,2021); [IEEE Access 梅西大学](Maddigan和Susnjak,2023); [IEEE VIS 马里兰大学](Bako等,2024;Gathani等,2024); [IEEE VIS 成均馆大学](Choi等,2024); [IEEE CGA 奥胡斯大学](Elmqvist和Klokose,2025); [IEEE TVCG 微软研究院](Chen等,2025b); [ACM CHI Workshop on AI Tools for Thought 慕尼黑大学](Schnizer和Mayer,2025);[IEEE TVCG 浙江大学](Tian等,2025a); [NeurIPS 香港科技大学(广州)](Luo等,2026); [arXivEnhans](Seo等,2026);[IEEE VIS 香港科技大学(广州)](Shuai等,2026);[FMS 中国石油大学(华东)](Li等,2025e); [IEEE TVCG 香港科技大学](Xie等,2026);[arXiv 浙江大学](Wen等,2025); [ACM CHI 微软研究院](Wang等,2025a); [ACM CHI 香港科技大学](Shen等,2026)	可视化分析指导、数据准备与探索性可视化、自然语言到可视化转换、模糊与歧义查询处理、可视化任务与意图识别、可视化生成与推理路径构建、可视化设计质量评估与优化、多模态可视化意图表达、用户自适应可视化与交互、迭代式数据可视化创作	机器学习、自然语言处理、大语言模型与多智能体架构、提示工程与少样本提示、分步推理与思维链推理、可视化生成与渲染、可视化设计优化与启发式规则、用户行为建模与自适应交互、多模态提示与交互增强、代码生成与自动化调试
agent-to-human	[IEEE TVCG 维也纳工业大学](Ceneda等,2017); [arXiv 罗马萨皮恩扎大学](Podo等,2024); [IEEE TVCG 复旦大学](Li等,2023);[FTD 马里兰大学](Zeng和Battle,2024); [Analytics 纽约州立大学石溪分校](Coelho等,2024); [IEEE VIS 马萨诸塞大学阿默斯特分校](Alexander等,2024); [Visual Informatics 浙江大学](Wang等,2024a); [PVLDB 香港科技大学(广州)](Xie等,2024b);[IEEE VIS 浙江大学](Li等,2026);[FFF 阿根廷泛美开放大学](Ceballos等,2025); [CGVC 伦敦城市大学](Jianu等,2025); [IEEE TVCG 加州大学伯克利分校](Stokes等,2026); [IEEE VIS 复旦大学](Zhao等,2025c);[Visual Informatics 浙江大学](Cai等,2025)	可视化推荐系统与多源信息融合、语义驱动可视化任务推荐、LLM驱动可视化生成与优化、多轮/主动式可视化交互、人机协同可视化与强化学习、可视化质量评估与误导检测、数据分析流程可视化与思维导图生成、文本与视觉元素协同分析、数据变换脚本理解与可视化追踪	提示词工程、语义网络构建、大语言模型、用户行为分析、强化学习、原位程序/数据谱系可视化、多模态视觉语言模型
人与Agent协同	[IEEE TVCG 弗劳恩霍夫 IAIS 研究所](Chen等,2020); [C&G 康斯坦茨大学](Sperrle等,2021); [arXiv 香港科技大学(广州)](Yan等,2025); [ACL 微软研究院](Dibia,2023);[IEEE VIS 马里兰大学](Gathani等,2024); [IEEE VIS 香港科技大学 & Kanaries Data](Yu等,2024); [UIST 香港科技大学](Xie等,2024a);[IEEE TVCG 同济大学](Chen等,2024b);[IEEE TVCG 复旦大学](Zhao等,2026); [IEEE CGA 奥胡斯大学](Elmqvist和Klokose,2025); [ACM CHI 慕尼黑大学](Schnizer和Mayer,2025); [ACM THS 同济大学](Guo等,2026); [IEEE VIS 微软亚洲研究院](Li等,2025c); [VIS x GenAI 奥胡斯大学](Wolter等,2025)	可视化分析指导与评估、探索性数据分析与洞察生成、自然语言驱动的可视化生成、可视化叙事与报告自动化、多智能体与生成式可视化交互、用户自适应与个性化可视化、图表组件与样式精细化编辑、数据分析过程可视化与监控、LLM增强的可视化工作流与智能推荐、生成式 AI 时代的人机协作与多阶段分析优化	机器学习、自然语言处理、提示工程、图像与可视化生成模型、可视化代码生成与执行、LLM及智能体、多智能体协作与人机协作框架、可视化分析与叙事生成、用户行为建模与自适应交互、自动化评估与可视化质量优化

续表2 大模型赋能可视分析涉及的主要任务以及主流技术

Continue Table 2 Key tasks and technologies for LLM-driven visual analytics

子类	主要文献	可视化任务	主要技术
Agent 与 Agent 协同	[arXiv 罗马萨皮恩扎大学](Podo等,2024); [Annual Meeting ACM 微软研究院](Dibia,2023); [IEEE TVCG 华盛顿大学](Wang等,2025c); [IEEE TVCG 复旦大学](Zhao等,2025b;Zhao等,2025c;Wang等,2025b); [arXiv 麻省理工学院(MIT)](Chen等,2024c); [BELIV 范德堡大学](Berger和Liu,2024); [IALP 北京语言大学](Yang等,2024a); [ACL 清华大学](Yang等,2024b);[arXiv 浙江大学](Pan等,2025); [VIS x GenAI 奥胡斯大学](Wolter等,2025); [IEEE VIS 香港科技大学(广州)](Shuai等,2026); [ACM 英迪拉·甘地女子技术大学](Goswami等,2025); [IEEE TVCG 浙江大学](Tian等,2025a); [ACM MM 中国科学院大学](Shi等,2025); [IEEE CG&A 复旦大学](Shao等,2025a); [COLING 德克萨斯大学圣安东尼奥分校](Ford等,2025)	自然语言到可视化转换、多模态可视化生成与交互、可视化代码自动生成、可视化设计与优化(参数、偏好、美学)、LLM驱动的可视化评估与自评、可视化叙事与报告生成、多智能体协作可视化、科学数据可视化与数据驱动分析、可视化迭代与错误修正、智能洞察推荐与分析历史追溯	图像生成模型、多模态基础模型、表示学习与微调技术、提示工程、思维链、LLM、智能体、多智能体协作架构、可微渲染
Agent 与外部 协同	[IEEE TVCG 清华大学](Luo等,2022); [ICLR 蒂宾根大学](Borisov等,2023); [IEEE TVCG 洛斯阿拉莫斯国家实验室](Biswas等,2026); [arXiv 浙江大学](Zhang等,2025); [arXiv 奥胡斯大学](Shin等,2025)	领域知识抽取与注释、LLM驱动知识采集与问答、可视化辅助知识发现、科学可视化自动化与特征导向渲染、视觉问答与领域自适应可视化、表格与异质数据生成与建模、合成数据真实性优化与多模态转换、数据协作与自动化分析 workflow、自然语言到可视化转换与序列生成优化	混合主动交互、提示工程、结构化知识库构建、检索增强生成、LLM微调与低秩适配、VTK/ParaView可视化代码生成、Transformer序列模型、智能体 workflow、多模态可视化界面与特征数据库

表3 大模型赋能可视化叙事涉及的主要任务以及主流技术

Table 3 Key tasks and technologies for LLM-driven visual storytelling

主要文献	可视化任务	主要技术
[IEEE TVCG 同济大学](Sun等,2023); [ACL 约克大学(加拿大)](Islam等,2024); [IEEE TVCG 香港科技大学](Shen等,2024;Xiao等,2024;Shen等,2025b); [UIST 北京理工大学](Aodeng等,2025); [IUI Tableau 研究院](Srinivasan等,2025); [UIST 浙江大学](Tian等,2025b); [IEEE TVCG 浙江大学](Cheng等,2025); [IEEE VIS 复旦大学](Gao等,2026); [IEEE TVCG 复旦大学](Shao等,2025b)	表格与图表驱动的数据故事构建、交互式叙事可视化创作、文本与图表联合写作、自动化数据报告与叙事生成、多模态叙事(图表+文本+视频)、人机协作式数据故事编辑、历史报告驱动的协作分析、体育及事件数据新闻生成、可视化动画与视频创作、图表语义增强与美化、旁白标注	大语言模型、多智能体 LLM 框架、检索增强生成与语义过滤、Acting/Reasoning 双模块交互 workflow、VLM、文本到图像生成模型、语义对齐与双向建议引擎

分析、大模型赋能可视化叙事和大模型赋能可视化评估4个方面分别总结了大模型时代可视化及可视分析应用所涉及的主要任务以及主流技术。可以预见,大模型技术与数据可视化的深度融合将成为未来的发展趋势,在这一新兴工具的不断

发展与应用背景下,数据可视化还有很多值得研究的问题。

1)面向数据可视化的跨任务、跨场景的通用可视化基础架构。未来研究需构建可支持多类型数据、任务与用户场景的可视化框架,实现从数据解

表4 大模型赋能可视化评估涉及的主要任务以及主流技术

Table 4 Key tasks and technologies for LLM-driven visualization evaluation

主要文献	可视化任务	主要技术
[IEEE TVCG 维也纳工业大学](Ceneda 等, 2024); [VLDB 加州大学伯克利分校](Lee 等, 2021); [IEEE Access 梅西大学](Maddigan 和 Susnjak, 2023); [arXiv 南洋理工大学](Han 等, 2023b); [arXiv 清华大学](Xu 等, 2024); [Analytics 纽约州立大学石溪分校](Coelho 等, 2024); [EMNLP 首尔大学](Zadeh 等, 2024); [PacificVis 埃默里大学](Wall 等, 2024); [ICDAR 纽约州立大学布法罗分校](Yan 等, 2024); [ACM MM 中国科学院大学](Chen 等, 2024a); [NeurIPS 普林斯顿大学](Wang 等, 2024b); [NAACL 延世大学](Koh 等, 2025); [ACL 加州大学洛杉矶分校](Li 等, 2025a); [IEEE TVCG 华盛顿大学](Wang 等, 2025c; Wang 等, 2025d); [IEEE TVCG 加州大学伯克利分校](Stokes 等, 2026); [TIP 上海人工智能实验室 / 上海交通大学](Xia 等, 2025); [IEEE TVCG 微软研究院](Chen 等, 2025b); [NeurIPS 清华大学](Xie 等, 2025); [IEEE TVCG 香港科技大学(广州)](Zeng 等, 2025); [CGF 香港科技大学(广州)](Wang 等, 2025e); [ICLR 清华大学](Xu 等, 2026)	图表数据抽取与结构 解析、图表生成质量 评估与反馈、可视化任 务推荐与偏好建模、交 互式分析过程引导与 导航、文本—视觉联合 数据传播、可信与负责 任的 AI 可视化设计	大语言模型、多模态视 觉语言模型、提示工程 与微调、多智能体、自 动反馈驱动强化学习微 调、可解释性与可置信 度评估、大规模可视化问 答与自然语言—可视化 数据集构建

析、特征提取到图形生成、交互呈现的端到端能力。兼顾可扩展性与模块化设计,使不同可视化任务能够共享底层知识表示、图形生成策略及交互接口,提升跨场景复用和多任务协作效率,为智能可视化系统提供统一基础支撑。

2)完善多智能体协同与人机共创可视分析及叙事 workflow。在复杂数据分析与可视化叙事中,多智能体协作能够实现任务分解、迭代优化和逻辑校验,而人机共创可确保分析意图与用户需求紧密匹配。未来研究应探索智能体之间及人机之间的高效交互机制、知识共享与反馈策略,结合动态任务调度与可解释提示,实现从数据理解到图表生成再到叙事构建的全流程协同优化。

3)建立多模态、个性化的智能叙事方法。大模型可将文本、表格、图像和视频等多模态信息融合,生成逻辑连贯且符合用户偏好的可视化叙事内容。未来研究需关注数据可视化叙事中的个性化偏好建模、情境感知与多模态数据对齐,通过智能摘要、标题生成和情节结构优化,实现叙事内容的可控性和多样性,使复杂数据不仅被可视化呈现,还能够以易理解、富有逻辑的方式传递洞见。

4)构建可解释、可控以及面向人类感知一致性

的可视化评估方法。随着可视化生成自动化和智能化,评估方法需从单一准确性评价扩展至认知契合、交互体验及语义一致性等多维指标。未来研究可结合人类感知规律、任务完成效率等因素,探索可解释、可控的量化评估体系,使可视化生成结果不仅在技术上可优化,更能够与用户认知和理解一致,从而提升数据呈现及决策支持效果。

**致谢:** 本文由中国图象图形学学会可视化与可视分析专业委员会组织撰写,该专委会链接为 <https://www.csig.org.cn/16/201801/49328.html>。

## 参考文献(References)

- Ai K S, Tang K Y and Wang C L. 2026. NLI4VolVis: natural language interaction for volume visualization via LLM multi-agents and editable 3D Gaussian splatting. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 46-56 [DOI: 10.1109/TVCG.2025.3633888]
- Alexander J, Nanda P, Yang K C and Sarvghad A. 2024. Can GPT-4 models detect misleading visualizations?//*Proceedings of 2024 IEEE Visualization and Visual Analytics (VIS)*. St. Pete Beach, USA: IEEE: 106-110 [DOI: 10.1109/VIS55277.2024.00029]
- Aodeng G, Li G Z, Feng Y S, Chen Q Y, Zhang Y and Liu C H. 2025. InReAcTable: LLM-powered interactive visual data story construc-

- tion from tabular data//Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology. Busan, Korea (South): Association for Computing Machinery: #139 [DOI: 10.1145/3746059.3747719]
- Bako H K, Bhutani A, Liu X Y, Cobbina K A and Liu Z C. 2024. Evaluating the semantic profiling abilities of LLMs for natural language utterances in data visualization//Proceedings of 2024 IEEE Visualization and Visual Analytics (VIS). St. Pete Beach, USA: IEEE: 261-265 [DOI: 10.1109/VIS55277.2024.00060]
- Berger M and Liu S S. 2024. The visualization JUDGE: can multimodal foundation models guide visualization design through visual perception? //Proceedings of 2024 IEEE Evaluation and Beyond—Methodological Approaches for Visualization (BELIV). St. Pete Beach, USA: IEEE: 60-70 [DOI: 10.1109/BELIV64461.2024.00012]
- Biswas A, Turton T L, Ranasinghe N R, Jones S, Love B, Jones W, et al. 2025. VizGenie: toward self-refining, domain-aware workflows for next-generation scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, 32 (1) : 1021-1031 [DOI: 10.1109/TVCG.2025.3634655]
- Borisov V, Seßler K, Leemann T, Pawelczyk M and Kasneci G. 2023. Language models are realistic tabular data generators [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2210.06280.pdf>
- Butt M A, Wang K, Vazquez-Corral J and van de Weijer J. 2025. Color-Peel: color prompt learning with diffusion models via color and shape disentanglement//Proceedings of the 18th European Conference on Computer Vision (ECCV 2024). Milan, Italy: Springer: 456-472 [DOI: 10.1007/978-3-031-72667-5\_26]
- Cai X W, Xiong K, Luo Z S, Weng D, Ye S N and Wu Y C. 2025. CodeLin: an in situ visualization method for understanding data transformation scripts. *Visual Informatics*, 9 (2) : #100233 [DOI: 10.1016/j.visinf.2025.03.002]
- Ceballos D, Zarate M, Delrieux C and Nuñez G. 2025. Enhancing data analytics with AI-powered visualization recommendations: a large language model approach [EB/OL]. [2026-01-05]. <https://10.36227/techrxiv.175303443.39749915/v1>
- Ceneda D, Ariel A, and El-Assady M. 2024. A heuristic approach for dual expert/end-user evaluation of guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 30 (1): 997-1007 [DOI: 10.1109/TVCG.2023.3327152]
- Chen J T, Wu J, Guo J J, Mohanty V, Li X M, Ono J P, et al. 2025a. InterChat: enhancing generative visual analytics using multimodal interactions. *Computer Graphics Forum*, 44 (3) : #e70112 [DOI: 10.1111/cgf.70112]
- Chen J Y, Kong L Y, Wei H R, Liu C L, Ge Z, Zhao L, et al. 2024a. OneChart: purify the chart structural extraction via one auxiliary token//Proceedings of the 32nd ACM International Conference on Multimedia. Melbourne, Australia: ACM: 147-155 [DOI: 10.1145/3664647.3681167]
- Chen N, Zhang Y G, Xu J H, Ren K and Yang Y Q. 2025b. VisEval: a benchmark for data visualization in the era of large language models. *IEEE Transactions on Visualization and Computer Graphics*, 31 (1): 1301-1311 [DOI: 10.1109/TVCG.2024.3456320]
- Chen Q, Cao S X, Wang J Z and Cao N. 2024b. How does automation shape the process of narrative visualization: a survey of tools. *IEEE Transactions on Visualization and Computer Graphics*, 30 (8) : 4429-4448 [DOI: 10.1109/TVCG.2023.3261320]
- Chen R, Shu X H, Chen J H, Weng D, Tang J X, Fu S W, et al. 2022. Nebula: a coordinating grammar of graphics. *IEEE Transactions on Visualization and Computer Graphics*, 28 (12) : 4127-4140 [DOI: 10.1109/TVCG.2021.3076222]
- Chen S M, Li J, Andrienko G, Andrienko N, Wang Y, Nguyen P H, et al. 2020. Supporting story synthesis: bridging the gap between visual analytics and storytelling. *IEEE Transactions on Visualization and Computer Graphics*, 26 (7) : 2499-2516 [DOI: 10.1109/TVCG.2018.2889054]
- Chen Y T, Li H Y, Shi N, Luo X H E, Xu W and Shen H W. 2025c. Explorable INR: an implicit neural representation for ensemble simulation enabling efficient spatial and parameter exploration. *IEEE Transactions on Visualization and Computer Graphics*, 31 (6) : 3758-3770 [DOI: 10.1109/TVCG.2025.3567052]
- Chen Z, Cao L, Madden S, Kraska T, Shang Z Y, Fan J, et al. 2024c. SEED: domain-specific data curation with large language models [EB/OL]. [2026-01-05] <https://arxiv.org/pdf/2310.00749.pdf>
- Cheng L Q, Deng D Z, Xie X, Qiu R H, Xu M L and Wu Y C. 2025. SNIL: generating sports news from insights with large language models. *IEEE Transactions on Visualization and Computer Graphics*, 31 (7) : 3973-3986 [DOI: 10.1109/TVCG.2024.3392683]
- Choi J, Lee J and Jo J. 2024. Bavisitter: integrating design guidelines into large language models for visualization authoring//Proceedings of 2024 IEEE Visualization and Visual Analytics (VIS). St. Pete Beach, USA: IEEE: 121-125 [DOI: 10.1109/VIS55277.2024.00032]
- Coelho D, Ghai B, Krishna A, Velez-Rojas M, Greenspan S, Mankovski S, et al. 2024. TaskFinder: a semantics-based methodology for visualization task recommendation. *Analytics*, 3 (3) : 255-275 [DOI: 10.3390/analytics3030015]
- Dibia V. 2023. LIDA: a tool for automatic generation of grammar-agnostic visualizations and infographics using large language models//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations). Toronto, Canada: Association for Computational Linguistics: 113-126 [DOI: 10.18653/v1/2023.acl-demo.11]
- Elmqvist N and Klokmoose C N. 2025. Automating the path: an R&D agenda for human-centered AI and visualization. *IEEE Computer Graphics and Applications*, 45 (3) : 73-81 [DOI: 10.1109/MCG.2025.3559374]
- Ford J, Zhao X M, Schumacher D and Rios A. 2024. Charting the future: using chart question-answering for scalable evaluation of

- LLM-driven data visualizations//Proceedings of the 31st International Conference on Computational Linguistics. Association for Computational Linguistics: 7494-7510
- Gao L, Lu J, Shao Z K, Lin Z Y, Yue S B, Leong C, et al. 2025. Fine-tuned large language model for visualization system: a study on self-regulated learning in education. *IEEE Transactions on Visualization and Computer Graphics*, 31 (1): 514-524 [DOI: 10.1109/TVCG.2024.3456145]
- Gao L, Shen L X, Zhao Y H, Lan J X, Qu H M and Chen S M. 2026. SceneLoom: communicating data with scene context. *IEEE Transactions on Visualization and Computer Graphics*, 32 (1): 1328-1338 [DOI: 10.1109/TVCG.2025.3634816]
- Gathani S, Crisan A, Setlur V and Srinivasan A. 2024. Groot: a system for editing and configuring automated data insights//Proceedings of 2024 IEEE Visualization and Visual Analytics (VIS). St. Pete Beach, USA: IEEE: 36-40 [DOI: 10.1109/VIS55277.2024.00015]
- Goswami K, Mathur P, Rossi R and Dernoncourt F. 2025. PlotGen: multi-agent LLM-based scientific data visualization via multimodal retrieval feedback//Companion Proceedings of the ACM on Web Conference 2025. Sydney, Australia: Association for Computing Machinery: 1672-1676 [DOI: 10.1145/3701716.3716888]
- Guo Y, Qi X Y, Li H Y, Zhang J, Shi D Q, Chen Q, et al. 2026. Ura-nia: visualizing data analysis pipelines for natural language-based data exploration. *ACM Transactions on Interactive Intelligent Systems*, 16(1): #9 [DOI: 10.1145/3770071]
- Han J, Tang K Y and Wang C L. 2026. MoE-INR: implicit neural representation with mixture-of-experts for time-varying volumetric data compression. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 254-264 [DOI: 10.1109/TVCG.2025.3633893]
- Han J and Wang C L. 2023a. CoordNet: data generation and visualization generation for time-varying volumes via a coordinate-based neural network. *IEEE Transactions on Visualization and Computer Graphics*, 29 (12): 4951-4963 [DOI: 10.1109/TVCG. 2022. 3197203]
- Han J and Yang F. 2025. DCINR: a divide-and-conquer implicit neural representation for compressing time-varying volumetric data in hours. *IEEE Transactions on Visualization and Computer Graphics*, 31(10): 8116-8128 [DOI: 10.1109/TVCG.2025.3564255]
- Han J, Zheng H and Bi C K. 2024. KD-INR: time-varying volumetric data compression via knowledge distillation-based implicit neural representation. *IEEE Transactions on Visualization and Computer Graphics*, 30 (10): 6826-6838 [DOI: 10.1109/TVCG. 2023. 3345373]
- Han J, Zheng H, Xing Y H, Chen D Z and Wang C L. 2021. V2V: a deep learning approach to variable-to-variable selection and translation for multivariate time-varying data. *IEEE Transactions on Visualization and Computer Graphics*, 27 (2): 1290-1300 [DOI: 10.1109/TVCG.2020.3030346]
- Han Y C, Zhang C, Chen X, Yang X, Wang Z B, Yu G, et al. 2023b. ChartLlama: a multimodal LLM for chart understanding and generation [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2311.16483.pdf>
- Hao J N, Yang M L, Shi Q, Jiang Y Z, Zhang G, and Zeng W. 2024. Finflir: Automating graphical overlays for financial visualizations with knowledge-grounding large language model. *IEEE Transactions on Visualization and Computer Graphics*, 31(9): 6353-6369 [DOI: 10.1109/TVCG.2024.3514138]
- He W B, Wang J P, Guo H Q, Shen H W and Peterka T. 2020. CECAV-DNN: collective ensemble comparison and visualization using deep neural networks. *Visual Informatics*, 4 (2): 109-121 [DOI: 10.1016/j.visinf.2020.04.004]
- Hou Y H, Yang M L, Cui H, Wang L, Xu J and Zeng W. 2024. C2Ideas: supporting creative interior color design ideation with a large language model//Proceedings of 2024 CHI Conference on Human Factors in Computing Systems. Honolulu, USA: ACM: #172 [DOI: 10.1145/3613904.3642224]
- Hou Y H, Zeng X C, Wang Y S, Yang M L, Chen X J and Zeng W. 2025. GenColor: generative color-concept association in visual design//Proceedings of 2025 CHI Conference on Human Factors in Computing Systems. Yokohama, Japan: ACM: #544 [DOI: 10.1145/3706598.3713418]
- Hu L M, Wang D K, Pan Y M, Yu J F, Shi Y X, Feng C, et al. 2024. NovaChart: a large-scale dataset towards chart understanding and generation of multimodal large language models//Proceedings of the 33rd ACM International Conference on Multimedia. Melbourne VIC, Australia: Association for Computing Machinery: 3917-3925 [DOI: 10.1145/3664647.36807]
- Islam M S, Laskar M T R, Parvez M R, Hoque E and Joty S. 2024. DataNarrative: automated data-driven storytelling with visualizations and texts//Proceedings of 2024 Conference on Empirical Methods in Natural Language Processing. Miami, USA: Association for Computational Linguistics: 19253-19286 [DOI: 10.18653/v1/2024.emnlp-main.1073]
- Jeong S, Li J X, Johnson C R, Liu S S and Berger M. 2024. Text-based transfer function design for semantic volume rendering//Proceedings of 2024 IEEE Visualization and Visual Analytics (VIS). St. Pete Beach, USA: IEEE: 196-200 [DOI: 10.1109/VIS55277. 2024.00047]
- Jia D G, Irger A, Besançon L, Strnad O, Luo D, Björklund J, et al. 2025. VOICE: visual oracle for interaction, conversation, and explanation. *IEEE Transactions on Visualization and Computer Graphics*, 31 (10): 8828-8845 [DOI: 10.1109/TVCG. 2025. 3579956]
- Jianu R, Hutchinson M, Andrienko N, Andrienko G, Elshehaly M and Slingsby A. 2025. Mind-mapping data analysis with LLMs: from vision to first steps//Proceedings of Computer Graphics and Visual Computing (CGVC) 2025. Liverpool, UK: The Eurographics Asso-

- ciation: #20251221 [DOI: 10.2312/cgvc.20251221]
- Koh W, Yoon J, Lee M H, Song Y, Cho J, Kang J, et al. 2025. C<sup>2</sup>: scalable auto-feedback for LLM-based chart generation//Proceedings of 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers). Albuquerque, USA: Association for Computational Linguistics: 4525-4566 [DOI: 10.18653/v1/2025.naacl-long.232]
- Kuang Y M, Guan J W, Liu H Y, Chen F, Wang Z H and Wang W D. 2025. PiCCL: a lightweight multiview contrastive learning framework for image classification. PLoS ONE, 20 (8) : #0329273 [DOI: 10.1371/journal.pone.0329273]
- Lee D J L, Tang D X, Agarwal K, Boonmark T, Chen C, Kang J, et al. 2021. Lux: always-on visualization recommendations for exploratory dataframe workflows. Proceedings of the VLDB Endowment, 15(3): 727-738 [DOI: 10.14778/3494124.3494151]
- Li B X, Wang Y W, Gu J X, Chang K W and Peng N Y. 2025a. METAL: a multi-agent framework for chart generation with test-time scaling//Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics. Vienna, Austria: Association for Computational Linguistics: 30054-30069 [DOI: 10.18653/v1/2025.acl-long.1452]
- Li C L, Dong X J, Liu W, Sheng S Y and Qian A J. 2020a. SSRDVis: interactive visualization for event sequences summarization and rare detection. Journal of Visualization, 23(1): 171-184 [DOI: 10.1007/s12650-019-00609-x]
- Li G, Liu Y, Shan G H, Cheng S Y, Cao W Q, Wang J P, et al. 2025b. *ParamsDrag*: interactive parameter space exploration via image-space dragging. IEEE Transactions on Visualization and Computer Graphics, 31(1): 624-634 [DOI: 10.1109/TVCG.2024.3456338]
- Li G Z, Tian M, Xu Q M, McGuffin M J and Yuan X R. 2020b. GoTree: a grammar of tree visualizations//Proceedings of 2020 CHI Conference on Human Factors in Computing Systems. Honolulu, USA: Association for Computing Machinery: 1-13 [DOI: 10.1145/3313831.3376297]
- Li H T, Wang Y and Qu H M. 2025c. Reflection on data storytelling tools in the generative AI era from the human-AI collaboration perspective//Proceedings of 2025 IEEE Visualization and Visual Analytics (VIS). Vienna, Austria: IEEE: 21-25 [DOI: 10.1109/VIS60296.2025.00009]
- Li H X, Wen Z, Jiang Q Q, Li C X, Wu Y W, Yang Y C, et al. 2026. ConceptViz: a visual analytics approach for exploring concepts in large language models. IEEE Transactions on Visualization and Computer Graphics, 32(1) : 57-67 [DOI: 10.1109/TVCG.2025.3634806]
- Li S M, Chen X A, Song Y F, Song Y Z, Zhang C J, Hao F, et al. 2025d. Prompt4Vis: prompting large language models with example mining for tabular data visualization. The VLDB Journal, 34(4): #38 [DOI: 10.1007/s00778-025-00912-0]
- Li X, Liao J X, Liu W, Miao Y, Wang L Y and Sun S Q. 2025e. Cold-start visualization recommendation driven by large language models for ocean data analysis. Frontiers in Marine Science, 12: #1554241 [DOI: 10.3389/fmars.2025.1554241]
- Li Y X, Qi Y S, Shi Y, Chen Q, Cao N and Chen S M. 2023. Diverse interaction recommendation for public users exploring multi-view visualization using deep learning. IEEE Transactions on Visualization and Computer Graphics, 29 (1) : 95-105 [DOI: 10.1109/TVCG.2022.3209461]
- Li Z L, Zhang W H and Tao J. 2025f. FlowLLM: large language model driven flow visualization. Visual Informatics, 9 (3) : #100241 [DOI: 10.1016/j.visinf.2025.100241]
- Lin Y N, Li H T, Yang L N, Wu A Y and Qu H M. 2024. InkSight: leveraging sketch interaction for documenting chart findings in computational notebooks. IEEE Transactions on Visualization and Computer Graphics, 30 (1) : 944-954 [DOI: 10.1109/TVCG.2023.3327170]
- Liu H X, Chen X Y, Zheng H Y, Li M Y, Liu Z F, Yang F M, et al. 2026. Self-supervised continuous colormap recovery from a 2D scalar field visualization without a legend. IEEE Transactions on Visualization and Computer Graphics, 32(1): 265-275 [DOI: 10.1109/TVCG.2025.3633886]
- Lu Y, Jiang K, Levine J A and Berger M. 2021. Compressive neural representations of volumetric scalar fields. Computer Graphics Forum, 40(3): 135-146 [DOI: 10.1111/cgf.14295]
- Luo T Q, Huang C H, Shen L X, Li B Y, Shen S Y, Zeng W, et al. 2026. nvBench 2.0: resolving ambiguity in text-to-visualization through stepwise reasoning [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2503.12880.pdf>
- Luo Y Y, Tang N, Li G L, Tang J W, Chai C L and Qin X D. 2022. Natural language to visualization by neural machine translation. IEEE Transactions on Visualization and Computer Graphics, 28(1) : 217-226 [DOI: 10.1109/TVCG.2021.3114848]
- Maddigan P and Susnjak T. 2023. Chat2VIS: generating data visualizations via natural language using ChatGPT, Codex and GPT-3 large language models. IEEE Access, 11: 45181-45193 [DOI: 10.1109/ACCESS.2023.3274199]
- Mao L J, Tang Z N and Suhr A. 2025. Evaluating model perception of color illusions in photorealistic scenes//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 7805-7814 [DOI: 10.1109/CVPR52734.2025.00731]
- Masson D, Malacria S, Casiez G and Vogel D. 2023. Charagraph: interactive generation of charts for realtime annotation of data-rich paragraphs//Proceedings of 2023 CHI Conference on Human Factors in Computing Systems. Hamburg, Germany: Association for Computing Machinery: #146 [DOI: 10.1145/3544548.3581091]
- Masson D, Malacria S, Casiez G and Vogel D. 2024. DirectGPT: a direct manipulation interface to interact with large language mod-

- els//Proceedings of 2024 CHI Conference on Human Factors in Computing Systems. Honolulu, USA: Association for Computing Machinery: #975 [DOI: 10.1145/3613904.3642462]
- Narechania A, Srinivasan A and Stasko J. 2021. NL4DV: a toolkit for generating analytic specifications for data visualization from natural language queries. *IEEE Transactions on Visualization and Computer Graphics*, 27 (2): 369-379 [DOI: 10.1109/TVCG.2020.3030378]
- Niu T H, Cui Y M, Wang B X, Xu X, Yao X, Zhu Q F, et al. 2025. Chart2Code53: a large-scale diverse and complex dataset for enhancing chart-to-code generation//Proceedings of 2025 Conference on Empirical Methods in Natural Language Processing. Suzhou, China: 15839-15855 [DOI: 10.18653/v1/2025.emnlp-main.799]
- Pan B, Fu Y X, Wang K, Lu J Y, Pan L K, Qian Z Y, et al. 2025. VIS-Shepherd: constructing critic for LLM-based data visualization generation [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2506.13326.pdf>
- Pan B, Lu J Y, Li H X, Chen W F, Wang Y Y, Zhu M F, et al. 2024. Differentiable design galleries: a differentiable approach to explore the design space of transfer functions. *IEEE Transactions on Visualization and Computer Graphics*, 30 (1): 1369-1379 [DOI: 10.1109/TVCG.2023.3327371]
- Podo L, Angelini M and Velardi P. 2024. V-RECS: a low-cost LLM4Vis recommender with explanations, captioning and suggestions [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2406.15259.pdf>
- Qiu R, Tu Y M, Wang Y S, Yen P Y and Shen H W. 2024. DocFlow: a visual analytics system for question-based document retrieval and categorization. *IEEE Transactions on Visualization and Computer Graphics*, 30 (2): 1533-1548 [DOI: 10.1109/TVCG.2022.3219762]
- Qiu T, Wang F, Huang S H, Guo M, Zhao Y H, Li J C, et al. 2025. SmartMLVs: LLM-enabled multiple linked views generation for interactive visualization//Proceedings of the 18th IEEE Pacific Visualization Conference. Taipei City, China: IEEE: 58-68 [DOI: 10.1109/PacificVis64226.2025.00012]
- Ren D H, Lee B and Brehmer M. 2019. Charticulator: interactive construction of bespoke chart layouts. *IEEE Transactions on Visualization and Computer Graphics*, 25 (1): 789-799 [DOI: 10.1109/TVCG.2018.2865158]
- Ren Y F, Li Z Q and Zhang S H. 2024. Survey of visualization methods for multiscene visual cue information in immersive environments. *Journal of Image and Graphics*, 29(1): 1-21 (任洋甫, 李志强, 张松海. 2024. 沉浸式环境中多场景视觉提示信息可视化方法综述. *中国图象图形学报*, 29(1): 1-21) [DOI: 10.11834/jig.221147]
- Salim S, Pial T and Mueller K. 2026. What is the color of serendipity? Investigating the use of language models for semantically resonant color generation. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 670-680 [DOI: 10.1109/TVCG.2025.3634243]
- Satyanarayan A, Moritz D, Wongsuphasawat K and Heer J. 2017. Vega-Lite: a grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1): 341-350 [DOI: 10.1109/TVCG.2016.2599030]
- Satyanarayan A, Russell R, Hoffswell J and Heer J. 2016. Reactive Vega: a streaming dataflow architecture for declarative interactive visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(1): 659-668 [DOI: 10.1109/TVCG.2015.2467091]
- Savva M, Kong N, Chhajta A, Li F F, Agrawala M and Heer J. 2011. ReVision: automated classification, analysis and redesign of chart images//Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. Santa Barbara, USA: Association for Computing Machinery: 393-402 [DOI: 10.1145/2047196.2047247]
- Schnizer K and Mayer S. 2025. User-centered AI for data exploration: rethinking GenAI's role in visualization [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2504.04253.pdf>
- Seo W, Kang D, An H, Kim T, Cho S, Lee S, et al. 2026. Automated visualization code synthesis via multi-path reasoning and feedback-driven optimization [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2502.11140.pdf>
- Shao Z K, Shan Y, He Y X, Yao Y X, Wang J H, Zhang X L, et al. 2025a. Do language model agents align with humans in rating visualizations? An empirical study. *IEEE Computer Graphics and Applications*, 45(6): 14-28 [DOI: 10.1109/MCG.2025.3586461]
- Shao Z K, Shen L X, Li H T, Shan Y, Qu H M, Wang Y, et al. 2025b. Narrative Player: reviving data narratives with visuals. *IEEE Transactions on Visualization and Computer Graphics*, 31(10): 6781-6795 [DOI: 10.1109/TVCG.2025.3530512]
- Shen J Y, Duan Y H and Shen H W. 2025a. SurroFlow: a flow-based surrogate model for parameter space exploration and uncertainty quantification. *IEEE Transactions on Visualization and Computer Graphics*, 31(1): 635-644 [DOI: 10.1109/TVCG.2024.3456372]
- Shen L X, Li H T, Wang Y, Luo T Q, Luo Y Y and Qu H M. 2025b. Data Playwright: authoring data videos with annotated narration. *IEEE Transactions on Visualization and Computer Graphics*, 31(9): 5884-5897 [DOI: 10.1109/TVCG.2024.3477926]
- Shen L X, Wang Y F, Qu H M, Xie X and Li H T. 2026. Interaction-augmented instruction: modeling the synergy of prompts and interactions in human-GenAI collaboration [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2510.26069.pdf>
- Shen L X, Zhang Y Z, Zhang H D and Wang Y. 2024. Data player: automatic generation of data videos with narration-animation interplay. *IEEE Transactions on Visualization and Computer Graphics*, 30(1): 109-119 [DOI: 10.1109/TVCG.2023.3327197]
- Shi C H, Cui W W, Liu C Z, Zheng C B, Zhang H D, Luo Q, et al. 2024. NL2Color: refining color palettes for charts with natural language. *IEEE Transactions on Visualization and Computer Graph-*

- ics, 30(1): 814-824 [DOI: 10.1109/TVCG.2023.3326522]
- Shi N, Xu J Y, Li H Y, Guo H Q, Woodring J and Shen H W. 2023. VDL-Surrogate: a view-dependent latent-based model for parameter space exploration of ensemble simulations. *IEEE Transactions on Visualization and Computer Graphics*, 29(1): 820-830 [DOI: 10.1109/TVCG.2022.3209413]
- Shi N, Xu J Y, Wurster S W, Guo H Q, Woodring J, Van Roekel L P, et al. 2022. GNN-Surrogate: a hierarchical and adaptive graph neural network for parameter space exploration of unstructured-mesh ocean simulations. *IEEE Transactions on Visualization and Computer Graphics*, 28(6): 2301-2313 [DOI: 10.1109/TVCG.2022.3165345]
- Shi S Z, Ren T, Zhu G L, Feng G D and Hu J. 2025. Closing the feedback loop in Text2Vis: refining visualization with vision-language models//*Proceedings of the 33rd ACM International Conference on Multimedia*. Dublin, Ireland: Association for Computing Machinery: 9053-9061 [DOI: 10.1145/3746027.3755862]
- Shin S, Jeon H, Hong S and Elmqvist N. 2025. Data therapist: eliciting domain knowledge from subject matter experts using large language models [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2505.00455.pdf>
- Shuai Z H, Li B Y, Yan S Y, Luo Y Y and Yang W K. 2026. DeepVIS: bridging natural language and data visualization through step-wise reasoning. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 868-878 [DOI: 10.1109/TVCG.2025.3634645]
- Song H, Dai Z, Xu P P and Ren L. 2022. Interactive visual pattern search on graph data via graph representation learning. *IEEE Transactions on Visualization and Computer Graphics*, 28(1): 335-345 [DOI: 10.1109/TVCG.2021.3114857]
- Sperle F, Jeitler A, Bernard J, Keim D and El-Assady M. 2021. Co-adaptive visual data analysis and guidance processes. *Computers and Graphics*, 100: 93-105 [DOI: 10.1016/j.cag.2021.06.016]
- Srinivasan A, Setlur V and Satyanarayan A. 2025. Pluto: authoring semantically aligned text and charts for data-driven communication//*Proceedings of the 30th International Conference on Intelligent User Interfaces*. Cagliari, Italy: Association for Computing Machinery: 1123-1140 [DOI: 10.1145/3708359.3712122]
- Stokes C, Arunkumar A, Hearst M A and Padilla L. 2026. An analysis of text functions in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 769-779 [DOI: 10.1109/TVCG.2025.3567122]
- Sun M D, Cai L G, Cui W W, Wu Y Q, Shi Y and Cao N. 2023. Erato: cooperative data story editing via fact interpolation. *IEEE Transactions on Visualization and Computer Graphics*, 29(1): 983-993 [DOI: 10.1109/TVCG.2022.3209428]
- Takahira K and Ueno Y. 2025. VisAider: AI-assisted context-aware visualization support for data presentations [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2510.14247.pdf>
- Tang K Y, Ai K S, Han J and Wang C L. 2026. TexGS-VolVis: expressive scene editing for volume visualization via textured Gaussian splatting. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 933-943 [DOI: 10.1109/TVCG.2025.3634643]
- Tao J, Zhang Y, Chen Q, Liu C, Chen S M and Yuan X R. 2023. Intelligent visualization and visual analytics. *Journal of Image and Graphics*, 28(6): 1909-1926 (陶钧, 张宇, 陈晴, 刘灿, 陈思明, 袁晓如. 2023. 智能可视化与可视分析. *中国图象图形学报*, 28(6): 1909-1921) [DOI: 10.11834/jig.230034]
- Tian Y, Cui W W, Deng D Z, Yi X J, Yang Y R, Zhang H D, et al. 2025a. ChartGPT: leveraging LLMs to generate charts from abstract natural language. *IEEE Transactions on Visualization and Computer Graphics*, 31(3): 1731-1745 [DOI: 10.1109/TVCG.2024.3368621]
- Tian Y, Zhang C H, Wang X T, Pan S T, Cui W W, Zhang H D, et al. 2025b. ReSpark: leveraging previous data reports as references to generate new reports with LLMs//*The 38th Annual ACM Symposium on User Interface Software and Technology*. Busan, Korea (South): ACM: #181 [DOI: 10.1145/3746059.3747644]
- Vaithilingam P, Glassman E L, Inala J P and Wang C L. 2024. DynaVis: dynamically synthesized UI widgets for visualization editing//*Proceedings of 2024 CHI Conference on Human Factors in Computing Systems*. Honolulu, USA: Association for Computing Machinery: #985 [DOI: 10.1145/3613904.3642639]
- Wall E, Matzen L, El-Assady M, Masters P, Hosseinpour H, Ender A, et al. 2024. Trust junk and evil knobs: calibrating trust in AI visualization//*Proceedings of the 17th IEEE Pacific Visualization Conference (PacificVis)*. Tokyo, Japan: IEEE: 22-31 [DOI: 10.1109/PacificVis60374.2024.00012]
- Wang C L, Lee B, Drucker S M, Marshall D and Gao J F. 2025a. Data Formulator 2: iterative creation of data visualizations, with AI transforming data along the way//*Proceedings of 2025 CHI Conference on Human Factors in Computing Systems*. Yokohama, Japan: ACM: #677 [DOI: 10.1145/3706598.3713296]
- Wang F, Wang B M, Shu X L, Liu Z, Shao Z K, Liu C, et al. 2025b. ChartInsighter: an approach for mitigating hallucination in time-series chart summary generation with a benchmark dataset. *IEEE Transactions on Visualization and Computer Graphics*, 31(6): 3733-3745 [DOI: 10.1109/TVCG.2025.3567122]
- Wang H W, Gordon M, Battle L and Heer J. 2025c. DracoGPT: extracting visualization design preferences from large language models. *IEEE Transactions on Visualization and Computer Graphics*, 31(1): 710-720 [DOI: 10.1109/TVCG.2024.3456350]
- Wang H W, Hoffswell J, Thane S M T, Bursztyn V S and Bearfield C X. 2025d. How aligned are human chart takeaways and LLM predictions? A case study on bar charts with varying layouts. *IEEE Transactions on Visualization and Computer Graphics*, 31(1): 536-546 [DOI: 10.1109/TVCG.2024.3456378]
- Wang J Z, Li X, Li C L, Peng D, Wang A Z, Gu Y H, et al. 2024a. AVA: an automated and AI-driven intelligent visual analytics

- framework. *Visual Informatics*, 8(2): 106-114 [DOI: 10.1016/j.visinf.2024.06.002]
- Wang L W, Wang Z, Xiao S S, Liu L, Tsung F and Zeng W. 2025e. VizTA: enhancing comprehension of distributional visualization with visual-lexical fused conversational interface. *Computer Graphics Forum*, 44(3): #70110 [DOI: 10.1111/cgf.70110]
- Wang Z R, Xia M Z, He L X, Chen H, Liu Y T, Zhu R, et al. 2024b. CharXiv: charting gaps in realistic chart understanding in multimodal LLMs//Proceedings of the 38th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: 113569-113697
- Weiss S, Hermüller P and Westermann R. 2022. Fast neural representations for direct volume rendering. *Computer Graphics Forum*, 41(6): 196-211 [DOI: 10.1111/cgf.14578]
- Wen Z, Weng L X, Tang Y H, Zhang R J, Liu Y X, Pan B, et al. 2025. Exploring multimodal prompt for visualization authoring with large language models[EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2504.13700.pdf>
- Wolter A, Vidalakis G, Yu M, Grover A and Dhanoa V. 2025. Multi-agent data visualization and narrative generation//[EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2509.00481.pdf>
- Wu Z K, Le V, Tiwari A, Gulwani S, Radhakrishna A, Radiček I, et al. 2022. NL2Viz: natural language to visualization via constrained syntax-guided synthesis//Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering. Singapore: Association for Computing Machinery: 972-983 [DOI: 10.1145/3540250.3549140]
- Xia R Q, Ye H C, Yan X C, Liu Q, Zhou H B, Chen Z J, et al. 2025. ChartX and ChartVLM: a versatile benchmark and foundation model for complicated chart reasoning. *IEEE Transactions on Image Processing*, 34: 7436-7447 [DOI: 10.1109/TIP.2025.3607618]
- Xiao S S, Huang S Z, Lin Y, Ye Y L and Zeng W. 2024. Let the chart spark: embedding semantic context into chart with text-to-image generative model. *IEEE Transactions on Visualization and Computer Graphics*, 30(1): 284-294 [DOI: 10.1109/TVCG.2023.3326913]
- Xie L W H, Lin Y N, Liu C, Qu H M and Shu X H. 2026. DataWink: reusing and adapting SVG-based visualization examples with large multimodal models. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 824-834 [DOI: 10.1109/TVCG.2025.3634635]
- Xie L W H, Zheng C B, Xia H J, Qu H M and Chen Z T. 2024a. WaitGPT: monitoring and steering conversational LLM agent in data analysis with on-the-fly code visualization//Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology. Pittsburgh, USA: Association for Computing Machinery: #119 [DOI: 10.1145/3654777.3676374]
- Xie T C, Lin M Z, Liu M C, Ye Y L, Chen C J and Liu S X. 2025. InfoChartQA: a benchmark for multimodal question answering on infographic charts//Proceedings of the 39th Annual Conference on Neural Information Processing Systems Datasets and Benchmark Track (Poster). San Diego, USA: NeurIPS: 1-13
- Xie Y P, Luo Y Y, Li G L and Tang N. 2024b. HAICChart: human and AI paired visualization system. *Proceedings of the VLDB Endowment*, 17(11): 3178-3191 [DOI: 10.14778/3681954.3681992]
- Xu Z Z, Du S N, Qi Y Y, Xu C J, Yuan C and Guo J. 2024. ChartBench: a benchmark for complex visual reasoning in charts [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2312.15915.pdf>
- Xu Z Z, Qu B W, Qi Y Y, Du S N, Xu C J, Yuan C, et al. 2026. ChartMoE: mixture of diversely aligned expert connector for chart understanding//Proceedings of the 14th International Conference on Learning Representations. Rio de Janeiro, Brazil: OpenReview.net [DOI: 10.48550/arXiv.2409.03277]
- Yan P Y, Bhosale M, Lal J, Adhikari B and Doermann D. 2024. ChartReformer: natural language-driven chart image editing//Proceedings of the 18th International Conference on Document Analysis and Recognition. Athens, Greece: Springer: 453-469 [DOI: 10.1007/978-3-031-70533-5\_26]
- Yan S Y, Liu T C, Yang W K, Tang N and Luo Y Y. 2025. ChartEditor: a human-AI paired tool for authoring pictorial charts [EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2501.07320.pdf>
- Yang H, Yang Z Y, Zhao R Y, Li X R and Rao G Q. 2024a. The implementation solution for automatic visualization of tabular data in relational databases based on large language models//Proceedings of 2024 International Conference on Asian Language Processing. Hohhot, China: IEEE: 175-180 [DOI: 10.1109/IALP63756.2024.10661162]
- Yang Z R, Pan B, Wang H, Wang Y Y, Liu X Y, Weng L X, et al. 2025. Multimodal DeepResearcher: generating text-chart interleaved reports from scratch with agentic framework//Proceedings of the 40th AAAI Conference on Artificial Intelligence. Singapore, Singapore: Association for the Advancement of Artificial Intelligence: 1-48
- Yang Z Y, Zhou Z H, Wang S, Cong X, Han X, Yan Y K, et al. 2024b. MatPlotAgent: method and evaluation for LLM-based agentic scientific data visualization//Proceedings of the Findings of the Association for Computational Linguistics: ACL 2024. Bangkok, Thailand: Association for Computational Linguistics: 11789-11804 [DOI: 10.18653/v1/2024.findings-acl.701]
- Yu Y, Shen L X, Long F, Qu H M and Chen H. 2024. PyGWalker: on-the-fly assistant for exploratory visual data analysis//Proceedings of 2024 IEEE Visualization and Visual Analytics (VIS). St. Pete Beach, USA: IEEE: 6-10 [DOI: 10.1109/VIS55277.2024.00009]
- Zadeh F P, Kim J, Kim J H and Kim G. 2024. Text2Chart31: instruction tuning for chart generation with automatic feedback//Proceedings of 2024 Conference on Empirical Methods in Natural Language Processing. Miami, USA: Association for Computational Linguistics

- tics: 11459-11480 [DOI: 10.18653/v1/2024.emnlp-main.640]
- Zeng X C, Lin H C, Ye Y L and Zeng W. 2025. Advancing multimodal large language models in chart question answering with visualization-referenced instruction tuning. *IEEE Transactions on Visualization and Computer Graphics*, 31(1): 525-535 [DOI: 10.1109/TVCG.2024.3456159]
- Zeng Z H and Battle L. 2024. A systematic review of visualization recommendation systems: goals, strategies, interfaces, and evaluations. *Foundations and Trends in Databases*, 14(1): 1-71 [DOI: 10.1561/19000000088]
- Zhang W H and Tao J. 2026. Automatic semantic alignment of flow pattern representations for exploration with large language models. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 835-845 [DOI: 10.1109/TVCG.2025.3634650]
- Zhang W Q, Shen Y L, Tan Z Q, Hou G Y, Lu W M and Zhuang Y T. 2025. Data-Copilot: bridging billions of data and humans with autonomous workflow[EB/OL]. [2026-01-05]. <https://arxiv.org/pdf/2306.07209.pdf>
- Zhao X and Tao J. 2026. Natural language-driven viewpoint navigation for volume exploration via semantic block representation. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 879-889 [DOI: 10.1109/TVCG.2025.3634651]
- Zhao Y, Wang Y H, Luo X, Wang Y Y and Fekete J D. 2025a. Libra: an interaction model for data visualization//*Proceedings of 2025 CHI Conference on Human Factors in Computing Systems*. Yokohama, Japan: ACM: #1169 [DOI: 10.1145/3706598.3713769]
- Zhao Y H, Shu X L, Fan L W, Gao L, Zhang Y and Chen S M. 2026. ProactiveVA: proactive visual analytics with LLM-based UI agent. *IEEE Transactions on Visualization and Computer Graphics*, 32(1): 451-461 [DOI: 10.1109/TVCG.2025.3642628]
- Zhao Y H, Wang J J, Xiang L B, Zhang X W, Guo Z F, Turkay C, et al. 2025b. LightVA: lightweight visual analytics with LLM agent-based task planning and execution. *IEEE Transactions on Visualization and Computer Graphics*, 31(9): 6162-6177 [DOI: 10.1109/TVCG.2024.3496112]
- Zhao Y H, Zhang Y X, Zhang Y, Zhao X Y, Wang J J, Shao Z K, et al. 2025c. LEVA: using large language models to enhance visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 31(3): 1830-1847 [DOI: 10.1109/TVCG.2024.3368060]

## 作者简介

汪云海,男,教授,主要研究方向为可视化与可视分析。

E-mail:wang.yh@ruc.edu.cn

张加万,通信作者,男,教授,主要研究方向为可视化与可视分析。E-mail:jwzhang@tju.edu.cn

曹楠,男,教授,主要研究方向为信息设计与可视化。

E-mail:nan.cao@gmail.com

陈思明,男,研究员,主要研究方向为可视分析。

E-mail:simingchen3@gmail.com

李晨辉,男,教授,主要研究方向为信息可视化。

E-mail:chli@cs.ecnu.edu.cn

曾伟,男,助理教授,主要研究方向为信息可视化。

E-mail:weizeng@ust.hk

陶钧,男,副教授,主要研究方向为科学可视化。

E-mail:taoj23@gmail.com

曾琼,女,教授,主要研究方向为科学可视化。

E-mail:qiong.zn@sdu.edu.cn

王长波,男,教授,主要研究方向为可视化与可视分析。

E-mail:cbwang@dase.ecnu.edu.cn