

智能交互驱动的数据可视化演进

摘要：随着大数据时代的深入发展，数据可视化作为连接数据量化内容与人类直觉的核心桥梁，其交互模式的智能化演进变得至关重要^[1]。本文综述了可视化领域智能交互技术，特别是自然语言界面（V-NLI）与多模态融合的现状、关键挑战及未来方向。研究表明，V-NLI显著降低了数据分析的技术门槛，但其在多轮对话支持、复杂场景适应等方面仍面临挑战。人工智能（AI）技术通过辅助可视表达生成、意图理解与交互优化，正深刻重塑可视化流程。

关键词：可视化分析、智能交互系统、可视化界面、V-NLI技术、人工智能驱动的可视化

1 引言与研究背景

随着大数据时代的发展，数据可视化已成为融合人类认知与机器智能的高效分析工具，它不仅限于简单的图形呈现，更在欺诈检测、风险管理等特定领域——例如生命保险行业——发挥着关键的支撑作用。

交互是可视化的核心，其正从传统的窗口、图标、菜单、指针（WIMP）模式向更自然的“后WIMP”范式演进。而V-NLI（Visual-Narrative Language Interface）技术的兴起，使得用户能够通过自然语言指令操控可视化系统，从而降低操作认知负担^[2]。同时，AI赋能的可视化系统能够学习用户习惯，并推荐最佳的视觉表达与交互路径^[3]。这种人机协同的方式，极大地提升了数据分析的效率和用户体验。

2 V-NLI 技术简介

当前可视化自然语言接口（V-NLI）系统的发展与自然语言处理（NLP）技术的进步紧密相关，并且日益完善的工具化支持为 V-NLI 的构建提供了坚实的基础^[4]。根据不同的应用目标，V-NLI 系统可以大致分为以下两类：

2.1 可视化生成型 V-NLI 系统

在实时数据分析领域，可视化生成型 V-NLI 系统是一种创新的数据交互方式。该系统利用自然语言处理技术和人工智能算法，将用户的自然语言查询转化为结构化数据操作指令，并自动生成相应的可视化结果。这种交互方式不仅降低了用户与数据之间的沟通门槛，还显著提升了数据解释的效率和准确性。例如，FlowSense^[5] 就是一个典型的例子，用户可以通过简单的自然语言输入（如“显示过去一年销售额的趋势图”），获得动态生成的折线图或柱状图，以此实现相应数据流的可视化。此外，V-NLI 系统还支持多维度数据分析和实时更新功能，能够在数据流变化时自动调整可视化内容，满足现代企业对快速决策的需求。

2.2 已有可视化交互型 V-NLI 系统

可视化交互型 V-NLI（Visual-Narrative Language Interface）系统聚焦于用户与已生成图表间的智能交互，通过构建图表问答框架实现数据探索。这类系统首先解析图表的视觉编码（如柱状图高度、饼图扇形面积等），建立视觉元素与底层数据的语义关联^[6]。用户可通过自然语言提问（如“5月销售额最高的产品是什么？”），系统基于编码解析结果精准定位答案区域并生成解释性响应。

为实现更高精度的交互，部分研究引入计算机视觉技术辅助理解。例如，基于 Mask R-CNN^[7] 的自动注释生成方法，通过像素级分割识别图表中的坐标轴、数据点、图例等元素，并将视觉实体与语言描述进行精准匹配。该技术能动态生成高亮标注与说明性文本，显著增强用户对复杂图表的认知效率。

在交互体验优化层面，语义链接技术被证明可降低认知负担。代表性工作如 Kori 系统^[8]，通过预训练词向量构建可视化元素与文本描述的语义映射关系。当用户输入自然语言查询时，系统自动定位关联的可视化组件并建立双向链接，有效减少用户在文本与图表间切换的认知成本。此类系统体现了自然语言处理与计算机视觉技术在可视化交互中的深度融合，为传统图表注入了动态响应能力。

3 V-NLI面临挑战

尽管 V-NLI 技术取得了显著进展，但仍然面临一些挑战：

3.1 自然语言的复杂性

自然语言本身具有模糊性和复杂性，这给 V-NLI 理解用户意图带来了挑战。用户可能使用不同的表达方式来描述相同的分析需求，或者提出的问题可能包含歧义，导致系统难以准确识别用户的真实意图。此外，口语化的表达、省略以及领域特定术语的使用，也增加了自然语言理解的难度。因此，V-NLI 需要更强大的语义分析能力，来处理复杂的语言结构和上下文信息。

3.2 对话连续性不足

多数 V-NLI 系统仅支持单次自然语言命令输入，缺乏多轮对话交互能力，限制了用户在复杂分析任务中的连续探索。这种局限性源于现有工具包无法有效维护上下文信息和对话状态^[9]。在数据分析场景中，用户通常需要通过多次交互迭代式细化需求，但现有系统的 NLP 模型难以处理省略、歧义或反馈性提问，导致对话连续性中断。这种单轮范式削弱了 V-NLI 的实用性，尤其在高阶分析如统计推理中，系统无法根据用户反馈动态调整可视化输出或数据解释，亟需开发上下文感知的多轮对话机制。

3.3 复杂场景适应性弱

在工程仿真、虚拟现实等复杂专业场景中，V-NLI 面临海量数据交互的显著技术瓶颈。这些场景涉及大规模数据集（如仿真输出或空间信息），但现有系统难以实时解析高维数据并生成准确的可视化表达^[10]。跨语言和跨领域适应性不足，语言差异及专业术语的歧义性导致模型无法可靠处理跨域查询，例如在工程仿真中融合多模态数据。同时，无障碍访问需求在复杂系统中更突出，V-NLI 需平衡数据密度与交互效率，但目前方案在解释海量数据的统计模式时易产生偏差，难以满足专业分析的可信度要求。

3.4 无障碍访问的挑战

对于有视觉障碍的人来说，访问统计图表是一个挑战。虽然可以通过自然语言接口 (NLI) 来实现对统计图表的无障碍访问，但这需要 NLI 能够理解用户的分析和导航查询，并以易于理解的方式呈现数据。因此，如何设计易于使用且能够满足不同用户需求的 NLI，是一个需要 V-NLI 领域学者们持续关注的问题。

4 未来方向

4.1 多模态交互的增强

V-NLI 的演进方向正从单一语言交互转向多模态协同交互范式，深度整合语音、手势、眼动追踪等多通道输入，以提供更自然、更直观的用户体验。这种融合突破传统交互限制，这使得用户可以通过最适合当前任务和环境的方式与数据进行交互。例如，在沉浸式环境中，用户可以通过语音指令动态筛选数据维度，同时结合手势操作实现空间探索，而无需依赖传统的键盘和鼠标等物理设备^[11]。此类系统需具备跨模态情境感知能力，通过传感器实时捕获环境状态（如用户位置、设备类型），动态适配最优交互组合。

未来的 V-NLI 系统可以考虑更加注重多模态输入的协同处理，实现更高级别的语义理解和意图推断。这意味着系统不仅能够识别用户的单个指令，还能理解指令之间的关系，从而提供更智能化的服务。

4.2 对话式分析的深入

对话式分析是 V-NLI 的另一个重要发展方向，它利用自然语言处理技术，使用户可以通过对话的方式与数据进行交互。这种方式降低了数据分析的门槛，使得非专业用户也能轻松地进行数据探索和分析。

对话式分析系统相较于传统的 V-NLI 系统更加注重语义理解和推理能力的提升，能够理解更复杂的查询和指令，提供更准确、更全面的回答。此外，系统还将具备一定的推理能力，能够根据用户的查询和数据之间的关系，自动推荐相关的信息和分析^[12]。

4.3 模型驱动的可视化

模型驱动的可视化是指利用机器学习和人工智能技术，自动生成可视化图表和分析结果。这种方式可以大大提高数据分析的效率，减少人工干预，降低人工成本的同时规避人工带来的数据处理误差。

在未来，模型驱动的可视化系统将更加注重可解释性和交互性，使用户能够理解模型的工作原理，并根据需要进行调整和优化。此外，系统还将具备一定的自动化能力，能够根据数据的特点和用户的需求，自动选择最合适的可视化方法^[13]。

这一演进趋势本质是双向赋能的技术融合：一方面，可视化技术为 AI 模型提供理解透明性，推动可解释机器学习在医疗诊断、金融风控等领域的落地；另一方面，AI 驱动的自动化引擎（如深度生成网络）显著提升可视分析的效率，使系统能够处理更大规模的数据（如千万级节点的网络关系），并通过智能故事叙述技术自动提炼关键叙事框架。未来的模型驱动系统将逐步发展为自适应的协同分析伙伴，在保障用户控制权的前提下，实现从被动响应到主动引导的范式升级。

5 小结

智能交互技术正推动数据可视化向更低门槛、更高沉浸感、更强分析效能的方向发展。V-NLI 与多模态融合代表了人机协同分析的范式革新。未来的研究需要着力突破语义理解、上下文建模与复杂场景适配等技术瓶颈，并结合 VR/AR、AI 大模型发展更具普适性的智能交互范式。随着技术与应用场景的深度融合，可视化将从被动呈现工具转变为主动智能分析伙伴。

参考文献：

1. 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(06):1909-1926 DOI: [10.11834/jig.230034](https://doi.org/10.11834/jig.230034).
2. Ni, M., Fan, Y., Zhang, L., & Zuo, W. (2024). *Visual-O1: Understanding Ambiguous Instructions via Multi-modal Multi-turn Chain-of-thoughts Reasoning* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2410.03321>
3. Ross, W., Morris, A., Ulieru, M., & Guyard, A. B. (2013). RECON: An Adaptive Human-Machine System for Supporting Intelligence Analysis. *2013 IEEE International Conference on Systems, Man, and Cybernetics*, 782–787. <https://doi.org/10.1109/smc.2013.138>
4. Kavaz, E., Puig, A., & Rodríguez, I. (2023). Chatbot-Based Natural Language Interfaces for Data Visualisation: A Scoping Review. *Applied Sciences*, 13(12), 7025. <https://doi.org/10.3390/app13127025>
5. Uddin, M. K. S. (2024). A REVIEW OF UTILIZING NATURAL LANGUAGE PROCESSING AND AI FOR ADVANCED DATA VISUALIZATION IN REAL-TIME ANALYTICS. *Global Mainstream Journal*, 1(4), 34–49. <https://doi.org/10.62304/ijmisds.v1i04.185>

6. Wang, Z., Guan, T., Fu, P., Duan, C., Jiang, Q., Guo, Z., Guo, S., Luo, J., Shen, W., & Yang, X. (2025). *Marten: Visual Question Answering with Mask Generation for Multi-modal Document Understanding* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2503.14140>
7. Manmadhan, S., & Kooroor, B. C. (2023). Object-Assisted Question Featurization and Multi-CNN Image Feature Fusion for Visual Question Answering. *International Journal of Intelligent Information Technologies*, 19(1), 1–19. <https://doi.org/10.4018/ijit.318671>
8. Ma, B., Zhang, J., Cao, F., & He, Y. (2020). MACD R-CNN: An Abnormal Cell Nucleus Detection Method. *IEEE Access*, 8, 166658–166669. <https://doi.org/10.1109/access.2020.3020123>
9. Mitra, R., Narechania, A., Endert, A., & Stasko, J. (2022). Facilitating Conversational Interaction in Natural Language Interfaces for Visualization (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.2207.00189>
10. Kavaz, E., Puig, A., & Rodríguez, I. (2023). Chatbot-Based Natural Language Interfaces for Data Visualisation: A Scoping Review. *Applied Sciences*, 13(12), 7025. <https://doi.org/10.3390/app13127025>
11. Su, C., Yang, C., Chen, Y., Wang, F., Wang, F., Wu, Y., & Zhang, X. (2021). Natural multimodal interaction in immersive flow visualization. *Visual Informatics*, 5(4), 56–66. <https://doi.org/10.1016/j.visinf.2021.12.005>
12. Shevchenko, V., Teney, D., Dick, A., & Hengel, A. van den. (2021). Reasoning over Vision and Language: Exploring the Benefits of Supplemental Knowledge (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2101.06013>
13. 夏佳志, 李杰, 陈思明, 秦红星, & 刘世霞. (2021). A survey on interdisciplinary research of visualization and artificial intelligence. *SCIENTIA SINICA Informationis*, 51(11), 1777. <https://doi.org/10.1360/ssi-2021-0062>