Structured AI Agents for Reliable Visualization Report Generation

Junhao Zhao* Lijie Yao†
Department of Computing, School of Advanced Technology, Xi'an Jiaotong-Liverpool University

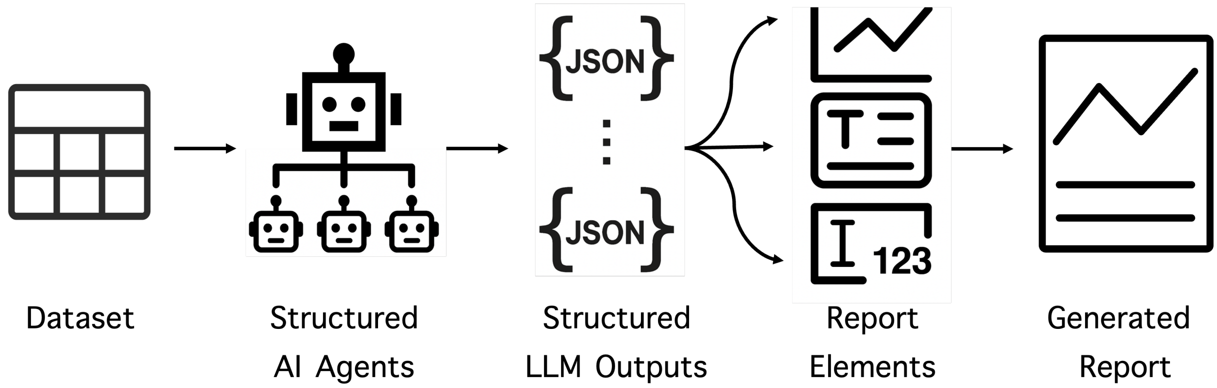

Fig. 1: Our structured pipeline for reliable visualization report generation.

Abstract

We introduce Structured AI Agents to tackle instability and hallucination in AI agent–based data visualization. AI agents promise faster, more accessible visualization reporting, but current methods remain inconsistent and struggle to deliver reliable results. We set multiple sub-tasks and form a structured pipeline, where each sub-task unit produces standardized JSON outputs encoding executable and deterministic code. Our approach enforces correctness in quantitative results, reduces hallucinations, and ensures reliable visualizations. In our final report, we demonstrate that the modular and verifiable outputs achieve strong reliability and maintain accuracy in visualization report generation.

Index Terms: Structured AI Agents, Reliable visualization, Hallucination reduction.

1 Introduction

AI agents have shown promise in automating visualization report generation, offering the potential to streamline data analysis and improve accessibility of insights. With the increasing availability of large-scale visualization datasets, the demand for automatic and reliable visualization reporting has become more significant. However, despite recent advances, current approaches often struggle to produce consistent, interpretable, and reliable outputs. Specific in this mini-challenge, conventional methods proved inadequate for handling the VisPubData dataset [1], as the data is large, diverse in format, and often inconsistent or incomplete. Directly prompting large language models (LLMs) for report generation also presents difficulties, frequently resulting in unstable execution, hallucinated numerical values, and inaccurate visualizations.

To address such challenges, we propose Structured AI Agents, a modular approach for visualization report generation. Instead of treating the report creation as a single black-box task, our method decomposes it into multiple sub-tasks, orchestrated in a structured pipeline. Each sub-task unit produces standardized JSON outputs encoding executable and deterministic code, which can be independently verified and executed. Our approach ensures the correctness of quantitative results, prevents hallucinations, and guarantees that visualizations are based strictly on underlying data rather than unconstrained model guesses.

In summary, our work makes two main contributions:

- **Structured Report Generation**: We introduce a structured approach to report generation by designing AI agents that operate through modular sub-tasks and standardized outputs, thereby improving stability and reproducibility.
- **Reliable Visualization Outcomes**: We ensure reliable visualization outcomes by integrating verifiable numeric analysis and deterministic code execution, which eliminates hallucinations and provides a robust foundation for trustworthy visualization reporting.

2 Method

Here we present the building process of our Structured AI Agents step by step.

2.1 Pipeline Overview

Our approach follows a structured pipeline that transforms raw dataset information into a reliable visualization report. As illustrated in Fig. 1, our process begins with the input dataset and proceeds through a structured AI agent with modular sub-task units. Each unit handles a specific sub-task and produces standardized outputs in JSON format. These standardized outputs contain both executable code and textual elements, which are later executed and composed into the final report. Our structured design ensures transparency, reproducibility, and robustness, avoiding the instability commonly observed in black-box prompting.

2.2 Structured AI Agent with Subtask Units

Instead of treating visualization report generation as a single end-to-end task, we decompose it into multiple structured sub-task units. These units are designed to reflect the main analytical

*e-mail: nemozjh@hotmail.com
†e-mail: yaolijie0219@gmail.com

dimensions of the dataset, and each dimension is further divided into two concrete tasks (8 sub-task units in total):

Overview provides a high-level characterization of the dataset, including overall statistics and distribution across conferences.

Temporal Trends captures dynamics over time, such as publication growth and awarded papers

Top Entities highlight influential actors and themes, including prolific authors and dominant keywords.

Cross-Metric Relationships analyze interactions between different metrics, such as citations vs. downloads and award share across conferences.

2.3 Structured LLM Output in JSON Format

We adopt JSON as the standardized output format for two main reasons: (1) it significantly improves the stability of LLM outputs by enforcing a fixed schema, and (2) it enables precise insertion of computed values into descriptive text, ensuring that narratives are grounded in verified numerical results rather than unconstrained generation. The following block shows an example JSON format corresponding to a single sub-task unit:

```
{
  "section": "<section>",
  "title": "<short title>",
  "goal": "<2-3 sentences>",
  "description_template":
    "<1-2 sentences with {placeholders} computed by METRICS>",
  "narrative_only": <true|false>,
  "compute_only_code": "<python>",
  "code": "<python code>"
}
```

2.4 Execution and Report Composition

Once the JSON outputs are produced, each sub-task unit is independently executed to ensure accuracy and stability. For units with the `code` or `compute_only_code` field, our *Python* codes execute directly on the dataset to compute quantitative results and, when required, generate an exact analytical chart. During execution, all relevant statistics are stored in the `METRICS` dictionary, which provides the values needed by the `description_template`. Our algorithm guarantees that every numerical statement in the final description is grounded in computation rather than unconstrained model generation.

Finally, the verified descriptions and the generated charts are composed into a coherent report. Each sub-task unit thus contributes a consistent block of text and visualization, and our full pipeline integrates these blocks into a structured, reliable visualization report.

3 Representative Results

We present the representative outputs from our structured pipeline. Considering the page limits, we highlight here two typical examples that illustrate the stability and reliability of our generated report. Our complete automatically generated report, including all four sections and full visualizations, is available at URL of REPORT.

Fig. 2 shows our method not only produces an accurate analytical chart but also generates a corresponding textual description with precise numerical values derived from verified computation. For instance, the summary states: *"From 1990 to 2024, the peak year for awards was 2021 with 20 awarded papers."*, which exactly matches the data shown in the chart. Fig. 3 illustrates the top keywords extracted from the VisPubData[1] dataset. A particular challenge of VisPubData is that the "AuthorKeywords" field contains inconsistent formatting, numerous missing entries, and inconsistent capitalization (e. g., "Visualization" vs. "visualization").

By guiding the agent through a specific subtask unit and enforcing constraints during report generation, our method successfully cleans

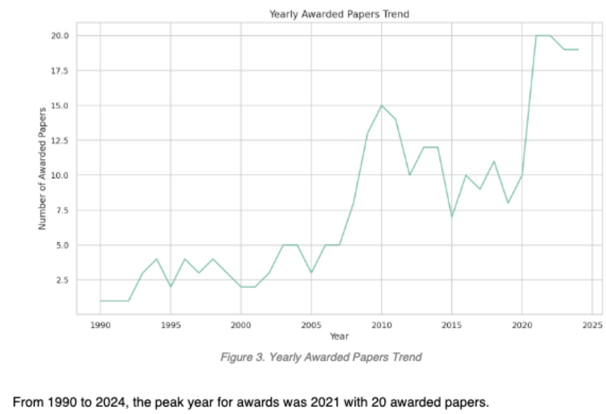

Fig. 2: Yearly awarded papers trend automatically generated by our pipeline.

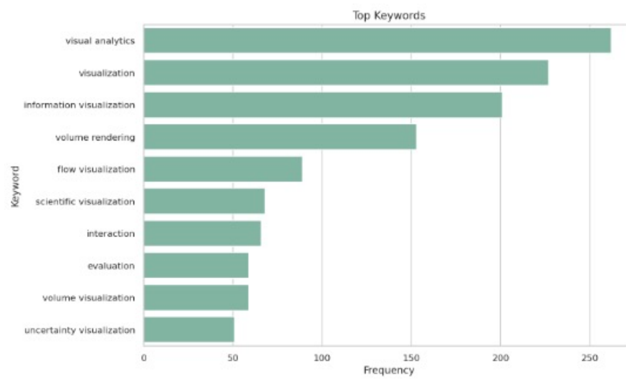

Fig. 3: Top keywords visualization automatically generated by our pipeline.

the data by removing invalid entries and consolidating keywords that differ only in letter case. As a result, our generated visualizations provide accurate and meaningful visual representations of the most frequent keywords in the dataset.

4 Conclusion

We propose Structured AI Agents, a structured pipeline for visualization report generation that decomposes the overall task into independent sub-task units, each producing a JSON output with executable code and descriptive templates. Our design improves the stability of the output from LLM and ensures precision through deterministic execution, producing reports where figures and narratives remain consistent and free from hallucinations. Our approach highlights the potential of Structured AI agents to deliver reliable and reproducible data analysis in future applications.

Acknowledgments

Lijie Yao is partially funded by the XJTLU RDF, grant № RDF-24-01-062, and XJTLU TDF, grant № TDF2425-R30-283.

References

[1] P. Isenberg, F. Heimerl, S. Koch, T. Isenberg, P. Xu, C. D. Stolper, M. Sedlmair, J. Chen, T. Möller, and J. Stasko. Vispubdata.org: A Metadata Collection About IEEE Visualization (VIS) Publications. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2199–2206, Sept. 2017. doi: 10.1109/TVCG.2016.2615308 1, 2