Towards Inline Natural Language Authoring for Word-Scale Visualizations

Paige So'Brien*
University of Calgary

Wesley Willett†
University of Calgary

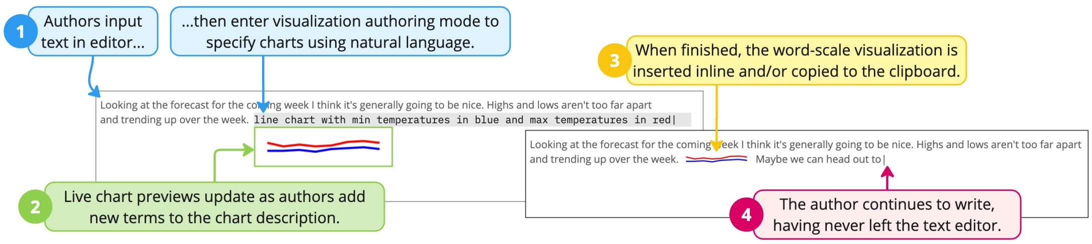


Figure 1: An overview of our natural language authoring workflow for inline word-scale visualization creation.

Abstract

We explore how natural language authoring with large language models (LLMs) can support the inline authoring of word-scale visualizations (WSVs). While word-scale visualizations that live alongside and within document text can support rich integration of data into written narratives and communication, these small visualizations have typically been challenging to author. We explore how modern LLMs—which are able to generate diverse visualization designs based on simple natural language descriptions—might allow authors to specify and insert new visualizations inline as they write text. Drawing on our experiences with an initial prototype built using GPT-4, we highlight the expressive potential of inline natural language visualization authoring and identify opportunities for further research.

Index Terms: word-scale visualization, large language models, visualization authoring

1 Introduction

Word-scale visualizations (WSVs) include a wide range of small data graphics that focus on representing data via compact representations that can be included inline with or alongside written text [4, 5]. Similar kinds of small word-scale graphics have been considered through many lenses, include Edward Tufte's sparklines [10], the micro visualizations outlined by Parnow and Dörk [9], and more general work on data glyphs [2]. These inline graphical representations can be expressive and data-dense, making it possible to communicate detailed trends, relationships, and values succinctly in ways that would be impossible with text alone. However, they are notoriously hard to author and usually require authors to design and implement their own bespoke visualization designs.

Meanwhile, large language models (LLMs) have advanced greatly in their capability to assist people and generate media, including visualizations, based on natural language inputs. Leveraging these new tools, we explore the potential for inline natural-language authoring of word-scale visualizations—demonstrating an example authoring workflow and results from initial prototypes that use GPT-4 to generate word-scale Vega-Lite visualizations. Based on our initial prototyping, we highlight several opportunities for future research related to WSV authoring.

*e-mail: paige.sobrien@alumni.ucalgary.ca
†e-mail: wesley.willett@ucalgary.ca

2 Exploring Inline Natural Language Authoring

In contrast to prior WSV authoring approaches, we propose a natural-language workflow (Figure 1) that unifies the text and visualization creation inside a text editing interface: ① While writing text, an author wishes to include an inline word-scale visualization. At the point in the text where they wish to add a graphic, they enter visualization authoring mode, either using a keyboard shortcut or by highlighting document text they wish to visualize (similar to Masson and colleagues' *delimited selection* approach [7]). This opens a visualization authoring text box at the location of the text cursor. Authors can then type in a natural language description of the word-scale visualization they wish to produce. ② As they enter new terms in their visualization description, live previews of the resulting visualization appear adjacent to the text, allowing them to refine and tune the design. ③ When satisfied, the author can dismiss the visualization authoring mode by pressing the Enter key. The resulting visualization can be inserted directly into the sentence with a keypress (replacing the natural language description), copied to the clipboard for insertion elsewhere, or deleted. ④ The author can then continue writing document text, having never left their text-entry environment. If desired, a visualization can also be dragged and dropped from the live preview into the document text directly.

This interaction model aims to transition the process of authoring WSVs from a complex one involving low-level manual visualization design and specification (typically separate from writing tools) to something more akin to an emoji picker—allowing authors to handle both text and visualization authoring in the same environment using only natural language.

2.1 Design Explorations

To examine the possibilities presented by these kinds of natural language authoring tools, we have conducted an ongoing series of design explorations—focusing both on the interaction workflow for inline authoring and the WSV generating capabilities of current LLMs. We began our exploration with visual prototyping that helped us to evaluate potential chart and placement types within documents of text. Next, we tested various current LLMs capable of generating visualizations from plain language prompts. Lastly, we began development of an application which can use an existing LLM to generate visualizations by users on their device and apply them to their own text.

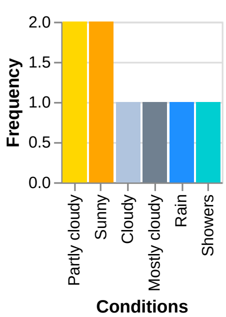
During the initial design of the application, we experimented with the existing natural language models NL4DV [8], and LIDA [1], both of which are trained to create data visualizations from plain language queries. We found that both of these NLMs had difficulty producing visualizations that are suitable for word-scale application, often producing blank results or incorrectly labeled charts which did not adhere to prompt instructions. Their failure to generate word-scale visualizations that are representative of data without labels, titles, and legends is likely due to the lack of training data for these kinds of minimalist visualizations.

Given these mixed results, our subsequent prototyping focused instead on using OpenAI's GPT-4 API to generate simple stripped down Vega-Lite visualizations. Results here were considerably better, likely reflecting the large volume of Vega-Lite specifications available in OpenAI's training data. Vega-Lite's relatively straightforward schema also makes it considerably simpler for LLMs to produce code that alters visualization sizes and removes elements like axes and labels, which are often undesirable at small scales.

2.2 Prototype Implementation

Our current prototype uses the GPT-4 API to translate natural language requests into Vega-Lite specifications, which we then render in-browser using the Vega-Lite renderer. This allows us to generate a wide range of small static visualizations that can be embedded as raster or vector images, as well as interactive graphics that can be embedded inline into HTML documents. Authors first upload a CSV containing a data table which we ingest into the GPT-4 API. Using the prototype interface, authors can then input natural language descriptions of the desired charts.

To simplify the creation of word-scale graphics, we prepend authors' natural language visualization descriptions with a default prompt that specifies the creation of a Vega-Lite schema, constrains the size of the visualization, and eliminates axes, labels, and other visual elements that are unlikely to function at small scales. This allows authors to specify visualizations using succinct natural language prompts. For example, using a simple 8-day weather forecast dataset (Table 1), an author can quickly generate visualizations a simple `‘‘bar chart with min values’’` , or a more complex `‘‘line chart with min temperatures in blue and max temperatures in red’’` . Authors can also take advantage of the latent semantic knowledge of the LLM with natural language specifications like `‘‘bubble chart with counts for each condition, colored to match the weather condition, sorted’’`  which assigns colors commonly associated with each weather condition.



Authors can also craft natural language descriptions that override the default parameters to create more complex visualizations, like the charagraph-sized chart [7] at left (generated using `‘‘200 x 200 px bar chart with counts for each condition, with axes, sorted, colored to match the weather condition’’`). Because charts are rendered using Vega, they can be embedded either as static images or as live objects with support for tooltips and other interactions.

In line with other recent examinations of Vega-Lite generation with ChatGPT (including Li and colleagues' recent evaluation using the NL2VIS corpus [6]), we observed that the LLM's schema generation performance is still somewhat inconsistent, particularly for more complex charts. In practice, we often needed to try multiple prompt variations in order to achieve the desired chart. However, we have found that the relatively simple nature of many word-scale charts, combined with live previews can make it possible to quickly iterate and arrive at a desired visualization.

Day	Weekday	Conditions	Max	Min
1	Mon	Partly cloudy	21	12
2	Tue	Cloudy	17	12
3	Wed	Mostly cloudy	19	13
4	Thu	Rain	17	11
5	Fri	Showers	18	14
6	Sat	Sunny	22	15
7	Sun	Sunny	23	16
8	Mon	Partly cloudy	21	14

Table 1: An example 8-day weather forecast.

3 Opportunities for Inline Visualization Authoring

While our initial exploration highlights some of the potential for inline visualization authoring using LLMs, it also points to a range of opportunities for further research.

Authoring graphically-diverse WSV designs. As showcased by Goffin and coauthors [4], Parnow and Dörk [9], and others, the expressive range and visual diversity of word-scale visualization designs is immense. However, it's unclear how varied a space of word-scale visualizations current LLMs can reproduce, especially since few examples of many of these smaller chart designs exist in the wild. More complex and interactive word-scale visualizations [3] may also pose challenges for natural language authoring.

Given these uncertainties, future research is necessary to explore the expressive range of visualizations that can be produced by current LLMs, as well as their specific limitations. Similarly, new approaches for augmenting LLM-based authoring tools with new bespoke chart designs may be necessary to support visualizations beyond common statistical graphics.

Streamlining authoring for visualization-rich documents. Crafting documents or news stories that include large numbers of word-scale visualizations (as in the results section of an academic paper or a data journalism piece), often entails creating large numbers of nearly identical charts—each showing slightly different slices of the same source data. These kinds of documents can also frequently include titles, legends, and even larger visualizations that share visual encodings and often the same underlying data.

This suggests considerable opportunities for natural language authoring approaches that support shortcuts and macros to let authors quickly re-use the same visualization designs with minor tweaks (a different slice of the data, colors, labels, etc.). Similarly, natural language authoring interfaces could be expanded to support the authoring of titles, legends, summary statistics, interactive links between text and visualizations, and other narrative building blocks.

Making visualizations a first-class text element. To date, the most common uses of WSVs have tended to be in scientific articles, journalism pieces, and web pages—all locales where typesetting and layout systems give authors the ability to place graphics directly inline with text. However, WSVs also have the potential to surface data in many other kinds of communication channels, including instant messaging, email, or social media—allowing writers to enrich their text with visualizations of appointment schedules, weather forecasts, sports statistics, maps, and a multitude of other data sources. Yet while almost all text authoring platforms now allow limited inline integration of text and imagery in the form of emoji, very few of these platforms support more general inline images, let alone interactive inline content.

Driving changes to the text encodings that underpin most applications is unlikely in the near term. However, research prototypes and examples built using web-based text editors (where HTML and CSS make inline image integration comparatively straightforward) might help demonstrate the benefits of richer text-image integration, paving the way for future standards. WSV authoring tools could also consider fallback approaches—including using emoji, links, and other references in the text to connect text with nearby images in cases where inline placements are not possible.

Ultimately, our initial experiments suggest that LLM-based natural language authoring is a viable and promising interaction paradigm for authoring word-scale graphics. We look forward to continuing to examine its breadth of potential applications, encouraging richer and deeper integration of text and data anywhere words are written.

Acknowledgments

The authors wish to thank the members of the Interactions Lab for their support when designing this project.

References

[1] V. Dibia. LIDA: A tool for automatic generation of grammar-agnostic visualizations and infographics using large language models. In D. Bollegala, R. Huang, and A. Ritter, eds., *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pp. 113–126. Association for Computational Linguistics, Toronto, Canada, July 2023. doi: 10.18653/v1/2023.acl-demo.11 1

[2] J. Fuchs, P. Isenberg, A. Bezerianos, and D. Keim. A systematic review of experimental studies on data glyphs. *IEEE transactions on visualization and computer graphics*, 23(7):1863–1879, 2016. 1

[3] P. Goffin, T. Blascheck, P. Isenberg, and W. Willett. Interaction techniques for visual exploration using embedded word-scale visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376842 2

[4] P. Goffin, J. Boy, W. Willett, and P. Isenberg. An exploratory study of word-scale graphics in data-rich text documents. *IEEE Transactions on Visualization and Computer Graphics*, 23(10):2275–2287, 2017. doi: 10.1109/TVCG.2016.2618797 1, 2

[5] P. Goffin, W. Willett, J.-D. Fekete, and P. Isenberg. Exploring the placement and design of word-scale visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 20:2291–2300, 12 2014. doi: 10.1109/TVCG.2014.2346435 1

[6] G. Li, X. Wang, G. Aodeng, S. Zheng, Y. Zhang, C. Ou, S. Wang, and C. H. Liu. Visualization generation with large language models: An evaluation. *arXiv preprint arXiv:2401.11255*, 2024. 2

[7] D. Masson, S. Malacria, G. Casiez, and D. Vogel. Charagraph: Interactive generation of charts for realtime annotation of data-rich paragraphs. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–18, 2023. 1, 2

[8] A. Narechania, A. Srinivasan, and J. Stasko. 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, Feb. 2021. doi: 10.1109/tvcg.2020.3030378 1

[9] J. Parnow and M. Dörk. Micro visualizations: Data-driven typography and graphical text enhancement. In *Proc. IEEE Conf. Inf. Vis. Posters*, pp. 12–13, 2015. 1, 2

[10] E. R. Tufte and E. R. Tufte. *Beautiful evidence*, vol. 1. Graphics Press Cheshire, CT, 2006. 1