Augmenting Human-AI Co-Writing with Interactive Visualization

Md Naimul Hoque
nhoque@umd.edu
University of Maryland, College Park
College Park, MD, USA

Niklas Elmqvist
elm@umd.edu
University of Maryland, College Park
College Park, MD, USA

ABSTRACT

Writing is a fundamental human activity, but today we have the opportunity to leverage Natural Language Processing (NLP) methods to help in this endeavor. Recent tools go beyond mere grammatical error-checking and use Large Language Models (LLMs) to support human-AI co-writing. While existing tools are helpful, many challenges remain: (1) mitigating ownership tensions between humans and AI; (2) enabling human autonomy in the process; (3) creating mechanisms for writers to understand and explore the reasoning of AI; and (4) applying NLP to complex and abstract narrative components (e.g., characterization, events, dialogue). In this paper, we hypothesize that some of these challenges can be resolved by introducing a communication interface between writers and AI. Further, we propose *interactive data visualization*, a prominent method for making sense of complex AI reasoning and revealing hidden patterns from text data, to be that interface. To demonstrate our proposal, we present two case studies where we combine NLP and interactive visualization to support creative writing. The first case study is on mitigating social biases in fiction writing and the second is on the design of dynamic characters and scenes. We conclude by outlining our future work and broader impact.

KEYWORDS

Creative writing, fiction writing, NLP, visualization.

1 BACKGROUND AND MOTIVATION

Modern creative writers use a range of computational tools. The first kind is text editor-like writing software such as Microsoft Word or Google Docs. There are several professional and paid software available to writers. For example, Scrivener [17] allows writers to organize a story in sections, add synopsis and notes to each section, and easily merge or swap between sections. Granthika [9] is a similar sort of paid service where writers can outline a more detailed narrative world, including character descriptions, major events in a timeline, and causal constraints on the events. These tools primarily enhance the organizational capabilities of writers with limited feedback on the actual writing.

More recently, researchers have proposed several emerging writing support tools. A dominant trend is the use of LLMs. These tools can generate text based on a prompt, often helping writers explore alternate narrative worlds and creative angles [2, 3, 8, 16, 21, 26]. Another set of tools uses NLP to extract patterns from the text that are otherwise difficult to notice. These tools typically focus on providing feedback while writers revise their text. For example, Du et al. [5, 6] proposed a tool to help writers in their iterative revising process. Sterman et al. [23] proposed an analytic model that allows writers to interact with the literary style of an article. DramatVis Personae [10] helps writers mitigate nuanced social biases in their writing. Dang et al. [4] integrated automatic text summarization in a text editor to help writers revise analytical essays.

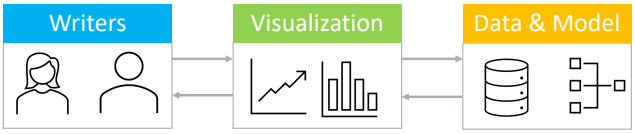

Figure 1: Visualization for co-writing. A workflow for supporting human-AI co-writing with visualization.

While existing tools are useful to writers, there exist several challenges for human-AI co-writing. For example, one critical challenge is enabling human autonomy and reducing ownership tensions in the process [27]. Another relevant problem is the lack of trust among writers and AI as LLMs and other NLP models are typically black boxes and are difficult to interpret and explore intuitively [21]. Other challenges include identifying plagiarism and bias in the AI-generated text and applying NLP to more abstract and complex narrative components such as characterization, narrative structure, and events [8].

In this work, we hypothesize that some of these challenges arise from a lack of communication interface between writers and AI. We further propose *interactive data visualization* to be that interface (Figure 1). There are abundant works on using visualization for text analytics [14, 24], human-AI collaboration [1, 11], explainable AI (XAI) [12, 25], and digital humanities [15, 18, 20]. However, it is still unknown how to translate and integrate these techniques to human-AI co-writing.

Our work aims to fill that void. In the next sections, we present two case studies, demonstrating how NLP and visualization can be combined to support creative writing.

2 CASE STUDY 1: MITIGATING SOCIAL BIASES IN CREATIVE WRITING

Implicit biases and stereotypes are often pervasive in different forms of creative writing such as novels, screenplays, and children's books [7, 13]. We interviewed 9 writers to understand how they mitigate unwanted biases in their writing. Writers stated that they actively look for such biases in their stories. However, these biases are often implicit and unconscious and difficult to wheedle out even for the best and most self-reflective of authors. The process is even more challenging for long and complicated stories where many characters take intersectional identities.

Based on the findings of the interviews, we designed DramatVis Personae (DVP) [10], a web-based visual analytics system to help

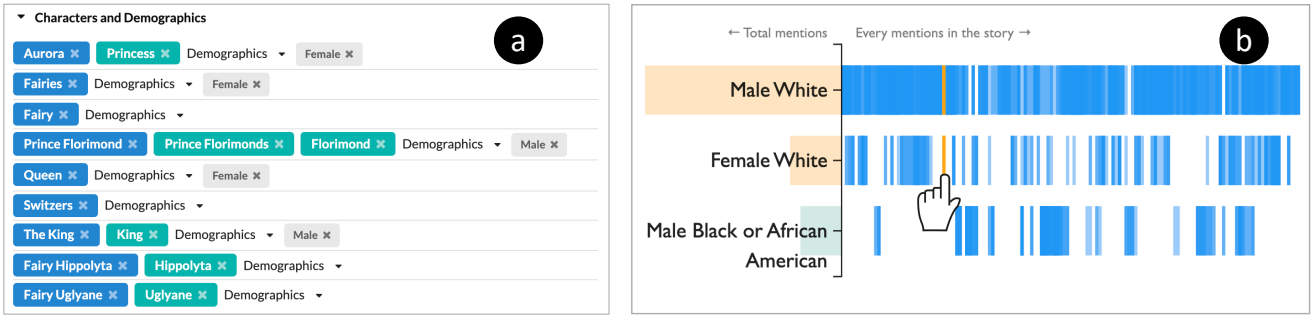

Figure 2: DramatVis Personae. (a) Characters and Demographics panel. A writer can assign social identities to their characters using this panel. In this instance, we are seeing the demographics panel for the children's book, *Sleeping Beauty*. (b) An example timeline showing the presence of three intersectional identities (*Male White, Male Black or African American,* and *Female White*) in the Movie *The Amazing Spiderman 2* (2014). Male White characters are present throughout the storyline. Further, there appears to be a lack of interactions between Male African-American and Female White characters except for a few aberrant ones. However, both groups have interactions with Male White characters. The orange bars show one such interaction.

writers identify stereotypes in creative writing. DVP is designed to integrate smoothly with the writer's own creative process, allowing them to analyze existing literature for research, upload their written content as it becomes available, or even write in the tool itself, and then have its text analytics and visualizations updated in real-time. DVP automatically detects characters in a text and allows writers to assign social identities such as age, ethnicity, and gender to the characters (Figure 2a). The DVP dashboard uses this continually growing dataset to visualize the presence of characters and social identities over time (Figure 2b).

3 CASE STUDY 2: DESIGNING DYNAMIC CHARACTERS AND SCENES

Characters are often central to narrative fiction. Many writers base their stories on rich and complex characters that drive the plot, almost as if they have a life of their own. As a case in point, witness Walter White slowly transforming from milquetoast chemistry teacher to insidious drug kingpin in AMC's *Breaking Bad* (2008).

Designing such engaging and dynamic characters is challenging for writers as they need to believably show the evolution of the character throughout the story. While the process for developing characters may vary from writer to writer, we believe an analytic tool at the time of editing or revising a draft can help writers to identify discrepancies between intended and written text, decide what edits to make, and how to make those desired edits. For example, writers often try to portray an emotional journey for the characters (i.e., character arcs). Sentiment analysis can capture the character arcs and help writers revise their draft to achieve the intended arcs. Another common technique for creating dynamic scenes is to let characters with opposite emotions interact with each other (i.e., Bakhtin's Polyphonic Theory). Sentiment and emotion analysis can again help writers design such dynamic scenes. We aim to integrate these models into a writing support tool using interactive visualization.

4 FUTURE RESEARCH DIRECTIONS

We envision several future directions for our research. We outline these directions below.

Empirical Studies. Our current works show promises for visualization as an interface between AI and humans. However, we have not properly evaluated the utility of visualization for mitigating the challenges listed in Section1. One promising future direction is conducting empirical research to evaluate the potential of visualization in those scenarios. For example, we could integrate an XAI component like SHAP visualization [19] in a writing tool with a generative model. Writers can understand the reason behind the generated text from a prompt by investigating the visualization. We can study how such XAI component could help improving human autonomy and trust between writers and AI.

Designing Novel Visualizations. Another promising approach is designing novel visualization to support complex creative writing tasks. Our first case study on mitigating social biases is one example of such research. Interactive visualization allowed us to integrate computational support for this sensitive and complex task. The tool empowered writers with access to advanced NLP methods without the need for algorithmic expertise. We encourage visualization researchers to invest efforts to design novel visual representations for other narrative tasks such as designing non-linear narrative structures and managing causal constraints between events, time, and settings.

Opportunities for NLP as an Analytic Partner. We believe our research has the potential to rekindle research on using NLP as an *analytic* assistant. This is in contrast with recent work on creative writing, which tends to focus on *generative* models [3, 16, 26]. However, we believe NLP as an assistant—when applied to reveal hidden patterns from stories rather than creating stories—still has powerful applications to offer. We hope our work will motivate future research to revisit how NLP can help writers analyze content. We believe that bridging the two directions—NLP both for story

generation and for analytic assistance—is the right way to cement NLP's importance in creative writing.

Scaling to Multiple Documents. The presented case studies support the analysis of a single document. We believe extending the analysis to multiple documents will lead to several novel applications. First, writers could perform version comparisons so that they can easily see the impact of the edits. Support for such iterative revision has previously been shown to be useful [5]. Another potential direction is supporting the analysis of a text corpus. For example, we can answer questions such as *what is the representation of females in stories written during 1900?* Such analysis could be useful to literary scholars, who use language to understand society and the literature borne from it [22].

REFERENCES

[1] Alex Bäuerle, Ángel Alexander Cabrera, Fred Hohman, Megan Maher, David Koski, Xavier Suau, Titus Barik, and Dominik Moritz. 2022. Symphony: Composing Interactive Interfaces for Machine Learning. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, 210:1–210:14. https://doi.org/10.1145/3491102.3502102

[2] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems* 33 (2020), 1877–1901. https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

[3] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching Stories with Generative Pretrained Language Models. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 209:1–209:19. https://doi.org/10.1145/3491102.3501819

[4] Hai Dang, Karim Benharrak, Florian Lehmann, and Daniel Buschek. 2022. Beyond Text Generation: Supporting Writers with Continuous Automatic Text Summaries. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, USA, 98:1–98:13. https://doi.org/10.1145/3526113.3545672

[5] Wanyu Du, Zae Myung Kim, Vipul Raheja, Dhruv Kumar, and Dongyeop Kang. 2022. Read, Revise, Repeat: A System Demonstration for Human-in-the-loop Iterative Text Revision. In *Proceedings of the Workshop on Intelligent and Interactive Writing Assistants*. Association for Computational Linguistics, Stroudsburg, PA, USA, 96–108. https://doi.org/10.18653/v1/2022.in2writing-1.14

[6] Wanyu Du, Vipul Raheja, Dhruv Kumar, Zae Myung Kim, Melissa Lopez, and Dongyeop Kang. 2022. Understanding Iterative Revision from Human-Written Text. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Stroudsburg, PA, USA, 3573–3590. https://doi.org/10.18653/v1/2022.acl-long.250

[7] Ethan Fast, Tina Vachovsky, and Michael S. Bernstein. 2016. Shirtless and dangerous: Quantifying linguistic signals of gender bias in an online fiction writing community. In *Proceedings of the International AAAI Conference on Web and Social Media*. Association for Computational Linguistics, Stroudsburg, PA, USA, 112–120.

[8] Katy Ilonka Gero, Vivian Liu, and Lydia B. Chilton. 2022. Sparks: Inspiration for Science Writing using Language Models. In *Proceedings of the ACM Conference on Designing Interactive Systems*. ACM, New York, NY, USA, 1002–1019. https://doi.org/10.1145/3532106.3533533

[9] Granthika. 2021. Granthika Writing Tool. https://granthika.co Accessed: 2022-02-04.

[10] Md Naimul Hoque, Bhavya Ghai, and Niklas Elmqvist. 2022. DramatVis Personae: Visual Text Analytics for Identifying Social Biases in Creative Writing. In *Proceedings of the ACM Conference on Designing Interactive Systems*. ACM, New York, NY, USA, 1260–1276. https://doi.org/10.1145/3532106.3533526

[11] Md. Naimul Hoque, Wenbin He, Arvind Kumar Shekar, Liang Gou, and Liu Ren. 2023. Visual Concept Programming: A Visual Analytics Approach to Injecting Human Intelligence at Scale. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2023), 74–83. https://doi.org/10.1109/TVCG.2022.3209466

[12] Md. Naimul Hoque and Klaus Mueller. 2022. Outcome-Explorer: A Causality Guided Interactive Visual Interface for Interpretable Algorithmic Decision Making. *IEEE Transaction on Visualization and Computer Graphics* 28, 12 (2022), 4728–4740. https://doi.org/10.1109/TVCG.2021.3102051

[13] Alexander Miserlis Hoyle, Lawrence Wolf-Sonkin, Hanna Wallach, Isabelle Augenstein, and Ryan Cotterell. 2019. Unsupervised Discovery of Gendered Language through Latent-Variable Modeling. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Stroudsburg, PA, USA, 1706–1716. https://doi.org/10.18653/v1/P19-1167

[14] Stefan Jänicke, Greta Franzini, Muhammad Faisal Cheema, and Gerik Scheuermann. 2015. On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In *State of the Art Reports of the Eurographics Conference on Visualization*. Eurographics Association, Geneva, Switzerland, 83–103. https://doi.org/10.2312/eurovisstar.20151113

[15] Nam Wook Kim, Benjamin Bach, Hyejin Im, Sasha Schriber, Markus H. Gross, and Hanspeter Pfister. 2018. Visualizing Nonlinear Narratives with Story Curves. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 595–604. https://doi.org/10.1109/TVCG.2017.2744118

[16] Mina Lee, Percy Liang, and Qian Yang. 2022. CoAuthor: Designing a Human-AI collaborative writing dataset for exploring language model capabilities. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 388:1–388:19. https://doi.org/10.1145/3491102.3502030

[17] Literature and Latte. 2021. Scrivener. https://www.literatureandlatte.com/scrivener/overview/ Accessed: 2022-02-04.

[18] Shixia Liu, Yingcai Wu, Enxun Wei, Mengchen Liu, and Yang Liu. 2013. StoryFlow: Tracking the Evolution of Stories. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2436–2445. https://doi.org/10.1109/TVCG.2013.196

[19] Scott M. Lundberg, Bala Nair, Monica S. Vavilala, Mayumi Horibe, Michael J. Eisses, Trevor Adams, David E. Liston, Daniel King-Wai Low, Shu-Fang Newman, Jerry Kim, et al. 2018. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature Biomedical Engineering* 2, 10 (2018), 749–760. https://doi.org/10.1038/s41551-018-0304-0

[20] Nina McCurdy, Julie Lein, Katherine Coles, and Miriah D. Meyer. 2016. Poemage: Visualizing the Sonic Topology of a Poem. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 439–448. https://doi.org/10.1109/TVCG.2015.2467811

[21] Piotr Mirowski, Kory W. Mathewson, Jaylen Pittman, and Richard Evans. 2022. Co-Writing Screenplays and Theatre Scripts with Language Models: An Evaluation by Industry Professionals. *CoRR* abs/2209.14958 (2022), 102 pages. https://doi.org/10.48550/arXiv.2209.14958 arXiv:2209.14958

[22] Stéfan Sinclair and Geoffrey Rockwell. 2012. Voyant Tools. https://voyant-tools.org/ Accessed: 2022-02-04.

[23] Sarah Sterman, Evey Huang, Vivian Liu, and Eric Paulos. 2020. Interacting with Literary Style through Computational Tools. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376730

[24] Fernanda B. Viégas, Martin Wattenberg, and Jonathan Feinberg. 2009. Participatory Visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 1137–1144. https://doi.org/10.1109/TVCG.2009.171

[25] Zijie J. Wang, Robert Turko, Omar Shaikh, Haekyu Park, Nilaksh Das, Fred Hohman, Minsuk Kahng, and Duen Horng (Polo) Chau. 2021. CNN Explainer: Learning Convolutional Neural Networks with Interactive Visualization. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1396–1406. https://doi.org/10.1109/TVCG.2020.3030418

[26] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: Story Writing With Large Language Models. In *Proceedings of the ACM Conference on Intelligent User Interfaces*. ACM, New York, NY, USA, 841–852. https://doi.org/10.1145/3490099.3511105

[27] Ying Yuan. 2022. Analysis of fictional characters in the context of artificial intelligence and big data: taking *A Dream of Red Mansions* for example. *Digit. Scholarsh. Humanit.* 37, 1 (2022), 289–297. https://doi.org/10.1093/llc/fqab064