My research focuses on designing, building, and validating AI-infused supertools using interactive visualization. AI-infused supertools are applications that amplify, augment, empower, and enhance human performance, by combining user experiences with **AI support services** [1]. In the past decades, we have made progress towards this vision of supertools. Generative AIs such as GPT-4 and DALL-E3 can generate human-like text, converse with humans, and produce futuristic images that amplify human creativity and imagination. However, at the same time, we see concerns about generative AIs threatening to overtake human **agency** and **ownership** [13]. For example, the Writers Guild of America (WGA) and the Screen Actors Guild - American Federation of Television and Radio Artists (SAG-AFTRA), were recently on strike, demanding clauses on their contract that they will not be replaced by AI.

In my research, I argue that to retain human agency and ownership in the design of *supertools* effectively, we need a **communication medium** between humans and AI that humans can use to steer AI toward their needs while having clear control over the final artifact (e.g., novel, code, etc.). I further propose ***interactive visualization*** to be that medium. To validate this premise, during my dissertation, I designed three AI-assisted writing tools and evaluated them with professional and novice creative writers, to study how interactive visualization and AI can amplify creative writing while retaining writers' agency and ownership. AI-assisted writing is a controversial area with high tensions between AI and humans, which makes the domain a suitable probe for validating my research agenda. My research summarizes lessons learned from the design process of the tools and proposes future works to cement *Designing AI-Infused Supertools with Interactive Visualization* as an emerging research area for the HCI and visualization community.

Hallmark: Supporting Provenance and Transparency in LLM-based Co-writing

A critical concern for writers who want to use LLMs is properly attributing contributions from the AI. This gives rise to "ownership tension" among writers and hampers their agency and control over the process. Identifying contributions from AI is also important for writers who need to conform to AI-assisted writing policies such as the one from the U.S. Copyright Office [2]. Authors currently have few mechanisms to track their accountability with regard to these rules and policies.

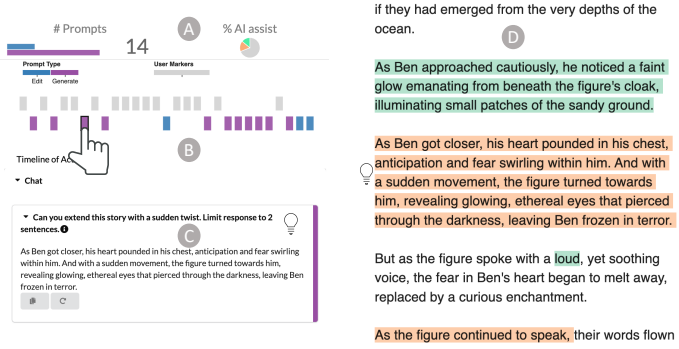

Figure 1: Visualization and interaction in Hallmark. a) Summary statistics: number of prompts and percentage of assistance from AI. b) Blue tiles indicate prompts seeking editorial assistance while purple tiles indicate prompts asking to generate new text. Grey bars show the user's writing behavior (e.g., writing a new sentence). Hovering over a colored tile will show the respective prompt (c) and text highlighted in the text editor (d). Text with green and orange colors indicates text influenced and written by

In this work, I argued that capturing interactions between AI and writers as the document evolves (i.e., provenance information) and supporting interactive exploration of that provenance information may improve the writer's agency, control, and ownership of the final artifact (e.g., short stories, novels, poems). Provenance can also help writers conform to AI-assisted writing policies and be transparent to publishers and readers. To explore this design space, I first reviewed existing guidelines and policy documents on the use of LLMs from several professional, educational, and academic organizations. This review informed me about the types of information that writers should be aware of in AI-assisted writing. I then developed **Hallmark** [3], a web-based design probe that integrates an authoring interface with LLM support that stores and visualizes a writer's interaction with an

LLM (Figure 1). The system facilitates writers to self-reflect on their use of AI by clearly highlighting AI contributions such as text written and influenced by AI and types of prompting (editing vs generating).

To validate Hallmark, I engaged 13 creative writers to use it to write a short story during remote evaluation sessions and collected their feedback and resulting stories. The findings suggest Hallmark encouraged writers to actively evaluate AI assistance from the onset of the writing process. As a result, it instilled a **sense of control** in the writer's mind and **improved ownership** of the final artifact. The findings also suggest Hallmark would help writers **conform to AI-assisted writing policies** without the need for manually generating disclosures. Writers were therefore confident that Hallmark will help them become **more transparent** and that the tool is an effective medium to communicate the use of AI in writing to external parties (i.e., publishers).

DramatVis-Personae: Visual Text Analytics for Identifying Social Biases in Writing

While LLMs are useful to writers for generating new text and narrative angles, it is still unknown how to best apply them to more complex and abstract narrative components such as characterization, dialogue, and narrative structures, and how that can help writers. In this work, I investigated how AI and visualization can help writers identify and mitigate implicit biases and stereotypes. We interviewed writers to understand the types of biases they were concerned about, how they mitigate those, and how AI can help them in this regard. The interviews suggested that despite a writer's best interest, tracking and managing implicit biases such as **a lack of agency, supporting or submissive roles, or harmful language** for characters representing marginalized groups is challenging as the story becomes longer and more complicated. What intrigued me even more about this research was the challenges associated with the problem. First, the task was sensitive, and we needed to ensure that the limitations of the AI (biases inherent to the AI models) did not further propagate to the story. Second, a writer might want to write about discrimination and biases intentionally and the tool should not prevent the writer from that. Finally, the system needed to be inclusive, supporting social identities beyond dominant identities in society (e.g., supporting LGTBQ identities).

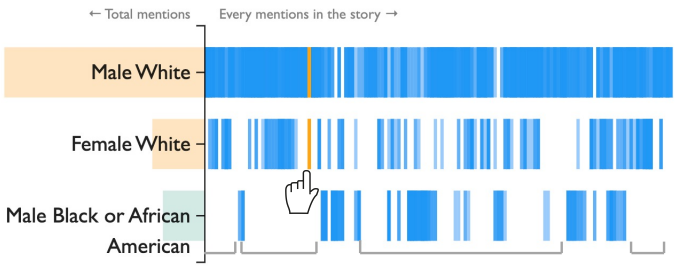

Figure 2*:* An example timeline showing the presence of three intersectional identities (Male White, Male Black or African American, and Female White) in the Movie "The Amazing Spiderman 2" (2014). Male White characters are present throughout the storyline.

We addressed all of the above challenges by using interactive visualization as a communication medium between the writers and AI in a system called **DramatVis Personae (DVP)** [4], a writing support tool for writers to investigate the representation of social identities (e.g., female white) in fiction. DVP allows writers to **assign social identities** of any kind to their characters. DVP can then measure and visualize the presence of social identities (Figure 2) and adjectives and actions used to describe different social identities.

Portrayal: Capturing and Visualizing Character Arcs

Portrayal [5] is an extension of DVP where we extend the supported character traits from only social identities to the character's sentiments, emotions, actions, and character arcs (Figure 3). I achieved a significant milestone in this project: conducting a longitudinal study with professional writers. This was essential to deeply understand the design of not just Portrayal, but also DVP and Hallmark. Twelve Writers used Portrayal to analyze one of their unpublished drafts independently for a week. In a debriefing meeting at the end of the week, we asked them to describe any new insights or thoughts arising from using Portrayal.

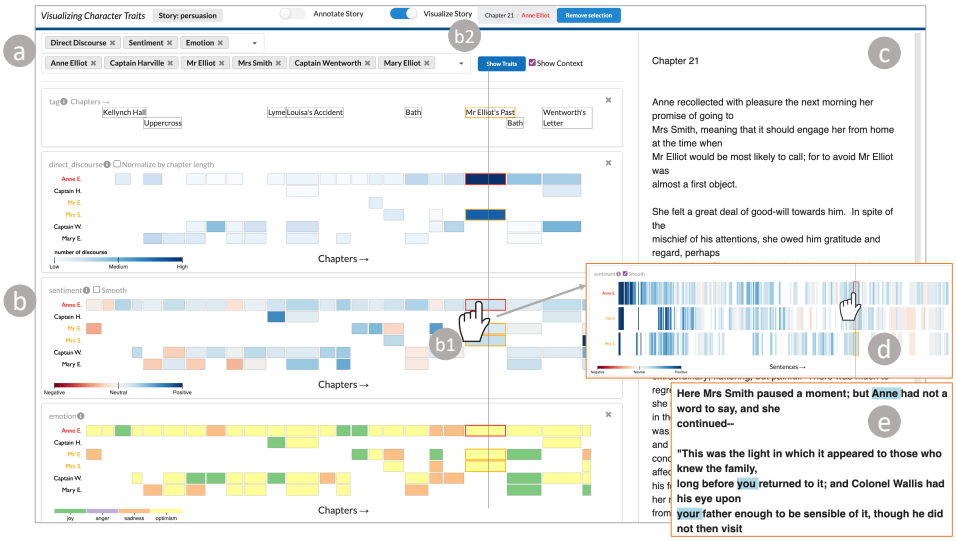

Figure 3: Overview of **Portrayal** using the novel *Persuasion* by Jane Austen. The central panel (b) shows the trait indicators for each character across the chapters. The panel shows indicators (direct discourse, sentiment, and emotion in this case) in three separate views, aligned vertically. Each row in these views represents a character while each column represents a chapter. The views are synchronized. For example, on hovering over a cell (b1), the system highlights the character of interest in red (Anne Elliot) and other co-occurring characters of that chapter in orange, and the text editor moves to the corresponding chapter (c). When a user clicks on a cell (b1), the views update to a sentence-wise visualization (d) where each cell represents a sentence of the selected chapter. On hovering over a tile in this view (d), the system highlights the selected sentence and the mentions of the character in the text editor (e).

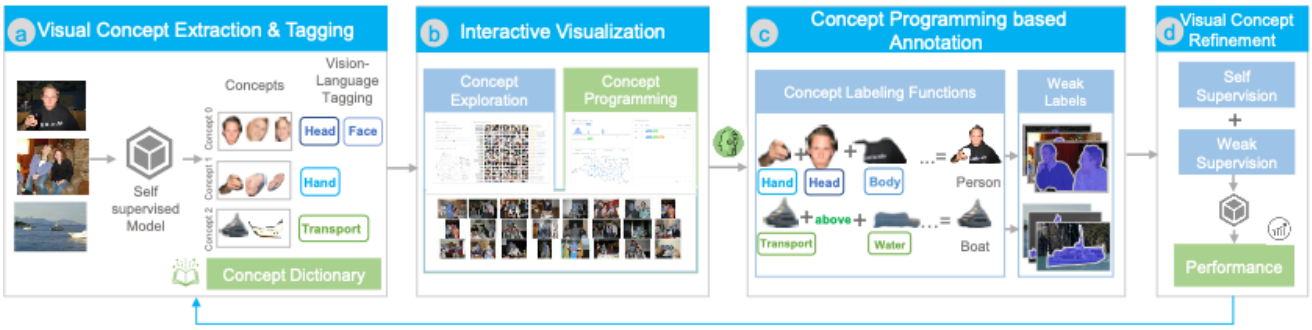

Figure 4: The framework of visual concept programming. a) Visual concept extraction and tagging component encode large-scale images into a dictionary of visual concepts through a self-supervised representation learning model. Each visual concept is associated with a group of image segments that share similar semantic meanings such as human head, hand, etc. The visual concepts are then tagged via a zero-shot vision-language model, CLIP; b) Interactive visualizations are provided to help users explore and understand the concepts, and also enable users to program concepts by supporting to compose, verify, and refine labeling functions with visual concepts; c) User generated labeling functions are applied to large-scale images and generate weak labels for image segments; d) Finally, weak labels are used to retrain the representation learning model to improve image segments and concepts.

The findings suggest that Portrayal helped writers (1) create **dynamic characters** that go through many emotional changes, (2) create **dynamic scenes** where characters with opposite sentiments and emotions interact, (3) evaluate character arcs, and (4) find unintentional prioritizing of social identities.

Future Work

My future work will focus on designing new visual representations for foundation models, novel AI-infused supertools, and a critical approach for evaluating the supertools. Here, I outline my vision.

Designing Visualization for Foundation Models. A critical step for using visualization as a communication medium between humans and AI is to devise visualization techniques capable of adapting to the new wave of foundation models. During my dissertation, I developed a series of visualization techniques [6-9] towards that

goal. For example, **Visual Concept Programming** [7] summarizes a large-scale image dataset by extracting visual concepts from the dataset by using a self-supervised representation learning model (Figure 4). The user can then interactively compose labeling functions (e.g., Head + Hand + Body = Person) and improve the learned representation. In my experiments, I showed that the learned representations perform better in downstream tasks like semantic segmentation than state-of-the-art models. In the future, I want to continue my research on developing **scalable and explainable visualization** techniques to help researchers and practitioners understand the internal mechanism of AI models. This thread will likely influence the design of visualizations in future supertools. My prior research in visual analytics and explainable AI (XAI) will work as a base for this thread in the future. I will seek external funding to support this research.

Designing Supertools for Diverse Domains. To establish designing supertools with visualization as a prominent HCI approach, we need to explore more domains and areas. My current research explores AI-assisted writing, a critical domain that provides a strong foundation to explore other critical domains. My future work will focus on areas where agency between humans and AI is often blurred and difficult to manage. One such domain I am particularly interested in is computational journalism. I recently explored how LLM-based chatbots can improve audience engagement and answer questions from the audience [10]. I conducted an online experiment to understand the types of questions readers want to ask the authors of an article and how the questions change when they ask the questions to a chatbot instead of the authors. Beyond computational journalism, I will explore domains with high-stakes decisions (e.g., health and medical support and financial decisions), and critical societal resource allocation scenarios such as police allocation [12]. I am excited to collaborate and write grant applications with researchers from these domains in my future workplace.

Accessible Data Visualization. The concept of Supertools with visualization rests on the assumption that users can see the visualization. What happens when a user cannot see the visualization, like a blind user? How can we design supertools if a large portion of the population cannot use them? Motivated by that, I investigated the design of accessible data visualization. I proposed **Susurrus** [11], a sonification (non-speech audio) method for translating common data visualization into a blend of natural sounds. Susurrus leverages people's familiarity with sounds drawn from nature, such as birds singing in a forest, and their ability to listen to these sounds in parallel, to represent a data visualization. This work was supported by the National Science Foundation (NSF). My future work will continue this research thread.

Evaluating Ethical Dimension of Supertools. It could well be argued that the central argument of my research—that modern AI technologies such as LLMs are here to stay, and that we should just learn how to best leverage them—is a technopositive, naive, and perhaps even actively harmful approach to the use of AI in human creativity, and that generative AI should be seen as dangerous technology that should be regulated or even banned. However, I would argue that this is true of virtually any technology. For example, photography was widely hailed as the end of painting but instead freed painters from the curse of realism [13]. However, I do believe that to harness these technologies as supertools, we need to thoroughly evaluate the supertools. Concepts relevant to supertools (e.g., agency, transparency, ownership) are mostly abstract concepts and are difficult to operationalize in research for evaluation [14]. My future work will focus on devising methods and experimental designs to evaluate supertools.

References

[1] Ben Shneiderman. 2022. Human-Centered AI. Oxford University Press, Oxford, United Kingdom.

[2] U.S. Copyright Office. 2023. Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence. https://copyright.gov/ai/ai_policy_guidance.pdf

[3] Md Naimul Hoque, Tasfia Mashiat, Bhavya Ghai, Cecilia D. Shelton, Fanny Chevalier, Kari Kraus, Niklas Elmqvist. The HaLLMark Effect: Supporting Provenance and Transparent Use of Large Language Models in Writing with Interactive Visualization. 2023. https://arxiv.org/abs/2311.13057 (Under review)

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

[5] Md Naimul Hoque, Bhavya Ghai, Kari Kraus, Niklas Elmqvist. Portrayal: Leveraging NLP and Visualization for Analyzing Fictional Characters. *Proceedings of the ACM Conference on Designing Interactive Systems (DIS),* ACM, New York, NY, USA, 2023. https://doi.org/10.1145/3563657.3596000

[6] Md Naimul Hoque, Niklas Elmqvist. Dataopsy: Scalable and Fluid Visual Exploration using Aggregate Query Sculpting. *IEEE Transaction on Visualization and Computer Graphics (TVCG)*, 2023.

[7] Md Naimul Hoque, Wenbin He, Shekar Arvind Kumar, Liang Gou, Liu Ren. Visual Concept Programming: A Visual Analytics Approach to Injecting Human Intelligence at Scale. *IEEE Transaction on Visualization and Computer Graphics (TVCG)*, 2022. https://doi.org/10.1109/TVCG.2022.3209466

[8] Bhavya Ghai, Md Naimul Hoque, Klaus Mueller. WordBias: An Interactive Visual Tool for Exploring Intersectional Social Biases Encoded in Word Embeddings. *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems (CHI)*, ACM, New York, NY, USA, 2021. https://doi.org/10.1145/3411763.3451587

[9] Md Naimul Hoque, Klaus Mueller. Outcome-Explorer: A Causality Guided Interactive Visual Interface for Interpretable Algorithmic Decision Making. *IEEE Transaction on Visualization and Computer Graphics (TVCG)*, 2021. https://doi.org/10.1109/TVCG.2021.3102051

[10] Md Naimul Hoque, Ayman A Mahfuz, Mayukha Kindi, Naeemul Hassan. Towards Designing a Question-Answering Chatbot for Online News: Understanding Questions and Perspectives. 2023. (Under review)

[11] Md Naimul Hoque, Md Ehtesham-Ul-Haque, Niklas Elmqvist, Syed Masum Billah. Accessible Data Representation with Natural Sounds. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI),* ACM, New York, NY, USA, 2023. https://doi.org/10.1145/3544548.3581087

[12] Danielle Ensign, Sorelle A. Friedler, Scott Neville, Carlos Scheidegger, Suresh Venkatasubramanian. Runaway feedback loops in predictive policing. In *Conference on fairness, accountability and transparency*, PMLR 81:160-171. 2018.

[13] Ziv Epstein, Aaron Hertzmann, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R. Frank, Matthew Groh, Laura Herman, Neil Leach, Robert Mahari, Alex "Sandy" Pentland, Olga Russakovsky, Hope Schroeder, and Amy Smith and. Art and the science of generative AI. *Science* 380, 6650, 1110–111. 2023. https://doi.org/10.1126/science.adh4451

[14] Mina Lee, Megha Srivastava, Amelia Hardy, John Tickstun, Esin Durmus, Ashwin Paranjape, Ines Gerard-Ursin, Xiang Lisa Li, Faisal Ladhak, Frieda Rong, Rose E. Wang, Minae Kwon, Joon Sung Park, Hancheng Cao, Tony Lee, Rishi Bommasani, Michael Bernstein, Percy Liang. Evaluating Human-Language Model Interaction. *Transactions on Machine Learning Research (TMLR),* 2023.