My research focuses on designing, building, and validating human-centered AI (HCAI) tools using interactive visualization. In the past decades, we have seen an abundance of AI-infused tools that amplify and empower human creativity and cognition. This advancement is mainly fueled by generative AIs such as GPT-4 and DALL-E3 that can generate human-like text, produce futuristic images, and support many other complex tasks with incredible efficiency. However, at the same time, we have seen concerns about generative AIs threatening to overtake human agency and ownership [1]. 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 design AI tools that are *human-centered*, we need a **communication medium** between humans and AI that people can use to steer AI towards their needs while having clear control over the resulting artifact (e.g., novel, code, or other outputs). I further propose **interactive visualization** to be that medium and call this overall approach *visualization-enabled HCAI tools*. To validate this premise, during my dissertation, I first developed an operational definition of HCAI tools based on their basic requirements, capabilities, and human concerns. I then showed how important design characteristics of visualization can address these concerns. Finally, I designed several visualization-enabled HCAI tools, including for creative writing, video production, and interactive machine learning. My research summarizes lessons learned from the design process of the tools and proposes future works to cement *visualization-enabled HCAI tools* as an emerging research area for the HCI, visualization, and AI community.

## *Definition: Human-Centered AI (HCAI) Tools*

There is no generally accepted definition for "HCAI tools" nor do we have a clear understanding of the requirements and components of such tools. Thus, instead of a canonical definition, I provide an operational definition of HCAI tools based on their basic requirements, capabilities, and human concerns (Figure 1). I start by enumerating some basic requirements of an HCAI tool such as that it should be an interactive software, involve human users, and use one or more AI models. I then outline the capabilities of Shneiderman's "supertools" [3], the closest notion relevant to HCAI tools. These capabilities include: 1) amplify- magnifying or strengthening existing abilities; 2) augment- adding new abilities not previously available; 3) empower- making tasks possible that were previously impossible; and 4) enhance- improving the quality of existing abilities or artifacts. Finally, to qualify as being "human-centered", HCAI tools must also incorporate and address key human concerns. There are seven key concerns in this list: 1) fairness; 2) transparency; 3) explainability; 4) understandability; 5) accountability; 6) provenance; and 7) privacy. To be clear, not all tools will address all human concerns, and not all human concerns can be listed exhaustively. Additionally, some dimensions of human concerns can be interdependent. For instance, an HCAI tool capable of explaining its decisions or reasoning process may satisfy the transparency needs of users to some extent.

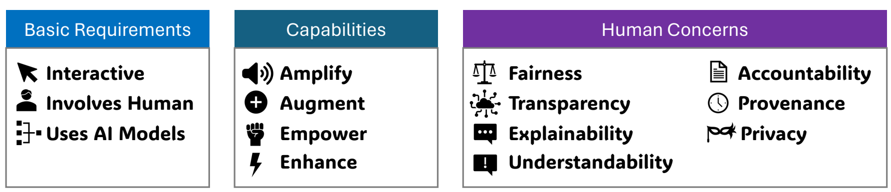


**Figure 1:** Requirements and components of HCAI tools. They should be interactive, involve human users, and use one or more AI models. They need to incorporate human concerns with the capabilities to be useful to the human users.

## *Interactive Visualization for HCAI Tools*

I have developed specific design characteristics (DC1-DC5) of Interactive visualizations that make them a key enabling technology for designing HCAI tools [2]. These characteristics address requirements outlined by

| HCAI Tool | Domain | Venue | Year | 🔊 | ⊕ | ✊ | ⚡ | ⚖ | ✺ | 💬 | ❗ | 📄 | 🕓 | 🎭 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **HaLLMark [9]** | AI-assisted writing | CHI | 2024 | ✓ | ✓ | ✓ | ✓ | ⊖ | ✓ | ⊖ | ⊖ | ✓ | ✓ | ⊖ |
| Portrayal [10] | AI-assisted writing | DIS | 2023 | ✓ | ✓ | ✓ | ✓ | ⊖ | ⊖ | ⊖ | ✓ | ⊖ | ⊖ | ⊖ |
| DramatVis-Personae [11] | AI-assisted writing | DIS | 2022 | ✓ | ✓ | ✓ | ✓ | ✓ | ⊖ | ⊖ | ✓ | ⊖ | ⊖ | ⊖ |
| Dataopsy [12] | Interactive ML | VIS | 2023 | ✓ | ⊖ | ⊖ | ✓ | ⊖ | ⊖ | ⊖ | ⊖ | ⊖ | ⊖ | ⊖ |
| VCP [13] | Interactive ML | VIS | 2022 | ✓ | ⊖ | ✓ | ✓ | ⊖ | ⊖ | ✓ | ✓ | ⊖ | ⊖ | ⊖ |
| **Outcome-Explorer [14]** | Interactive ML | VIS | 2021 | ✓ | ⊖ | ✓ | ✓ | ⊖ | ✓ | ✓ | ✓ | ✓ | ✓ | ⊖ |
| Screen-balancer [19] | Video Production | CSCW | 2020 | ✓ | ⊖ | ⊖ | ✓ | ✓ | ⊖ | ⊖ | ✓ | ⊖ | ✓ | ⊖ |

**Table 1:** Examples of visualization-enabled HCAI tools developed during my PhD. These tools are classified according to four distinct capabilities—amplify (🔊), augment (⊕), empower (✊), and enhance (⚡)—as well as seven human concerns: fairness (⚖), transparency (✺), explainability (💬), understandability (❗), accountability (📄), provenance (🕓), and privacy (🎭). Tools highlighted in bold are discussed in detail below.

several existing guidelines on HCAI and relevant technology, including Eric Horvitz's guidelines for mixed-initiative interfaces [4], Microsoft and Google's guidelines for human-AI interaction designs [5, 6], Heer's discussion on shared representations [7], and Shneiderman's Prometheus Principles [8]. Below, I briefly introduce the design characteristics.

**DC1. Open-ended and data-driven:** A direct strength of data visualization is that it—unlike confirmatory tools—supports open-ended exploratory data analysis informed by the data. This supports Horvitz's guideline on considering *uncertainty about user goals* [4].

**DC2. Facilitates user-computer conversations:** Modern data visualization tools are fundamentally interactive, thus supporting seamless exchanges between the AI and the user. This interactive nature undergirds Shneiderman's *consistent interfaces to allow users to form, express, and revise intent* as well as *rapid, incremental, and reversible actions* [8].

**DC3. Externalizes data:** Visualizations serve as external representations of data, thereby offloading memory, facilitating re-representation, and simplifying computation. As per Microsoft's guideline, this can reduce cognitive load by *showing contextually relevant information* and *remembering recent interactions* [5].

**DC4. Shared data representation:** Visual representations can serve as a representation of the data common to both user and AI models. A *shared representation* is central to Heer's argument for enabling both user and AI control of a common task [7]. These representations can also be used to, as Google PAIR puts it, *determine how to show model confidence* as well as *explain for understanding* [6].

**DC5. Shared task representation:** The interactions enabled by the visualization capture the user's potential actions. Similar to DC4, this scaffold's Heer's *shared representation* of the user's actions on the data [7].

Building on these design characteristics, I have developed several visualization-enabled HCAI tools during my PhD (Table 1). These tools include three creative writing tools and several interactive machine learning tools. I also explored several other novel HCAI domains, including video production [19], news audience engagement [15], and accessible data visualization [16]. As a demonstration of my approach to designing HCAI tools, in the following, I discuss two tools I have developed: 1) HaLLMark [9], a creative writing tool; and 2) Outcome-Explorer [14], an interactive tool for AI decision-making.

### *HaLLMark: An HCAI tool for Supporting Transparency in LLM-based Co-writing*

HaLLMark is a web-based co-writing tool for large language models (LLM) that stores and visualizes a writer's interaction with the LLM (Figure 2). The system facilitates writers to self-reflect on their use of the LLM by clearly highlighting AI writing and prompting activities (editing vs. generating). The motivation is that by capturing interactions between AI and writers as the document evolves and by supporting interactive exploration of that provenance, a writer will have an enhanced sense of agency, control, and ownership of the final artifact. Provenance can also help writers conform to AI-assisted writing policies and be transparent to publishers and readers.

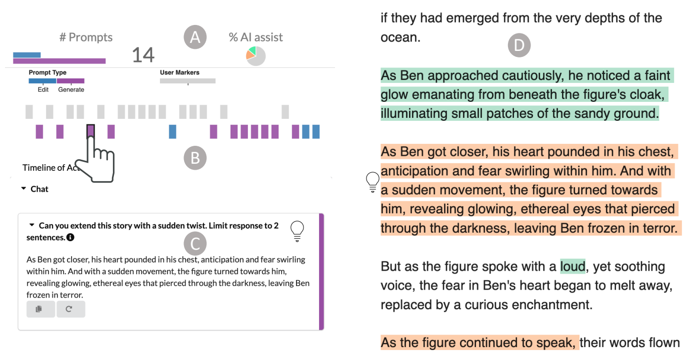


**Figure 2:** Visualization and interaction in Hallmark. a) Summary statistics: number of prompts and percentage of assistance from AI. b) Blue tiles indicate prompts seeking editorial assistance while purple tiles indicate prompts asking to generate new text. Grey bars show the user's writing behavior (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

From evaluations with 13 creative writers, we found that 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 the sense of ownership over the final artifact. Writers also felt that HaLLMark would help them become more transparent in communicating AI co-writing to external parties.

**HCAI Characteristics.** HaLLMark supports all four capabilities of an HCAI tool. It amplifies () and enhances ( ) writing with AI and visualization. It augments () several new capabilities, including tracking provenance as well as ensuring transparency and accountability with regard to AI co-writing policies. It also empowers ( ) writers to control and reflect on how and when they want to use AI support. HaLLMark addresses several of the human concerns: transparency (), accountability (), and provenance ().

### *Outcome-Explorer: An HCAI Tool for Interpretable and Interactive AI Decision-Making*

Outcome-Explorer is a tool for interpretable AI decision-making (Figure 3). Its goal is to help users get answers to explanation queries such as "Why does this model make such decisions?", "What if I change a particular input feature?" or "How will my action change the decision?'" To facilitate this, the system uses a causal model for

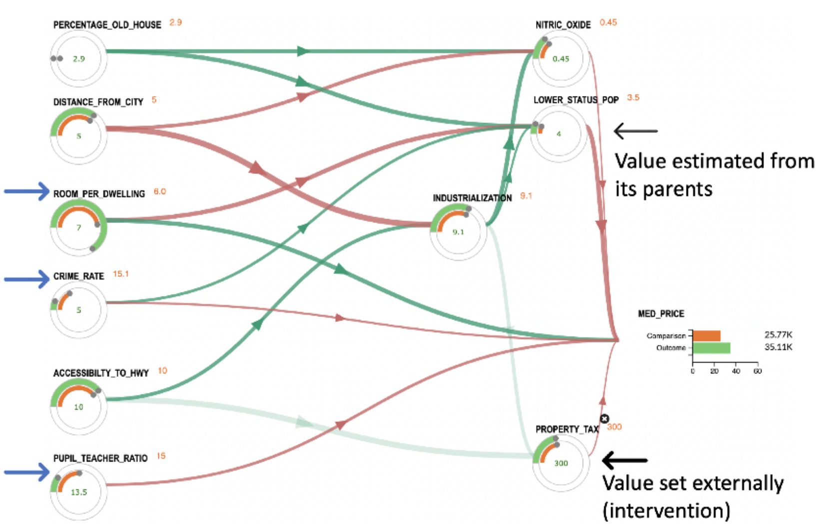


**Figure 3:** Asking what-if questions in Outcome-Explorer. The directed acyclic graph shows causal relations between variables that determine median housing prices in a neighborhood. A user can create a user profile by moving the circular knobs in the nodes (variables). A user can keep one profile (green) fixed and change the other profile (orange) to ask what-if questions. The blue arrows indicate the changes in the orange profile. Note that property tax is set to 300 by the user. As a result, changing its parent nodes will not affect property tax. The other variables are estimated from their parents.

prediction as it is "inherently interpretable": its inner workings are directly observable through a Directed Acyclic Graph (DAG). Users (both expert and non-expert) can interact with the causal model by setting specific values to variables, visualizing changes in the underlying causal model, and then receiving the AI's decision.

**HCAI Characteristics.** Outcome-Explorer facilitates transparency () by clearly representing the relationships between the variables and how they impact the outcome variable (e.g., loan approval decisions). It improves explainability (), understandability (), accountability (), and provenance () through features such as what-if analysis, neighborhood exploration (decisions for similar data profiles), and provenance visualization.

## *Future Work*

My future work will focus on a critical approach to evaluating HCAI tools, designing novel HCAI tools, and inventing new visual representations for foundation models. Here, I outline my vision.

***Evaluating Ethical Dimension of HCAI Tools.*** 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 [1]. However, I do believe that to harness these technologies as supertools, we need to thoroughly evaluate the HCAI tools. Human concerns relevant to HCAI tools (e.g., agency, transparency, ownership) are mostly abstract concepts and are difficult to operationalize in research for evaluation [18]. My future work will focus on devising methods and experimental designs to evaluate HCAI tools.

***Designing HCAI Tools for Diverse Domains.*** To establish visualization-enabled HCAI tools as a prominent HCI approach, we need to explore more domains and areas. My current research mainly explores AI-assisted writing and human-in-the-loop machine learning, critical domains that provide 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 [15]. 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 [17]. I am excited to collaborate and write grant applications with researchers from these domains in my future workplace.

***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 [12-14] towards that goal. For example, **Visual Concept Programming** [13] summarizes a large-scale image dataset by extracting visual concepts from the dataset by using a self-supervised representation learning model. 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 HCAI tools. 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.

***Accessible Data Visualization.*** The concept of visualization-enabled HCAI tools 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 HCAI tools if a large portion of the population cannot use them? Motivated by that, I investigated the design of accessible data visualization. I proposed **Susurrus** [16], 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.

## References


[1] 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. Doi: 10.1126/science.adh4451

[2] Md Naimul Hoque, Sungbok Shin, Niklas Elmqvist. Interactive Visualization for Human-Centered AI Tools. 2024. (Under review)

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

[4] Eric Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 159–166, 1999. Doi: 10.1145/302979.303030

[5] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz Guidelines for Human-AI interaction. . In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 3:1–3:13. ACM, New York, 2019. doi: 10.1145/3290605.3300233

[6] Google PAIR. People + AI Guidebook. https://pair.withgoogle.com/guidebook/, 2019. Accessed on March 15

[7] Jeffrey Heer. Agency plus automation: Designing artificial intelligence into interactive systems. Proc. National Academy of Sciences U.S.A.,116(6):1844–1850, 2019. Doi: 10.1073/pnas.1807184115

[8] B. Shneiderman. Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction, 36(6):495–504, 2020. Doi: 10.1080/10447318.2020.1741118

[9] 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. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).* 2024. Doi: 10.48550/arXiv.2311.13057

[10] Md Naimul Hoque, Bhavya Ghai, Niklas Elmqvist. DramatVis Personae: Visual Text Analytics for Identifying Social Biases in Creative Writing. In *Proceedings of the ACM Conference on Designing Interactive Systems (DIS).* 2022. Doi: 10.1145/3532106.3533526

[11] Md Naimul Hoque, Bhavya Ghai, Kari Kraus, Niklas Elmqvist. Portrayal: Leveraging NLP and Visualization for Analyzing Fictional Characters. In *Proceedings of the ACM Conference on Designing Interactive Systems (DIS).* 2023. Doi: 10.1145/3563657.3596000

[12] Md Naimul Hoque, Niklas Elmqvist. Dataopsy: Scalable and Fluid Visual Exploration using Aggregate Query Sculpting. *IEEE Transaction on Visualization and Computer Graphics (TVCG),* 30, 1, 186–196. 2023. Doi: 10.1109/TVCG.2023.3326594.

[13] 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)*, 29, 1, 74-83. 2023, doi: 10.1109/TVCG.2022.3209466

[14] 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)*. 28, 12, 4728-4740. 2022, doi: 10.1109/TVCG.2021.3102051

[15] Md Naimul Hoque, Ayman A Mahfuz, Mayukha Kindi, Naeemul Hassan. Towards Designing a Question-Answering Chatbot for Online News: Understanding Questions and Perspectives. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).* 2024. Doi: 10.48550/arXiv.2312.10650

[16] Md Naimul Hoque, Md Ehtesham-Ul-Haque, Niklas Elmqvist, Syed Masum Billah. Accessible Data Representation with Natural Sounds. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).* 2023. Doi: 10.1145/3544548.3581087

[17] 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.

[18] 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.

[19] Md Naimul Hoque, Nazmus Saquib, Syed Masum Billah, and Klaus Mueller. 2020. Toward Interactively Balancing the Screen Time of Actors Based on Observable Phenotypic Traits in Live Telecast. *Proc. ACM Hum.-Comput. Interact*. 4, CSCW2, Article 154, 19 pages. Doi: 10.1145/3415225