



# RevTogether: Supporting Science Story Revision with Multiple AI Agents

Yu Zhang

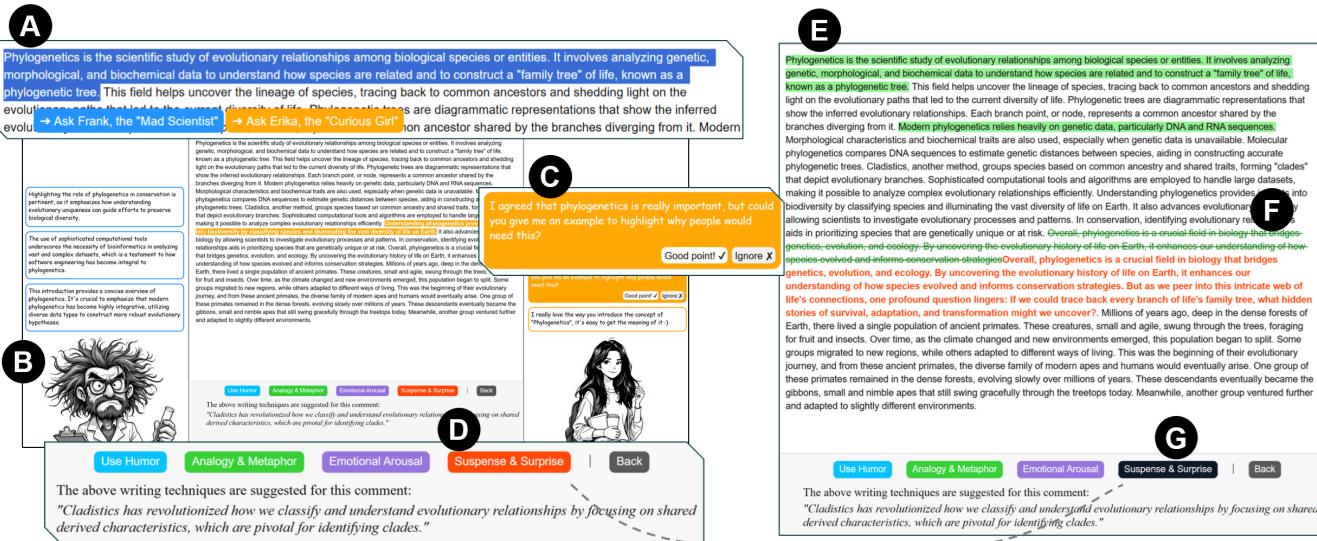
Department of Computer Science  
City University of Hong Kong  
Hong Kong, Hong Kong  
yui.zhang@my.cityu.edu.hk

Kexue Fu

School of Creative Media  
City University of Hong Kong  
Hong Kong, Hong Kong  
kexuefu2-c@my.cityu.edu.hk

Zhicong Lu

Department of Computer Science  
George Mason University  
Fairfax, Virginia, USA  
zlu6@gmu.edu



**Figure 1: Overview of the RevTogether System. (A)** The writer can ask the commentator agent for feedback on a randomly selected segment of the texts. **(B)** The commentator agents present human-like avatars with emotional appearances matching the sentiment of comments. **(C)** The writer can review and respond to each comment. **(D)** Once the writer agrees with a comment, a writing assistant agent will provide writing technique suggestions for addressing the problems mentioned in comments. **(G)** The writer can further choose a technique to see **(E)** which parts of the texts are suggested to apply the technique and **(F)** review detailed revision suggestions provided by AI agents.

## Abstract

As a popular form of science communication, science stories attract readers because they combine engaging narratives with comprehensible scientific knowledge. However, crafting such stories requires substantial skill and effort, as writers must navigate complex scientific concepts and transform them into coherent and accessible narratives tailored to audiences with varying levels of scientific literacy. To address the challenge, we propose RevTogether, a multi-agent system (MAS) designed to support revision of science stories with human-like AI agents (using GPT-4o). RevTogether allows AI agents to simulate affects in addition to providing comments and

writing suggestions, while offering varying degrees of user agency. Our preliminary user study with non-expert writers (N=3) highlighted the need for transparency in AI agents' decision-making processes to support learning and suggested that emotional interactions could enhance human-AI collaboration in science storytelling.

## CCS Concepts

- Human-centered computing → Interactive systems and tools.

## Keywords

Writing Support, Human-AI Collaboration, Science Storytelling

## ACM Reference Format:

Yu Zhang, Kexue Fu, and Zhicong Lu. 2025. RevTogether: Supporting Science Story Revision with Multiple AI Agents. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26–May 01, 2025, Yokohama, Japan*. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3706599.3719888>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI EA '25, Yokohama, Japan

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1395-8/25/04

<https://doi.org/10.1145/3706599.3719888>

## 1 Introduction

Science stories have long captivated broad audiences with varying levels of scientific literacy by blending engagement, entertainment, and comprehensible scientific knowledge [10, 19, 43]. However, crafting quality stories from complex scientific facts highly relies on writers' expertise and mastery of advanced writing strategies [2, 12, 16]. For example, even scientists, who are usually the primary source of new scientific discoveries, also struggle to write engaging and accessible narratives to share their research insights with the general public [9, 33, 37]. Fortunately, advanced computing and artificial intelligence (AI) technologies could bring opportunities to lower the barriers to science story creation, empowering a wider range of individuals to participate in science communication.

With the rapid development of Large Language Models (LLMs) [6] and Generative Artificial Intelligence (GenAI) [25], researchers in multiple disciplines including Human-computer Interaction (HCI) have noticed the potentials of applying AI to support various tasks in science communication [1, 4, 29, 40], including writing for audiences from the general public [14, 18, 31, 34]. Previous work has explored approaches to scaffolding the creation of science news and social media content using AI [13, 21, 22]. For example, LLMs are used to apply particular writing strategies (e.g., “hook” in Long et al. [24] and “metaphor” in Kim et al. [21]), provide creative inspirations [13], and retrieve additional information from external sources [22]. However, these approaches neglect feedback from other perspectives, focusing solely on fulfilling the writer's demands. Feedback from different perspectives is essential in writing because it helps to build an iterative revision process for writers to continuously improve the outcomes [3, 11, 26]. Lacking feedback prevents writers from making changes with clear directional guidance [24, 39], especially for science stories that can have a wide range of readers from scientists to the general public [7, 12].

Inspired by Benharak et al.'s previous research that highlights the potential of social factors (e.g., relation, valence) in designing AI personas for writing feedback [3], we propose RevTogether, a multi-agent system (MAS) that supports writers in revising science stories together with two human-like commentator AI agents and a writing assistant agent. Specifically, we designed the personas of the two commentator agents as a “mad scientist” and a “curious girl”, representing the two sides of the “dialogue model” in science communication—scientists and ordinary people [30, 35]—for providing feedback in diverse perspectives. To enhance emotional expression, we introduced character avatars with dynamic emotional responses for the commentator agents in RevTogether, providing non-verbal feedback to engage the writers.

Through a preliminary user study with three non-expert writers, we highlighted the importance of the transparency of AI agent's suggestion generation process because it could be helpful for writers who want to learn from AI. This finding aligned with the previous work on human-AI alignment in writing tasks [8, 14, 15, 38], revealing *learning* as a new demand in human-AI collaborative writing. Besides, we also found that emotional reactions simulated by AI agents could serve as effective feedback to stimulate writers' thoughts in science storytelling. Our findings could inspire future research on AI agent design for broad tasks in content creation in the context of science communication.

## 2 System Design and Implementation

RevTogether aims to support science story revision with on-demand feedback provided by two commentator agents and a writing assistant agent. To begin with, we chose the Human-AI teaming (HAT) analysis framework [27, 32] as design guidelines to operationalize our idea into a feasible prototype system. According to the HAT framework, RevTogether is a *multi-agent-single-human* system in which the two commentator agents and the writing assistant agent collaborate and interact with a single human writer. With different roles played and tasks performed by the agents, RevTogether offers multiple levels of user agency.

The overall layout of RevTogether's user interface (UI) is similar to common document editing tools (e.g., Microsoft Word and Google Docs) where the centric area is used for text editing and the side columns are used for supportive information such as commentary and the character avatars of commentator agents (Figure 1). We design the UI more like an editing tool rather than chatbot is because we hope to highlight science story writing as the main task. Direct conversation with agents are intentionally not supported as it may interrupt writers' thought processes.

### 2.1 Designing the Commentator Agents

In order to simulate feedback from potential readers with different perspectives, we design and implement two commentator agents that role-play a “mad scientist” and a “curious girl”. The two specific personas are derived from the two types of major stakeholders (i.e., scientists and the general public), as well as typical readers of science stories, in science communication. Based on the HAT framework, the commentator agents are *critic* type [32] of agents who take the responsibilities for giving feedback on demand.

**2.1.1 Asking for Comments.** When editing a science story using RevTogether, the writer can ask commentator agents to comment on user-specified segments (e.g., sentences, paragraphs, and full texts) of the manuscript. Precisely, when selecting a piece of texts, a floating menu will appear beneath the selection area to let the writer decide whom to ask for (Figure 1 (A)). Once the writer chooses either of the commentators, the corresponding agent will take the selected texts as a focus to make a comment. Newly-generated comments will appear above the agent's character avatars next to the editing area (Figure 1 (B)). Multiple comments are placed in a bottom-up direction from the latest to the earliest. When hovering mouse on a piece of comment, the writer can either accept (clicking on “Good point!”) or reject (clicking on “ignore”) it (Figure 1 (C)). The actions will trigger agents to respond with emotional reactions.

**2.1.2 Emotional Reactions.** To establish the emotional engagement between the writer and commentator agents by leveraging the human-likeness perspective in the HAT framework [5, 20], we introduced comic-style character avatars for the two commentators displaying beside the editing area in RevTogether (Figure 1 (B)).

Particularly, for each commentator we created three avatars in different affects (emotional states): positive (i.e., happy), neutral (i.e., calm), and negative (i.e., angry or disappointed). Thus, we use the avatars as a non-verbal cue to convey emotional reactions (Figure 2). When asked to generate comments, we require the LLM to indicate the sentiment of each comment. The sentiment value can



**Figure 2: Character Avatars with different affects for Commentator Agents.** Each character has three avatars with positive, neutral, and negative affect respectively. They share “happy” and “calm” for positive and neutral affects, while the “mad scientist” character has “angry” while the “curious girl” character has “disappointed” as negative affect.

be positive, neutral, or negative. With the sentiment information associated with each comment, while the writer hovers the mouse over a comment for reviewing, the commentator character will change to the avatar of the corresponding affect. While the writer not hovering or there is not comment yet, the character will remain in the neutral avatar by default.

Besides reflecting on the sentiment of comments, we additionally allow the commentator characters to respond to the writer’s acceptance/rejection actions with affect changing actions. If the writer accepts a comment, the character will turn to the positive avatar staying for one second and turn back. The action performs like the avatar smiles to the writer as a compliment. On the contrary, rejection action will triggers the negative avatar as a complain. The actions are designed to be in a non-verbal way to avoid interruptions.

## 2.2 Designing the Writing Assistant Agent

While human-like comments can provide some clues for how the manuscript could be changed, the writer still needs to figure out specific writing strategies for improvement. To provide clearer guidance of composing compelling science stories, we introduce the writing assistant agent. Unlike the commentators, the assistant agent is designed to be of *companion* type [32], that is less active and passively responds to the writer’s requests to provide suggestions with particular writing techniques. Therefore, the writing assistant agent does not have an avatar as it does not need to provide human-like comments and conveys emotional reactions to the writer.

The assistant agent will only be activated when the writer explicitly accepts a comment made by the commentator agent. Once a comment is accepted, the suggestions will be displayed below the editing box as a series of tags (Figure 1 (D)) indicating which writing techniques can be used for addressing the comment. When the writer clicks on a tag, the agent will highlight all the places that could be changed by adding a green background to the texts, giving visual hints to the writer for further editing (Figure 1 (E)). Although each comment is associated with a piece of user-specified text, the writing suggestions will still consider the full story. Because only make changes in the specified texts could be limited and break the consistency through the entire story.

In addition to highlighting the places for changing, the assistant agent also provides an AI-revised version of each highlighted place.

Once the writer clicks on a highlighted segment, the auto-revised version will be displayed inline using a red color and a weighted font (Figure 1 (F)). The writer can easily adopt the auto-revised version by double clicking on it.

**2.2.1 Embedding Writing Techniques in Agent’s Suggestions.** To incorporate effective feedback for revision, we analyzed the common techniques in science story writing from the literature and embedded these techniques into prompts so that the writing assistant agent can make suggestions with those concrete techniques. In this way the agent can reduce the chance to give verbose or unspecific feedback [3]. Since the writing techniques and strategies of science stories refers to an endless space of creativity, we only adopted four well-established techniques commonly mentioned in previous research [2, 10, 17, 21, 41–43] in RevTogether at this moment. Specifically, we summarize and present the detailed definitions, purposes, and literature sources for those techniques (Table 1).

## 2.3 Supporting Revision in Different Levels of User Agency

Using RevTogether, the writer can always take the initiative to freely edit the manuscript, while RevTogether enables an additive AI-supported iterative revision process, allowing the writer to receive supportive information on three levels of user agency (Figure 3).

**2.3.1 Human-like comments for high agency.** (Figure 3 (A)) The human-like comments made by commentator agents leave high agency to users. The commentator agents will not interfere when editing texts, but will only share subjective feedback like humans. The writer can use the provided comments at will to figure out a strategy to improve the draft.

**2.3.2 Suggestions on writing techniques for medium agency.** (Figure 3 (B)) While the clues obtained from subjective comments are usually ambiguous and unclear, the writing assistant agent thus is introduced to provide suggestions about specific writing techniques. First, when the writer explicitly accepts a comment, the writing assistant agent will reveal which techniques are possibly useful for addressing the issues mentioned by the comment (Figure 1 (D)). At this stage, the agent still does not edit any texts but provide clearer directions for the writer to think about and move forward. This

**Table 1: Techniques for Science Story Writing. We use four techniques in RevTogether: (1) Humor, (2) Analogy and Metaphor, (3) Emotional Arousal, and (4) Suspense and Surprise.**

Technique	Definition	Purpose	Source
Humor	The use of wit, jokes, or light-hearted language to make complex topics more engaging and enjoyable.	(1) Capture attention, (2) Simplify understanding, (3) Make the content relatable to readers.	[41, 42]
Analogy and Metaphor	Compare complex ideas to familiar concepts to enhance understanding.	(1) Simplify obscure topics by relating them to everyday experiences or imagery, (2) Make the content memorable.	[2, 21]
Emotional Arousal	The use of evocative language or storytelling to trigger readers' emotions and create a deeper connection	(1) Engage readers, (2) Make the content memorable, (3) Inspire curiosity.	[10, 42, 43]
Suspense and Surprise	Build anticipation through uncertainty and captivate readers by delivering unexpected twists or revelations.	(1) Engage readers, (2) Stimulating curiosity, (3) Make the content memorable.	[2, 17, 43]

leaves a medium level of user agency in the human-AI collaboration in RevTogether.

**2.3.3 Suggestions on revisions for low agency.** (Figure 3 (C)) Once the user chooses to let the agent to provide revision suggestions (Figure 1 (G)), the agent will then give hints on which parts of the texts the writing technique can apply to (Figure 1 (E)). Only when the writer clicks on a hint will the system completely show an AI-revised version of texts (Figure 1 (F)) allowing the writer to simply accept by clicking on the revised texts (shown in red color). The revision suggestion feature of writing assistant agent leaves the lowest user agency in RevTogether as AI takes the control of editing and directly involved in story generation.

## 2.4 Implementation

We built RevTogether as a Web application using Flask<sup>1</sup> on Python for its back-end and React<sup>2</sup> for the front-end. Particularly, we employ react-quill<sup>3</sup> to implement the text editor for achieving customized text selection, highlighting and editing. For the large language model (LLM) used to power the AI agents, we adopted GPT-4o<sup>4</sup> offered by a commercial cloud service provider that is accessible and stable in our region. All the features of the AI agents in RevTogether are achieved by our customized prompts and the conversations between the back-end and the remote LLM, where the back-end is responsible for assembling dynamic parameters (e.g., user-selected texts) from the front-end into predefined prompt templates and process the result from the remote LLM. Particularly, the avatars of the commentator agents are manually generated using the FLUX-pro-1.1<sup>5</sup> model.

## 3 Preliminary User Study

To explore the potential of human-AI collaborative writing in science story creation with RevTogether, we conducted a preliminary user study with 3 participants (1 female, 2 male, aged between 22 to

30) who are all PhD students from our local university to evaluate RevTogether. The study results highlight the importance of the transparency of AI agent's suggestion generation and the effects of emotional reactions simulated by AI agents in RevTogether. These findings provide design implications for future iterations of RevTogether and reveal opportunities for future research on human-AI collaborative content creation in science communication.

### 3.1 Procedure

For each participant, the study lasted for about 60 minutes and the session was conducted in a one-on-one manner with the first author of this paper. The procedure started from signing the consent (~5 min), getting familiar with RevTogether, and knowing the backgrounds and motivations of building RevTogether (~10 min). Next, we required the participant to generate a short story based on an academic paper they like, which can also be one of their own published papers, using GPT-4o with a fixed prompt given by us (only replacing the PDF file uploaded to the conversation). Using this story as an initial draft, each participant was instructed to request comments from both commentator agents at least twice and to accept at least one comment from each for further interaction with the writing assistant. The ultimate goal is to compose a better science story with RevTogether. During the editing process, we encouraged the participants to think-aloud about their experience using RevTogether. The editing process lasted about 30 minutes, and after which we conducted a short post-study interview with each participant, asking them to reflect on their thoughts and the most impressive experience with RevTogether. The entire study procedure was screen-recorded and audio-recorded for analysis.

### 3.2 Findings and Future Directions

Given the exploratory nature of our study and the limited sample size of three participants (referred to as P1, P2, and P3), our findings do not allow for definitive conclusions regarding the effectiveness of our design. Instead, we focus on reporting the key insights that emerged from the preliminary user study. These findings highlight potential areas of interest and provide a foundation for future research directions to further investigate and refine RevTogether.

<sup>1</sup><https://flask.palletsprojects.com/en/stable/>

<sup>2</sup><https://react.dev/>

<sup>3</sup><https://github.com/zenoamaro/react-quill>

<sup>4</sup><https://openai.com/index/hello-gpt-4o/>

<sup>5</sup><https://www.fluxpro.ai/flux-1.1-pro>



**Figure 3: The Iterative Revision Workflow.** RevTogether enables an AI-supported iterative revision process, which starts from free editing and allows the writer to receive supportive information on three levels of user agency throughout the process: (A) Human-like comments with high agency; (B) Writing technique suggestions with medium agency; and (C) Detailed Revision suggestions with low agency.

**3.2.1 Highlighting the Transparency of the Decision-making Process of AI Agents.** Explicitly showing a “thought process” of AI agents can help writers make sense of AI-generated suggestions and learn from the suggestions to improve their own skills in science story writing. When asked about system features to extend in post-study interview, P1 asked “*Can the system show me how AI reads the story line by line to decide which parts of the texts are applicable for adding humor? Because I want to know the thoughts behind a suggestion to determine if it is reasonable or just an illusion of the language model behind.*”. P1 believed a visible step-by-step process showing how AI agent comes up with a writing suggestion would help him to learn from AI in science story writing, and also make him more confident with his later decision to either accept or reject an revision offered by the writing assistant agent. P3 also mentioned about the idea of learning from AI through a more transparent process of obtaining suggestions from the writing assistant agent. “*By using the system for writing such a story, I always want to have improvement on my own writing skills rather than letting AI write for me, otherwise why didn't I just let GPT generate full stories and keep improving by itself? I hope to see more details and reasons behind to have deeper understandings of the writing techniques*”, said P3.

In summary, we found that more information regarding how and why AI makes a suggestion (i.e., the transparency of decision-making of AI agents) need to be exposed in the interactions between the human writer and AI agents, so that the writer can learn from and rely on this information to make their own decisions effectively. This finding echoes previous research on human-AI alignment in AI-supported writing systems [8, 14, 15, 38], showing that allowing users to better learn from AI should also be a design consideration for building AI agents for writing.

**3.2.2 The Effects of Emotional Expressions of AI agents in Science Story Writing.** Although our participants knew that the emotional reactions of AI agents’ avatars were simulated by programmed rules and LLMs rather than real human emotional expressions, they could still feel strong feedback from the changing face expressions of the avatars, which is a kind of non-verbal cue in human-AI communication [28]. P3 mentioned, “*Seeing the agent appears unhappy, I felt*

*like I was motivated to carefully re-consider the comment again. This reminds me that my story could be read by people with very different opinions and might affect their emotions.*” All of the three participants agreed that the emotional reactions as feedback affected their decision making during story editing. “*It's really interesting and funny to see the avatar change faces when accepting or rejecting a comment*”, P2 commented.

Previous research have revealed the importance of emotion in science communication in terms of how emotion of the audiences could be affected and considered in design and content creation to enhance the public engagement of science [17, 23, 36, 43]. However, how the emotion of content creators (e.g. science story writers) and the simulated “emotion” of AI agents can affect science communication content creation, remains under-explored. Future research in this line may extend the knowledge of the effects of emotion in human-AI co-creation in the contexts of science communication.

## 4 Limitations and Future Work

While our preliminary user study demonstrates the potential of RevTogether in supporting science story revision, there are several limitations to address. First, the small sample size of three participants, all of whom were non-expert writers, limits the generalizability of our findings. Secondly, while emotional expressions of AI agents were perceived by the participants in the preliminary study, their actual impact on user engagement and human-AI collaboration was not systematically examined, leaving their effectiveness as an open question.

In future work, we plan to expand the scale of our user study by involving a larger and more diverse group of participants to obtain a more comprehensive and robust evaluation of the system’s usefulness. Additionally, we aim to invite professional science story writers to assess RevTogether, providing expert perspectives on its design and functionality to further refine the system. Furthermore, we intend to conduct ablation studies to investigate whether the current emotional expressions of AI agents effectively elicit user responses and enhance their collaboration and engagement with the agents. These efforts will help us gain deeper insights into

RevTogether's impact and identify areas for improvement to better support human-AI collaboration in science storytelling.

## 5 Conclusion

In this paper, we propose RevTogether, a multi-agent system (MAS) that facilitates the iterative revision process in science story creation by employing Large Language Model (LLM)-based AI agents to provide feedback and suggestions with different levels of user agency. Through a preliminary user study with three non-expert writers, we revealed the potential of RevTogether and future research directions regarding understanding and supporting human-AI collaboration in writing tasks within the contexts of scientific communication.

## References

- [1] Amanda Alvarez, Aylin Caliskan, MJ Crockett, Shirley S Ho, Lisa Messeri, and Jevin West. 2024. Science communication with generative AI. *Nature Human Behaviour* 8, 4 (2024), 625–627.
- [2] Tal August, Lauren Kim, Katharina Reinecke, and Noah A Smith. 2020. Writing Strategies for Science Communication: Data and Computational Analysis. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Online, 5327–5344. doi:10.18653/v1/2020.emnlp-main.429
- [3] Karim Benharrak, Tim Zindulka, Florian Lehmann, Hendrik Heuer, and Daniel Buschek. 2024. Writer-Defined AI Personas for On-Demand Feedback Generation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 1049, 18 pages. doi:10.1145/3613904.3642406
- [4] Sibusiso Biyela, Kanta Dihal, Katy Ilonka Gero, Daphne Ippolito, Filippo Menczer, Mike S Schäfer, and Hiromi M Yokoyama. 2024. Generative AI and science communication in the physical sciences. *Nature Reviews Physics* 6, 3 (2024), 162–165.
- [5] Justin W. Bonny and Kevin T. Wynne. 2024. Increasing Human-Likeness and Acceptance of Conversational Autonomy through Experience. In *2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS)*. 1–6. doi:10.1109/ICHMS59971.2024.10555683
- [6] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.
- [7] Terry W Burns, D John O'Connor, and Susan M Stocklmayer. 2003. Science communication: a contemporary definition. *Public understanding of science* 12, 2 (2003), 183–202.
- [8] Tuhin Chakrabarty, Philippe Laban, and Chien-Sheng Wu. 2024. Can AI writing be salvaged? Mitigating Idiosyncrasies and Improving Human-AI Alignment in the Writing Process through Edits. *arXiv preprint arXiv:2409.14509* (2024).
- [9] Isabelle M Côté and Emily S Darling. 2018. Scientists on Twitter: Preaching to the choir or singing from the rooftops? *Facets* 3, 1 (2018), 682–694.
- [10] Sarah R Davies, Megan Halpern, Maja Horst, David A Kirby, and Bruce Lewenstein. 2019. Science stories as culture: experience, identity, narrative and emotion in public communication of science. *Journal of Science Communication* 18, 5 (2019).
- [11] Wanyu Du, Vipul Raheja, Dhruv Kumar, Zae Myung Kim, Melissa Lopez, and Dongyeop Kang. 2022. Understanding iterative revision from human-written text. *arXiv preprint arXiv:2203.03802* (2022).
- [12] Sara J ElShafie. 2018. Making science meaningful for broad audiences through stories. *Integrative and comparative biology* 58, 6 (2018), 1213–1223.
- [13] Katy Ilonka Gero, Vivian Liu, and Lydia Chilton. 2022. Sparks: Inspiration for Science Writing using Language Models. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference (Virtual Event, Australia) (DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1002–1019. doi:10.1145/3532106.3533533
- [14] Katy Ilonka Gero, Tao Long, and Lydia B Chilton. 2023. Social Dynamics of AI Support in Creative Writing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 245, 15 pages. doi:10.1145/3544548.3580782
- [15] Nitesh Goyal, Minsuk Chang, and Michael Terry. 2024. Designing for Human-Agent Alignment: Understanding what humans want from their agents. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '24). Association for Computing Machinery, New York, NY, USA, Article 106, 6 pages. doi:10.1145/3613905.3650948
- [16] Changyang He, Yue Deng, Lu He, Qingyu Guo, Yu Zhang, Zhicong Lu, and Bo Li. 2024. Engage Wider Audience or Facilitate Quality Answers? a Mixed-methods Analysis of Questioning Strategies for Research Sensemaking on a Community Q&A Site. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 50 (April 2024), 31 pages. doi:10.1145/3637327
- [17] Tianle Huang and Will J Grant. 2020. A good story well told: Storytelling components that impact science video popularity on YouTube. *Frontiers in Communication* 5 (2020), 86.
- [18] Gongyao Jiang, Xinran Shi, and Qiong Luo. 2024. Llm-collaboration on automatic science journalism for the general audience. *arXiv preprint arXiv:2407.09756* (2024).
- [19] Marina Joubert, Lloyd Davis, and Jenni Metcalfe. 2019. Storytelling: the soul of science communication. E pages.
- [20] Joohee Kim and Il Im. 2023. Anthropomorphic response: Understanding interactions between humans and artificial intelligence agents. *Computers in Human Behavior* 139 (2023), 107512.
- [21] Jeongyeon Kim, Sangho Suh, Lydia B Chilton, and Haijun Xia. 2023. Metaphorian: Leveraging Large Language Models to Support Extended Metaphor Creation for Science Writing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) (DIS '23). Association for Computing Machinery, New York, NY, USA, 115–135. doi:10.1145/3563657.3595596
- [22] Philippe Laban, Jesse Vig, Marti Hearst, Caiming Xiong, and Chien-Sheng Wu. 2024. Beyond the Chat: Executable and Verifiable Text-Editing with LLMs. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) (UIST '24). Association for Computing Machinery, New York, NY, USA, Article 20, 23 pages. doi:10.1145/3654777.3676419
- [23] Huann-shyang Lin, Zuway-R Hong, and Tai-Chu Huang. 2012. The role of emotional factors in building public scientific literacy and engagement with science. *International Journal of Science Education* 34, 1 (2012), 25–42.
- [24] Tao Long, Dorothy Zhang, Grace Li, Batool Taraif, Samia Menon, Kynnedy Simone Smith, Sitong Wang, Katy Ilonka Gero, and Lydia B Chilton. 2023. Tweeterial hooks: generative AI tools to motivate science on social media. *arXiv preprint arXiv:2305.12265* (2023).
- [25] Michael Muller, Lydia B Chilton, Anna Kantosalo, Charles Patrick Martin, and Greg Walsh. 2022. GenAIChI: Generative AI and HCI. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI EA '22). Association for Computing Machinery, New York, NY, USA, Article 110, 7 pages. doi:10.1145/3491101.3503719
- [26] Melissa M Nelson and Christian D Schunn. 2009. The nature of feedback: How different types of peer feedback affect writing performance. *Instructional science* 37 (2009), 375–401.
- [27] Amanda K. Newendorp, Mohammadamin Sanaei, Arthur J Perron, Hila Sabouni, Nikoo Javadpour, Maddie Sells, Katherine Nelson, Michael Dorneich, and Stephen B. Gilbert. 2024. Apple's Knowledge Navigator: Why Doesn't that Conversational Agent Exist Yet?. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 153, 14 pages. doi:10.1145/3613904.3642739
- [28] Joshua Newn, Ronal Singh, Fraser Allison, Prashan Madumal, Eduardo Veloso, and Frank Vetere. 2020. Nonverbal Communication in Human-AI Interaction: Opportunities and Challenges. *Human Computer Interaction and Emerging Technologies: Adjunct Proceedings from* (2020), 221.
- [29] Sachita Nishal, Jasmine Sinhai, and Nicholas Diakopoulos. 2024. Understanding Practices around Computational News Discovery Tools in the Domain of Science Journalism. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 142 (April 2024), 36 pages. doi:10.1145/3637419
- [30] Cathelijne M Reinke, Annelien L Bredenoord, and Marc HW van Mil. 2020. From deficit to dialogue in science communication: the dialogue communication model requires additional roles from scientists. *EMBO reports* 21, 9 (2020), e51278.
- [31] Mike S Schäfer, Bastian Kremer, Niels G Mede, and Lilianne Fischer. 2024. Trust in science, trust in ChatGPT? How Germans think about generative AI as a source in science communication. *Journal of Science Communication* 23, 9 (2024), A04.
- [32] Nathan Sepich, Michael C Dorneich, and Stephen Gilbert. 2021. Human-Agent Team Game Analysis Framework: Case Studies. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 65, 1 (2021), 1146–1150. doi:10.1177/1071181321651188 arXiv:https://doi.org/10.1177/1071181321651188
- [33] C. Estelle Smith, Xinyi Wang, Raghav Pavan Karumur, and Haiyi Zhu. 2018. [Un]breaking News: Design Opportunities for Enhancing Collaboration in Scientific Media Production. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3173955
- [34] Mico Tatalovic. 2018. AI writing bots are about to revolutionise science journalism: we must shape how this is done. *Journal of Science Communication* 17, 1 (2018), E.
- [35] Brian Trench. 2008. *Towards an Analytical Framework of Science Communication Models*. Springer Netherlands, Dordrecht, 119–135. doi:10.1007/978-1-4020-8598-7\_7

- [36] Brian Trench and Massimiano Buccchi. 2021. Rethinking science communication as the social conversation around science. *Journal of Science Communication* 20, 3 (2021), 1–11.
- [37] Matthew J Tuttle, Daniela Cejas, Donghoon Kang, Francis Muchaamba, Britta Goncarovs, Yaprak Ozakman, Fatima Aziz, and Arnaud Orelle. 2023. Promoting Science Literacy and Awareness across the Globe: the Role of Scientists as Science Ambassadors. *Journal of Microbiology and Biology Education* 24, 2 (2023), e00041–23.
- [38] Sitong Wang, Xuanming Zhang, Jenny Ma, Alyssa Hwang, Zhou Yu, and Lydia B. Chilton. 2025. JumpStarter: Getting Started on Personal Goals with Adaptive Personal Context Curation. arXiv:2410.03882 [cs.HC] <https://arxiv.org/abs/2410.03882>
- [39] Gillian Wigglesworth and Neomy Storch. 2012. What role for collaboration in writing and writing feedback. *Journal of second language writing* 21, 4 (2012), 364–374.
- [40] Spencer Williams, Ridley Jones, Katharina Reinecke, and Gary Hsieh. 2022. An HCI Research Agenda for Online Science Communication. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 490 (Nov. 2022), 22 pages. doi:10.1145/3555591
- [41] Sara K Yeo, Ashley A Anderson, Amy B Becker, and Michael A Cacciatore. 2020. Scientists as comedians: The effects of humor on perceptions of scientists and scientific messages. *Public Understanding of Science* 29, 4 (2020), 408–418.
- [42] Sara K Yeo and Meaghan McKasy. 2021. Emotion and humor as misinformation antidotes. *Proceedings of the National Academy of Sciences* 118, 15 (2021), e2002484118.
- [43] Yu Zhang, Changyang He, Huachen Wang, and Zhicong Lu. 2023. Understanding Communication Strategies and Viewer Engagement with Science Knowledge Videos on Bilibili. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 668, 18 pages. doi:10.1145/3544548.3581476