# Navigating Generative AI Disclosure, Ownership, and Accountability in Co-Creative Domains

Hyo Jin Do
IBM Research
Cambridge, MA, USA
hjdo@ibm.com

Molly Q Feldman
Oberlin College
Oberlin, OH, USA
mfeldman@oberlin.edu

Jessica He
IBM Research
Seattle, WA, USA
jessicahe@ibm.com

Angel Hsing-Chi Hwang
University of Southern California
Los Angeles, CA, USA
angel.hwang@usc.edu

Seyun Kim
Carnegie Mellon University
Pittsburgh, PA, USA
seyunkim@cs.cmu.edu

## Abstract

The increasing integration of generative AI into work has amplified issues of disclosure, ownership, and accountability, including whether and how to acknowledge AI use, who owns AI-generated or co-created work, and who is accountable for risks. In response, governments, organizations, and researchers are introducing new policies, guidelines, and methods for enhanced transparency. However, the complex interplay between multiple stakeholders and technologies, coupled with growing AI agency, continues to spark debates about ownership and accountability of co-created work, leading to open questions about whether, when, and how to disclose and attribute human-AI co-created work. To address these emergent issues, this workshop aims to gather interdisciplinary researchers, practitioners, and experts to discuss key questions from law, technology, design, and HCI research standpoints, with the ultimate goal of promoting responsible generative AI use for work.

## Keywords

Generative AI, Disclosure, Attribution, Ownership, Authorship, Accountability, Responsibility, Co-creation

**ACM Reference Format:**
Hyo Jin Do, Molly Q Feldman, Jessica He, Angel Hsing-Chi Hwang, and Seyun Kim. 2025. Navigating Generative AI Disclosure, Ownership, and Accountability in Co-Creative Domains. In *CHIWORK '25 Adjunct: Adjunct Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work (CHIWORK '25 Adjunct), June 23–25, 2025, Amsterdam, Netherlands.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3707640.3729212

## 1 Introduction and Motivation

The development of generative AI, including large language models (LLMs), has enabled people to co-create with AI, transforming the way we work in numerous domains such as programming [9, 42], writing [29], and art [31]. However, these advancements have also raised new concerns. LLMs are known to produce outputs that are plausible but not necessarily correct, potentially spreading misinformation [48]. AI systems also have the capability to learn, encode, and disseminate harmful biases [4, 13, 26]. Since AI-generated content is often not easily distinguishable from human-generated content, risks of deception and obscuring responsibility are significantly increased [20, 41]. Consequently, legal disputes over ownership and accountability for AI-assisted work are intensifying [2, 15, 49], and a 2024 Stack Overflow Developer Survey [10] reveals a growing gap between the use of AI and user trust in AI-generated outputs. These issues highlight the need for new forms of *AI disclosure, ownership* and *accountability* that are tailored to, or calibrated for, the unique information needs of individual (e.g., co-creators, developers, end-users, managers) and organizational (e.g., companies, regulatory agencies) stakeholders. The objective of this workshop is to initiate multi-disciplinary conversations and collaborations to drive responsible outcomes in work domains involving *human-AI co-creation*. In the following sections, we describe relevant developments and gaps that we aim to address.

**Laws, policies, and regulations:** Governments and organizations have sought to address emerging disclosure, ownership, and accountability needs by introducing new laws, policies, and regulations. For example, the EU AI Act and the California AI Transparency Act require deployers or content providers to disclose the use of AI [24, 36]. Many academic journals and publishers similarly require the use of AI to be acknowledged, but specify that AI does not meet the criteria for authorship [1, 3, 11, 33, 34, 43]. Copyright laws have begun to define ownership and accountability over AI-assisted work in the US, EU, and China, which have agreed that copyright requires meaningful, original, and significant human contributions [2, 15]. However, these policies and regulations are constantly shifting as AI capabilities and uses evolve, and they are often ambiguous in defining "meaningful" contributions or providing guidance on how disclosures should appear, leaving a gap in understanding how to implement them across work domains.

**Technological and design approaches:** Researchers and practitioners have started developing technological and design approaches for AI disclosure and attribution. These tend to provide detailed information about the outputs, such as what was AI-generated and how, as well as its underlying sources. One common technological approach is watermarking. Examples include Meta's Stable Signature [14] and Google's SynthID [7], both of



*CHIWORK '25 Adjunct, Amsterdam, Netherlands*

ACM ISBN 979-8-4007-1397-2/25/06
https://doi.org/10.1145/3707640.3729212

which create digital watermarks for AI-generated material. Another approach is retrieval augmented generation (RAG), which identifies AI sources by augmenting LLM outputs with references retrieved from a reliable knowledge base [5, 30, 47]. The implementation of methods for detecting AI contributions and tracing their sources remains an active area of research.

Design approaches offer ways to visually differentiate or label AI contributions [41, 44]. IBM's Carbon for AI framework provides reusable style elements, including an AI label and explainability panel, to lend a unique visual identity to AI-generated content [23]. iA Writer, a web plugin, enables authors to denote text provenance and stylizes AI text in a distinct color [22]. Design research in HCI has recommended various approaches and information to include in an AI disclosure based on user needs [44]. While these approaches offer a starting point, understanding when and how to apply them poses a challenge to workers using AI. AI attribution does not follow a one-size-fits-all approach [19, 27], and the development of tailored, context-specific disclosures remains an open challenge.

**Perspectives on ownership and accountability:** While disclosure is essential for understanding AI involvement, ownership plays a critical role in determining who is accountable for providing transparency and ensuring responsible use. There is ongoing research and debate on who owns AI-generated or co-created work [8, 12, 16, 18, 21, 25, 39, 40, 45, 46]. The involvement of multiple stakeholders along the AI pipeline, including data labelers, data owners, developers, and users, further complicates this question [6]. The concept of ownership can be categorized into two perspectives: legal and psychological [45]. Copyright laws often dictate legal ownership [2, 15, 35, 49]. In contrast, psychological ownership refers to the feeling of possessiveness and of being psychologically tied to an object [37, 38, 45]. Psychological ownership enhances a sense of *responsibility* over outcomes involving AI and *accountability* for what happens with these outcomes [28]. However, people have growing concerns about AI's potential threat to psychological ownership, as co-created work may not preserve the creator's *authenticity* [21]. Moreover, sense of ownership may not align with disclosure requirements. Draxler et al. [8] found an AI Ghostwriter effect in which users do not consider themselves the authors of AI-generated text, but also refrain from declaring AI authorship due to expected negative reactions from readers. In AI-assisted code, Weisz et al. [42] posited that "new mechanisms may be needed to track co-creative activities and ensure that each party's contributions...are properly attributed." Developing disclosure approaches that align with the ownership and accountability perceptions of different stakeholders remains an unmet need.

## 2 Workshop Objectives

AI disclosure, ownership, and accountability are interdisciplinary issues involving legal, technological, design, social, and philosophical perspectives. The central goal of this workshop is to bring together diverse stakeholders to initiate conversations and collaborations in addressing emergent issues in this space. Questions that could be explored include:

- **Disclosure & Attribution:** How do we determine and evaluate what disclosure and/or attribution means? What are the next steps to operationalize those definitions for stakeholders?
- **Ownership, Authenticity, and Perceptions:** How can we determine who owns co-created work? How might we preserve human ownership and authenticity? How can we shape our work environments to reduce potential negative perceptions of AI use?
- **Accountability & Responsibility:** Who benefits from AI and who is accountable for the harms of AI outcomes? How should we develop regulations to reduce harms and promote responsibility over AI outcomes?
- **Implications for Work:** How and when should AI disclosure be provided in various co-creative work contexts? How might these choices privilege some workers over others?

While the general topics of human-AI co-creation and generative AI have been addressed in multiple workshops (e.g., [17, 32]), the topics we focus on, particularly in the context of work and the worker, have not been addressed.

## 3 Workshop Mode

We will have a *hybrid* workshop to accommodate participants who may not be able to attend in-person. We will use a videoconferencing platform to stream the full workshop. At least one organizer will be online to coordinate remote participants. We will invite all who are interested in the workshop to submit short papers that address their perspectives toward the workshop themes. Remote participants will present their submissions live or via asynchronous video recordings. We will also provide an online canvas for all attendees to engage with presentations. If at least two participants are online in the group activity, we will group them and facilitate their participation with our online organizer. Otherwise, we will provide a laptop for our in-person participants to interact with remote participants.

## 4 Workshop Activities

Our planned half-day workshop schedule is as follows: Introduction (9:00-9:10), Keynote (9:10-10:10), Participant Presentations (10:10-11:10), Social Coffee Break with Interactive Posters (11:10-11:40), Group Activity (11:40-12:50), and Closing Remarks (12:50-13:00). For the keynote, we will invite a speaker who can share insightful perspectives on related topics. After the keynote, we will ask participants to present an overview of their submission in person, via a videoconferencing platform, or through a video recording. For those unable to present, we will ask them to present their poster either in-person or online during the coffee break. Following the break, we will hold a collaborative co-design activity for participants to further engage with the workshop topics and build connections. Participants will work in groups of 3 to 5, allocated based on their submission topics, to address one of the questions described in Section 2 or another question that arises from their interests. Using a design thinking approach, workshop organizers will guide participants through stakeholder mapping, rapid ideation, prototyping, and share-outs. We will dedicate sufficient time for open discussions after each activity so that participants who share similar interests can interact and learn.

## 5 Plans to Publish Workshop Proceedings

If the authors provide their consent, accepted papers will be published in the workshop proceedings (optional). If they choose, authors may opt-out of archival proceedings and choose a non-archival option. In this case, we will link to their submissions on our workshop website only.

## 6 Call for Participation

Recent advancements in generative AI have raised questions about the disclosure, ownership, and accountability of human-AI co-created work. Specifically, when and how should we disclose AI usage to different stakeholders in various domains, and how do their perceptions of ownership influence the design, implementation, and evaluation of these disclosures? Furthermore, who is accountable for the benefits and risks of AI-assisted outcomes? These questions have profound implications for organizations and individuals navigating emerging policies, risks, and user needs in this space. The goal of this hybrid half-day workshop is to bring together researchers and practitioners from diverse disciplines to establish a collaborative community and explore research, policies, designs, technologies, and practices in this emerging field. We welcome submissions of full or in-progress studies, design fictions, literature reviews, demonstrations, encores of published work, and position papers. Accepted papers will be published in the workshop proceedings with the authors' consent. At least one author of each accepted submission must attend the workshop and register for at least one day of the conference. For more information, please visit our website: https://chiwork-aidisclosure.github.io/.

## 7 Organizers


- *Hyo Jin (Gina) Do* is a Research Scientist at IBM Research in Cambridge, MA, USA. Her research interests lie at the intersection of HCI and AI, particularly in designing trustworthy AI interfaces and systems that cater to users' needs and improve their experience while mitigating potential risks and biases.
- *Molly Q Feldman* is an Assistant Professor of Computer Science at Oberlin College, USA. She studies the challenges of adapting novel technologies, such as generative AI, for use by non-experts. She focuses on programming and writing as work tasks and is interested in workers throughout the technology sector.
- *Jessica He* is a UX Designer at IBM Research based in Seattle, WA, USA. Her work involves applying user-centered design methods to bridge the gap between user expectations and emerging AI technologies, spanning topics such as AI attribution, co-creation, and risk mitigation.
- *Angel Hsing-Chi Hwang* is an Assistant Professor at USC Annenberg School for Communication. Her research explores the societal impact of artificial intelligence (AI) on work practices in various applied domains, such as small-group teamwork, content production, and mental healthcare services. Her work also aims to provide implications for practitioners to design, build, and apply AI-powered technologies for better work futures.
- *Seyun Kim* is a Ph.D. student at Carnegie Mellon University, School of Computer Science, Human-Computer Interaction Institute, USA. Her research is on human-AI iteration, focusing on integrating values and insights from those of impacted stakeholders in the design and development of AI systems.


## References


[1] ACM. 2023. ACM Policy on Authorship. https://www.acm.org/publications/policies/new-acm-policy-on-authorship.

[2] Cooley alert. 2024. Copyright Ownership of Generative AI Outputs Varies Around the World. https://www.cooley.com/news/insight/2024/2024-01-29-copyright-ownership-of-generative-ai-outputs-varies-around-the-world.

[3] arXiv. Date unknown. arXiv moderation. https://info.arxiv.org/help/moderation/index.html#policy-for-authors-use-of-generative-ai-language-tools.

[4] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Virtual Event Canada, 610–623. https://doi.org/10.1145/3442188.3445922

[5] Deng Cai, Yan Wang, Lemao Liu, and Shuming Shi. 2022. Recent advances in retrieval-augmented text generation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 3417–3419.

[6] Mark Coeckelbergh. 2020. Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and engineering ethics* 26, 4 (2020), 2051–2068.

[7] Sumanth Dathathri, Abigail See, Sumedh Ghaisas, Po-Sen Huang, Rob McAdam, Johannes Welbl, Vandana Bachani, Alex Kaskasoli, Robert Stanforth, Tatiana Matejovicova, et al. 2024. Scalable watermarking for identifying large language model outputs. *Nature* 634, 8035 (2024), 818–823.

[8] Fiona Draxler, Anna Werner, Florian Lehmann, Matthias Hoppe, Albrecht Schmidt, Daniel Buschek, and Robin Welsch. 2024. The AI ghostwriter effect: When users do not perceive ownership of AI-generated text but self-declare as authors. *ACM Transactions on Computer-Human Interaction* 31, 2 (2024), 1–40.

[9] Christof Ebert and Panos Louridas. 2023. Generative AI for software practitioners. *IEEE Software* 40, 4 (2023), 30–38.

[10] Ryan Polk Ellen Brandenberger. 2024. Attribution as the foundation of developer trust. https://stackoverflow.blog/2024/09/30/attribution-as-the-foundation-of-developer-trust/.

[11] Elsevier. [n. d.]. The use of generative AI and AI-assisted technologies in writing for Elsevier. https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier.

[12] Ziv Epstein, Sydney Levine, David G Rand, and Iyad Rahwan. 2020. Who gets credit for AI-generated art? *Iscience* 23, 9 (2020).

[13] Xiao Fang, Shangkun Che, Minjia Mao, Hongzhe Zhang, Ming Zhao, and Xiaohang Zhao. 2024. Bias of AI-generated content: an examination of news produced by large language models. *Scientific Reports* 14, 1 (March 2024), 5224. https://doi.org/10.1038/s41598-024-55686-2

[14] Pierre Fernandez, Guillaume Couairon, Hervé Jégou, Matthijs Douze, and Teddy Furon. 2023. The stable signature: Rooting watermarks in latent diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 22466–22477.

[15] Dan Fitzpatrick. 2025. New Copyright Ruling Just Made AI Skills The Biggest Advantage. https://www.forbes.com/sites/danfitzpatrick/2025/02/03/new-copyright-ruling-just-made-ai-skills-the-biggest-advantage/

[16] Katy Ilonka Gero and Lydia B Chilton. 2019. Metaphoria: An algorithmic companion for metaphor creation. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–12.

[17] Werner Geyer, Mary Lou Maher, Justin D Weisz, Daniel Buschek, and Lydia B Chilton. 2024. HAI-GEN 2024: 5th Workshop on Human-AI Co-Creation with Generative Models. In *Companion Proceedings of the 29th International Conference on Intelligent User Interfaces*. 122–124.

[18] Jessica He, Stephanie Houde, Gabriel E Gonzalez, Darío Andrés Silva Moran, Steven I Ross, Michael Muller, and Justin D Weisz. 2024. AI and the Future of Collaborative Work: Group Ideation with an LLM in a Virtual Canvas. In *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work*. 1–14.

[19] Jessica He, Stephanie Houde, and Justin D. Weisz. 2025. Which Contributions Deserve Credit? Perceptions of Attribution in Human-AI Co-Creation. arXiv:2502.18357 [cs.HC] https://arxiv.org/abs/2502.18357

[20] Todd C Helmus. 2022. Artificial intelligence, deepfakes, and disinformation. *RAND Corporation* (2022), 1–24.

[21] Angel Hsing-Chi Hwang, Q Vera Liao, Su Lin Blodgett, Alexandra Olteanu, and Adam Trischler. 2024. " It was 80% me, 20% AI": Seeking Authenticity in Co-Writing with Large Language Models. *arXiv preprint arXiv:2411.13032* (2024).

[22] iA. 2023. https://ia.net/topics/ia-writer-7. https://ia.net/topics/ia-writer-7.

[23] ibm. 2025. Carbon for AI. https://carbondesignsystem.com/guidelines/carbon-for-ai/.

[24] California Legislative Information. 2024. SB-942 California AI Transparency Act. https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB942.

[25] Nikhita Joshi and Daniel Vogel. 2024. Writing with AI Lowers Psychological Ownership, but Longer Prompts Can Help. *arXiv preprint arXiv:2404.03108* (2024).

[26] Heidy Khlaaf, Pamela Mishkin, Joshua Achiam, Gretchen Krueger, and Miles Brundage. 2022. A Hazard Analysis Framework for Code Synthesis Large Language Models. http://arxiv.org/abs/2207.14157 arXiv:2207.14157 [cs].

[27] Donal Khosrowi, Finola Finn, and Elinor Clark. 2023. Diffusing the creator: Attributing credit for generative AI outputs. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. 890–900.

[28] Anastasia Kuzminykh and Jessica R Cauchard. 2020. Be mine: Contextualization of ownership research in HCI. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–9.

[29] Mina Lee, Katy Ilonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thiemo Wambsganss, David Zhou, Emad A Alghamdi, et al. 2024. A Design Space for Intelligent and Interactive Writing Assistants. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–35.

[30] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems* 33 (2020), 9459–9474.

[31] Uwe Messer. 2024. Co-creating art with generative artificial intelligence: Implications for artworks and artists. *Computers in Human Behavior: Artificial Humans* 2, 1 (2024), 100056.

[32] Michael Muller, Anna Kantosalo, Mary Lou Maher, Charles Patrick Martin, and Greg Walsh. 2024. GenAICHI 2024: Generative AI and HCI at CHI 2024. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[33] Springer Nature. Date unknown. Artificial Intelligence (AI). https://www.nature.com/nature-portfolio/editorial-policies/ai.

[34] International Committee of Medical Journal Editors. Date unknown. Defining the Role of Authors and Contributors. https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#four.

[35] U.S. Copyright Office. 2025. Copyright and Artificial Intelligence. https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf.

[36] European Parliament. 2023. EU AI Act: first regulation on artificial intelligence. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

[37] Jon L Pierce, Tatiana Kostova, and Kurt T Dirks. 2001. Toward a theory of psychological ownership in organizations. *Academy of management review* 26, 2 (2001), 298–310.

[38] Jon L Pierce, Tatiana Kostova, and Kurt T Dirks. 2003. The state of psychological ownership: Integrating and extending a century of research. *Review of general psychology* 7, 1 (2003), 84–107.

[39] Jeba Rezwana and Mary Lou Maher. 2023. User Perspectives of the Ethical Dilemmas of Ownership, Accountability, Leadership in Human-AI Co-Creation. (2023).

[40] Advait Sarkar. 2023. Exploring perspectives on the impact of Artificial Intelligence on the Creativity of Knowledge Work: Beyond Mechanised Plagiarism and Stochastic parrots. In *Proceedings of the 2nd Annual Meeting of the Symposium on Human-Computer Interaction for Work*. 1–17.

[41] Alejo Jose G Sison, Marco Tulio Daza, Roberto Gozalo-Brizuela, and Eduardo C Garrido-Merchán. 2024. ChatGPT: More than a "weapon of mass deception" ethical challenges and responses from the human-centered artificial intelligence (HCAI) perspective. *International Journal of Human–Computer Interaction* 40, 17 (2024), 4853–4872.

[42] Justin D Weisz, Shraddha Kumar, Michael Muller, Karen-Ellen Browne, Arielle Goldberg, Ellice Heintze, and Shagun Bajpai. 2024. Examining the Use and Impact of an AI Code Assistant on Developer Productivity and Experience in the Enterprise. *arXiv preprint arXiv:2412.06603* (2024).

[43] Wiley. 2023. *Generative Artificial Intelligence Guidelines for Authors*. Retrieved 15-Aug-2024 from https://onlinelibrary.wiley.com/pb-assets/assets/15405885/Generative%20AI%20Policy_September%202023-1695231878293.pdf

[44] Chloe Wittenberg, Ziv Epstein, Adam J Berinsky, and David G Rand. 2024. Labeling AI-Generated Content: Promises, Perils, and Future Directions. (2024).

[45] Yuxin Xu, Mengqiu Cheng, and Anastasia Kuzminykh. 2024. What Makes It Mine? Exploring Psychological Ownership over Human-AI Co-Creations. In *Graphics Interface 2024 Second Deadline*.

[46] Catherine Yeh, Gonzalo Ramos, Rachel Ng, Andy Huntington, and Richard Banks. 2024. GhostWriter: Augmenting Collaborative Human-AI Writing Experiences Through Personalization and Agency. *arXiv preprint arXiv:2402.08855* (2024).

[47] Burak Yetiştiren, Işık Özsoy, Miray Ayerdem, and Eray Tüzün. 2023. Evaluating the code quality of ai-assisted code generation tools: An empirical study on github copilot, amazon codewhisperer, and chatgpt. *arXiv preprint arXiv:2304.10778* (2023).

[48] Jiawei Zhou, Yixuan Zhang, Qianni Luo, Andrea G Parker, and Munmun De Choudhury. 2023. Synthetic Lies: Understanding AI-Generated Misinformation and Evaluating Algorithmic and Human Solutions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3544548.3581318

[49] Christopher T Zirpoli. 2023. Generative artificial intelligence and copyright law. *CRS Legal Sidebar* (23 Feb 2023).