Exploring Human-AI Co-Creativity under Human Control: Framing, Reframing, Brainstorming, and Future Challenges (workshop keynote)

Michael Muller

IBM Research, Cambridge MA USA

Abstract

Generative AI is increasingly important in many human activities. In this keynote, I describe a series of experiments in generative AI, using a highly-conversational user interface (UI) to large language models (LLMs). I specify multiple strategies through which the conversational paradigm can be used to achieve ethical outcomes that promote AI humility, AI brevity, human control, and co-creativity among human and AI.

Keywords

HCAI (Human Centered AI), Generative AI, LLMs (Large Language Models), Conversational AI, Co-creativity, Distributed creativity, Ethical AI practices

Generative AI has the potential to support human creativity. Over the past 60 years, scholars in Human Centered AI (HCAI) have proposed diverse models of how systems for human-computer co-creativity can be designed. In 1961, Rhodes proposed a "4Ps" model in which "A Person engages in a computerized Process to make a Product in an environment (Press)" [1]. In this early view, the computer was primarily a tool. Almost 60 years later, Kantosalo and Takala made the most recent update in their "5Cs" model: "A Collective (a Human and an AI) Collaborate to make a Contribution for a Community in a Context" [2].

Within the framework of Kantosalo and Takala, and Glăveanu's work in distributed creativity [3], we investigate how one or more humans can collaborate with an AI agent to co-create their contributions, while maintaining human control over process and outcomes. In earlier work, we had developed a conversational UI to large language models (LLMs) for software engineering tasks [4]. In a quantitative experiment, Ross and colleagues showed that a well-tuned UI could make a back-end LLM behave in a humble, polite, and highly supportive way [5].

We re-used this architecture to explore creativity and co-creativity opportunities, through careful prompt-engineering. After surveying human-human co-creativity strategies [6, 7, 8], we conducted three informal experiments [9] using the well-known strategy of framing a problem with a productive representation [10, 11]. Next, in a fourth experiment, we explored the more powerful concept of reframing a problem [12] after the initial frame had been found to be flawed or insufficient in some way [13]. The conversational UI allowed the human to control how

EthAIcs workshop, IJCAI 2023, August 21, 2023, Macao

michael<underscore-character>muller@us.ibm.com (M. Muller)

https://research.ibm.com/people/michael-muller (M. Muller)

0000-0001-7860-163X (M. Muller)

the conversation developed, and which aspects of the conversation would be preserved in an analogy-based design.

In our fifth, unpublished experiment, we moved from specialist methods to the more generally-adopted processes of brainstorming. A human was able to guide the UI+LLM in exercises based on divergent-thinking, convergent-thinking, summarization, and structured organization/re-organization of outcomes. The user structured the activity, and the AI provided content. The user could question, critique, and reject certain content, and the AI could (when requested) provide alternatives to that content. The user could also tell the AI to re-organize its proposed higher-level structuring of content.

While these initial experiments were successful, we were only able to implement a dialog between one human and one AI. Our next projects will use a specialized environment in which multiple humans can interact with the UI+LLM configuration, with preservation of each human's identity, thus adding aspects of Mutual Theory of Mind [14, 15, 16] to the co-creative exercises. Based on the *Library of Mixed-Initiative Creative Interfaces* [17, 18], we are applying mixed-initiative models [19] to human-AI dialogs. After that, we hope to revisit multi-agent symbiotic cognitive computing architectures for a richer configuration of multiple humans and multiple AI agents [20].

Throughout this work, we have focused on principles of IBM's Augmented Human Intelligence, in which AI is used to support and extend the work of humans – not to replace humans [21]. Following a recent debate [22], we label all AI conversational turns with an "AI" or "APP" marker – i.e., we explicitly avoid any so-called Turing test confusions about who or what is speaking or acting. We maintain human control of both process and outcomes.

As we showed in a recent CHIWORK paper [23], these are design choices. It is possible to create interactive AI solutions that channel and control the work of humans [24]. Recent work by many researchers have documented the potential and actual harms of such systems (e.g., [25, 26, 27]). We make a different choice: We design for AI applications that support, educate, and enable human abilities and human agency.

Michael Muller works at IBM Research in the role of Senior Research Scientist. His work takes place at the intersection of human-computer interaction, AI, and social justice. Michael is known for early work in participatory design, and for co-proposing and co-leading the CHI conference program subcommittee on Critical and Sustainable Computing and Social Justice. His more recent work has explored possible futures for human interactions with generative AI applications. Michael has led and participated in mentorship programs for students and early-career scholars at multiple ACM SIGCHI conferences, including the CHI Early Career Symposium and the CSCW Student Reviewer Mentoring program. ACM recognizes Michael as a Distinguished Scientist. Michael co-chairs the SIGCHI CARES committee, and is a member of the SIGCHI Research Ethics committee. He holds membership in Fempower.tech and AccessSIGCHI, and has begun a term of service for the US National Academies of Science, Engineering and Medicine on the Board on Human-Systems Integration (BOHSI).

References

[1] M. Rhodes, An analysis of creativity, The Phi Delta Kappan 42 (1961) 305–310.

[2] A. Kantosalo, T. Takala, Five C's for human-computer co-creativity-an update on classical creativity perspectives., in: ICCC, 2020, pp. 17–24.

[3] V. P. Glăveanu, Distributed creativity: What is it?, Springer, 2014.

[4] S. I. Ross, F. Martinez, S. Houde, M. Muller, J. D. Weisz, The programmer's assistant: Conversational interaction with a large language model for software development, in: Proceedings of the 28th International Conference on Intelligent User Interfaces, 2023, pp. 491–514.

[5] S. I. Ross, M. Muller, F. Martinez, S. Houde, J. D. Weisz, A case study in engineering a conversational programming assistant's persona, arXiv preprint arXiv:2301.10016 (2023).

[6] M. Mose Biskjaer, P. Dalsgaard, K. Halskov, Understanding creativity methods in design, in: Proceedings of the 2017 conference on designing interactive systems, 2017, pp. 839–851.

[7] J. Schleith, M. Norkute, M. Mikhail, D. Tsar, Cognitive strategy prompts: Creativity triggers for human centered ai opportunity detection, in: Creativity and Cognition, 2022, pp. 29–37.

[8] Z. Wu, D. Ji, K. Yu, X. Zeng, D. Wu, M. Shidujaman, Ai creativity and the human-ai co-creation model, in: Human-Computer Interaction. Theory, Methods and Tools: Thematic Area, HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I 23, Springer, 2021, pp. 171–190.

[9] M. Muller, H. Candello, J. Weisz, Interactional co-creativity of human and ai in analogy-based design, in: International Conference on Computational Creativity, 2023.

[10] A. Liapis, G. N. Yannakakis, C. Alexopoulos, P. Lopes, Can computers foster human users' creativity? theory and praxis of mixed-initiative co-creativity (2016).

[11] C. Walsh, K. Chappell, A. Craft, A co-creativity theoretical framework to foster and evaluate the presence of wise humanising creativity in virtual learning environments (vles), Thinking skills and creativity 24 (2017) 228–241.

[12] N. Kelly, J. S. Gero, Creative systems that generate and explore, 2015. URL: https://api.semanticscholar.org/CorpusID:55272279.

[13] M. Muller, J. Weisz, Analogies-based design using a generative ai application: A play in three acts, 2023. URL: https://www.researchgate.net/publication/372060579_Analogies-based_Design_using_a_Generative_AI_Application_A_Play_in_Three_Acts.

[14] F. Cuzzolin, A. Morelli, B. Cirstea, B. J. Sahakian, Knowing me, knowing you: theory of mind in ai, Psychological medicine 50 (2020) 1057–1061.

[15] Q. Wang, K. Saha, E. Gregori, D. Joyner, A. Goel, Towards mutual theory of mind in human-ai interaction: How language reflects what students perceive about a virtual teaching assistant, in: Proceedings of the 2021 CHI conference on human factors in computing systems, 2021, pp. 1–14.

[16] A. F. Winfield, Experiments in artificial theory of mind: From safety to story-telling, Frontiers in Robotics and AI 5 (2018) 75.

[17] S. Deterding, J. Hook, R. Fiebrink, M. Gillies, J. Gow, M. Akten, G. Smith, A. Liapis, K. Compton, Mixed-initiative creative interfaces, in: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, 2017, pp. 628–635.

[18] A. Spoto, N. Oleynik, Library of mixed-initiative creative interfaces, 2017. URL: http:

//mici.codingconduct.cc/.
[19] E. Horvitz, Principles of mixed-initiative user interfaces, in: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 1999, pp. 159–166.
[20] R. Farrell, J. Lenchner, J. Kephart, A. Webb, M. Muller, T. Erickson, D. Melville, R. Bellamy, D. Gruen, J. Connell, et al., Symbiotic cognitive computing, ai Magazine 37 (2016) 81–93.
[21] R. Mariani, F. Rossi, R. Cucchiara, M. Pavone, B. Simkin, A. Koene, J. Papenbrock, Trustworthy ai—part 1, Computer 56 (2023) 14–18.
[22] B. Shneiderman, M. Muller, On AI anthropomorphism. Medium 4 April (2023). URL: https://medium.com/human-centered-ai/on-ai-anthropomorphism-abff4cecc5ae.
[23] M. Muller, J. Weisz, Frameworks for collaborating humans and ais: Sequence and sociality in organizational applications, in: CHIWORK, 2022.
[24] M. Muller, J. He, J. D. Weisz, The trouble with ai-based workflows (2023). URL: https://www.researchgate.net/profile/Michael-Muller-12/publication/368242833_The_Trouble_with_AI-Based_Workflows/links/63dd67fc62d2a24f92f5a04d/The-Trouble-with-AI-Based-Workflows.pdf.
[25] N. Holten Møller, G. Neff, J. G. Simonsen, J. C. Villumsen, P. Bjørn, Can workplace tracking ever empower? collective sensemaking for the responsible use of sensor data at work, Proceedings of the ACM on human-computer interaction 5 (2021) 1–21.
[26] C. Lin, S. Margot Lindtner, Techniques of use: Confronting value systems of productivity, progress, and usefulness in computing and design, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 2021, pp. 1–16.
[27] R. Shelby, S. Rismani, K. Henne, A. Moon, N. Rostamzadeh, P. Nicholas, Y.-A. N'MAH, J. Gallegos, A. Smart, G. VIRK, Identifying sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction, arXiv preprint arXiv:2210.05791 (2022).