

CHAI-DT: Prompting Conversational Generative AI Agents to Actively Participate in Co-Creation

BRANDON HARWOOD

CCS Concepts: • **Human-centered computing** → **Collaborative interaction**.

Additional Key Words and Phrases: HCI, Artificial Intelligence, Generative AI, Co-Creation, Design Thinking,

ACM Reference Format:

Brandon Harwood. 2023. CHAI-DT: Prompting Conversational Generative AI Agents to Actively Participate in Co-Creation. 1, 1 (April 2023), 5 pages.

1 INTRODUCTION

In recent years, advancements in AI technology and accessibility have enabled the use of generative AI models such as ChatGPT [7] as versatile tools for productivity, research, and creative application across a wide range of contexts and domains [6] as well as the ability to successfully contribute to UX design functions. [4] While these use-cases are often examined within the context of one-on-one human + AI interactions, little research explores how participants and facilitators might utilize generative creative agents effectively within group focused co-creative frameworks such as Design Thinking, which this paper posits could bring value to problem solving and ideation in business innovation and co-creation contexts. [5]

This paper seeks to explore the space of group + AI co-creative partnerships in IBM, and propose a prompting technique for conversational agents (i.e. ChatGPT) which employ methods inspired by traditional 'human-to-human' facilitation and instruction typically seen in Design Thinking, a co-creative framework which brings multidisciplinary participants together in groups to ideate user-focused business solutions. [2] Through experiments using this prompting technique, I have gathered evidence that conversational generative transformers like ChatGPT have the capability to contribute context-specific, useful, and creative input into Design Thinking activities. I will also discuss the frameworks potential value in fostering 'human/s + AI' co-creative partnership within Kantosalo's 5Cs Framework for Human-Computer Co-Creativity [3] and IBM's Design Thinking Framework, described in [2]. [5]

Finally, I consider potential benefits, limitations, and risks associated with using generative AI in co-creative ideation, as well as speculate on opportunities for future research into how we might mitigate risks and measure effectiveness of 'human/s + AI' co-creative partnership within Design Thinking contexts. By addressing these limitations and exploring effective ways to partner with creative AI agents in group settings, I hope to set a path forward in advancing our understanding and use of AI systems/agents as helpful, effective partners and participants in co-creative pursuits.

Author's address: Brandon Harwood.

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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2023 Association for Computing Machinery.

Manuscript submitted to ACM

Manuscript submitted to ACM

2 BACKGROUND

Human-computer co-creativity is described by Kantosalo et. al. as “the interactions within a human–computer *collective*, the collective’s *collaboration* process and creative *contributions* to a *community*, all situated within a rich *context*.” [3] This is contrasted with, but closely aligns to traditional human co-creative frameworks such as Design Thinking (DT), differentiated by the explicit account for the contribution of computational agents in co-creation. DT, as it is compared to Kantosalo’s model, can be described as employing a *collective* group of individuals to utilize their “individual subjective” *context* in reference to a *community* (e.g. user or user-group) in a narrowly-scoped “idealized objective” *context* defined in the setting of the DT session (e.g. problem space, process, etc.), so they may *collaborate* with each other through the Design Thinking framework described in [2], and ultimately *contribute* creative artifacts (e.g. ideas, observations etc.).

Kantosalo notes that “in a successful creative collaboration, the collective benefits from the profound communication and sharing of contributions within the collective.” [3] This framework is differentiated from DT by specifying the “*human-initiative vs. computational-initiative*” dichotomy, but this measure of success can also be used to describe a successful DT session, the only difference being that humans are the only participants with creative initiative in traditional DT. However, with recent research into the ability of generative AI agents to contribute to design spaces creatively through conversation [4], it is suggested that we are able to build a system respective of human vs. computational initiative, and integrate creative AI agents into group ideation spaces as active participants.

Drawing on this background, I intend to explore the potential of integrating useful, creative AI agents into Design Thinking through a novel prompting framework, and in doing so, explore how we might be able to build environments which foster greater creative output in group-focused co-creation settings through mixed human-computer initiative.

3 METHODS

OpenAI’s ChatGPT [7] was chosen as the testing grounds for experimentation due to it’s availability and relatively advanced capabilities, as well as the conversational interface, which provide an environment that enables both DT facilitators and participants to engage with the model during live ideation. During initial testing it became apparent that a loose conversational style for ‘facilitating’ the model through DT exercises resulted in inconsistent output quality, requiring multiple back-and-forth explanations or re-framing of the prompts. This is sufficient (and sometimes ideal) for one-on-one ideation, but could be distracting in a live session. Through observing the outputs of these loose interactions and iterative prompt-design, I have identified 6 criteria or “sub-prompts” that result in better, more structured co-creative output from ChatGPT, and developed a prompting method that balances consistency and efficiency.

The method I am introducing as the **Collaborative Human-AI Design Thinking (CHAI-DT) Prompting Method** is a framework for building a single prompt that initiates the interaction, explains the activity and instructions, and provides necessary context through a composition of six sub-prompts. The first four sub-prompts are ‘static instructions’ which introduce the structure and purpose of the chosen activity in the same way we might introduce it to human participants, and does not change between interactions with the agent. The final two are ‘variable instructions’, which are dependent on the session context and interaction needs from the facilitator and participants.

Static Instructions:

- Introduction: “We are conducting a(n) '[Activity Name]' Design Thinking exercise.”
- Definition: Explanation of the exercise, it’s intended outcome, and why it is useful for co-creation.
- Example(s): One or multiple examples of an ideal end-state artifact.
- Instructions: Step-by Step instructions for the activity.

Table 1. ChatGPT's Responses to the 3 "Hills" DT Exercise Steps

"Who" (Potential Users)	"What" (Enablers)	"Wow" (Impact)
Retail store managers	Accurately predict sales trends	Dramatically reduce stockouts and overstocking, resulting in increased sales and profitability
Inventory managers	Optimize inventory levels in real-time	Improve customer satisfaction by ensuring products are always in stock
Supply chain managers	Identify underperforming products	Increase efficiency and productivity for store and inventory managers
Sales associates	Identify overstocked products	Reduce waste and optimize resource utilization
Customers (indirectly impacted by the system)	Determine reorder quantities and timing	Provide a competitive advantage in the retail industry through advanced data analytics and artificial intelligence
Executives/decision-makers at RetailInc	Monitor stock levels and alert managers when stock falls below a certain threshold	Ensure accuracy and reliability of sales and inventory data, leading to better decision-making
	Provide insights into customer demand and behavior	Enhance the overall shopping experience for customers through better inventory management and product availability
	Generate automated reports and analytics for inventory and sales data	Enable RetailInc to respond quickly to changing market trends and customer demands
	Minimize stockouts and overstocking	Foster a culture of innovation and continuous improvement at RetailInc
	Enable data-driven decision-making for inventory management	

Variable Instructions:

- Context: Appropriate context to inform the interaction. (i.e. Why are we co-creating? Who is participating? What background information can we provide?)
- Execute: Prompt the model to perform the activity, either in its entirety or gradually (i.e. "perform the entire Hills exercise.", or "perform step 1.").

4 RESULTS

Table 1, above, provides a record of outputs from ChatGPT (Feb 13 Release [8]), prompted with the CHAI-DT framework to perform a "Hills" DT exercise. Hills are "concise statements of the goals we aim to help our users accomplish" [2], and are built by participants observing and reflecting on a narrow context to individually contribute artifacts (often in the form of sticky notes, one idea written on each) within 3 criteria; **"Who"** (relevant people within the context), **"What"** (enablers provided to them), and **"Wow"** (value differentiator/impact). Typically, participants cluster the individual artifacts into thematic groups, and reflect on emergent themes to build ideas collaboratively. This exercise, performed one step at a time, tests the models capability to reflect on the context provided, as well as the groups/its own previous artifacts, to generate *observational* artifacts to the group e.g. users derived from the context, as well as *creative* artifacts such as assumed problems they face and new ideas valuable to them. The Context sub-prompt provided in this test was a fake scenario generated by ChatGPT (Jan 9 Release [8]), solely intended for testing different DT activities:

RetailInc, a large retail chain, and IBM have come together to improve their inventory management system. The current system is outdated and often leads to stockouts and overstocking, resulting in lost sales and wasted resources. IBM brings their expertise in data analytics and artificial intelligence to the partnership, while RetailInc provides their extensive data on sales and inventory levels. Together, they aim to create a new system that can accurately predict sales trends and optimize inventory levels in real-time, leading to increased efficiency and profits for the retail chain.

Interestingly, after ChatGPT completed each step of the Hills exercise, it also made a point to add disclaimers: "Note: These are just some potential users [or "enablements that the system could provide" or "potential market values or differentiators that the system could provide"], and the team may need to further refine the list based on their research and understanding of the user needs.". This behavior frames the content generated as *intended* for collaboration with others, and in doing so, aligns to Kantosalo's framework by creating an interaction between the Human/s + AI collective to *collaborate* through individual *contributions* for the *community* identified in the *context* provided.

While the Table 1 outputs are from a single one-on-one test, they are representative of typical, repeatable patterns seen in multiple tests. Given these results, we can see that ChatGPT was able to generate content relevant to the context of the session with reasonable quality, creativity, and usefulness to cooperative ideation. The models response was also in a manner that follows the instructions provided and can easily be incorporated into a live Design Thinking session.

5 DISCUSSION

Through the development of the CHAI-DT prompting framework, I have demonstrated that ChatGPT has the capability to provide useful, creative artifacts in live co-creation activities. This method of prompt-engineering is limited in that it is applied to the context of corporate co-creation often utilized for business value creation, and further research covering integration into other group-focused co-creative activities is needed. However, the implications for this mode of Human/s + AI co-creation are wide reaching and provide a space for future research into how people or groups of people collaborate, partner, and interact with generative AI agents to enhance collective creativity in other domains and contexts, as well as informs future attempts to design AI agents intended for collaboration, cooperation, and co-creation.

By introducing a real-world implementation for co-creative generative agents which align to Kantosalo's 5Cs framework, it is suggested that through similar or iterative methods, there is opportunity to further explore this space to build entirely new forms of creative collaboration, as well as environments or methods to enhance communication in ideation and co-creativity. It is also recommended that future research address potential risk of utilizing generative AI agents in these environments, covered well by Buschek et. al. such as bias, misinformation, conflict of responsibility, and potential exposure to private or inappropriate data [1], as well as the potential for participant distraction in-session, less useful or harmful ideation, misuse of the framework by malicious actors, or potential harm to the communities we are ideating for.

6 CONCLUSION

In this paper, I have introduced CHAI-DT, a novel approach to prompting conversational agents such as ChatGPT to engage in 'Human/s + AI' co-creative efforts, and have explored the potential of integrating creative generative AI agents within live Design Thinking, as well as the implications of this approach and some of the potential risks associated. I also suggest more research be done considering the risks, benefits, and potential future application or evolution of similar frameworks (both useful and harmful) to better understand the future this technology and the framework implies, so we may shape the use of these models in co-creative contexts to be useful, safe, and ultimately enhance the creative capabilities of people and communities.

REFERENCES

- [1] Daniel Buschek, Lukas Mecke, Florian Lehmann, and Hai Dang. 2021. Nine Potential Pitfalls when Designing Human-AI Co-Creative Systems. *ArXiv abs/2104.00358* (2021).
- [2] IBM Design. "n.d.". Enterprise Design Thinking. <https://www.ibm.com/design/approach/design-thinking/>
- [3] Anna Kantosalo and Tapio Takala. 2020. Five C's for Human-Computer Co-Creativity - An Update on Classical Creativity Perspectives. In *International Conference on Innovative Computing and Cloud Computing*.
- [4] A. Baki Kocaballi. 2023. Conversational AI-Powered Design: ChatGPT as Designer, User, and Product. <https://doi.org/10.48550/ARXIV.2302.07406>
- [5] Jieun Kwon, Younghyun Choi, and Yura Hwang. 2021. Enterprise Design Thinking: An Investigation on User-Centered Design Processes in Large Corporations. *Designs* 5, 3 (2021). <https://doi.org/10.3390/designs5030043>
- [6] Zhicheng Lin. 2023. Why and how to embrace AI such as ChatGPT in your academic life. <https://doi.org/10.31234/osf.io/sdx3j>
- [7] OpenAI. 2023. CHATGPT: Optimizing language models for dialogue. <https://openai.com/blog/chatgpt/>
- [8] Natalie Staudacher. 2023. ChatGPT — Release Notes. <https://help.openai.com/en/articles/6825453-chatgpt-release-notes>

Received 20 February 2023; accepted 9 March 2023