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

ANONYMOUS AUTHOR(S)

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

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

ACM Reference Format:

Anonymous Author(s). 2018. CHAI-DT: Prompting Conversational Generative AI Agents to Actively Participate in Co-creation. 1, 1 (February 2018), 5 pages. https://doi.org/XXXXXXX.XXXXXXX

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 SWCorp, 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, we have gathered evidence that conversational generative transformers like ChatGPT have the capability to contribute context-specific, useful, and creative input into Design Thinking activities. We 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 SWCorp's Design Thinking Framework, described in [2]. [5]

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

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 concept is contrasted with, but closely aligns to traditional human co-creative frameworks such as Design Thinking (DT), the most important differentiator being 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 such as ideas.

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 actors 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], we are able to build a framework respective of human vs. computational initiative to integrate creative agents into the Design Thinking space.

Drawing on this background, this paper intends to explore the potential of integrating creative AI agents within Design Thinking, and by doing so, explore how we might be able to build environments which fosters 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 iterative experimentation, we developed a prompting method that balances consistency and efficiency, and can be introduced in a single prompt to quickly start the interaction.

This prompt we are introducing as the **Collaborative Human-AI Design Thinking Prompting Method (CHAI-DT)** is a sum total of six sub-prompts, initiated at the beginning of the "conversation" with ChatGPT. The first four sub-prompts are 'static' in that they 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 sessions. The final two are 'variable' sub-prompts, which are dependent on the session context and interaction needs from the facilitator and participants. The sub-prompts are as follows:

Static Instructions:

- `Introduction`: "We are conducting a(n) '[Activity Name]' Design Thinking exercise."
- `Definition`: Explanation of what the exercise is and what it is for.
- `Example(s)`: One or multiple examples of an ideal end-state artifact.

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

"Who" (Potential Users)	"What" (Enablements)	"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	

- `Instructions`: Step-by Step instructions for the activity.

Variable Instructions:

- `Context`: Appropriate context for the session. Why are we co-creating? Who is participating? What background information can we provide?
- `Execute`: Prompt the model to perform the activity, either in it's entirety or one step at a time.

4 RESULTS

Table 1, above, provides a record of response outputs of ChatGPT using the above framework for the DT "Hills" exercise described in [2] as "concise statements of the goals we aim to help our users accomplish." Hill Statements are made up of 3 criteria which are collaboratively defined. These criteria are "Who" (users), "What" (the enablement) and "Wow" (the value differentiator or impact). This exercise was chosen because it tests whether the models is capable of generating logical content like probable user-types and tasks a scenario implies, as well as creative content that examines the needs of the users and provides potential value in context. The context sub-prompt provided was a fake scenario (generated by ChatGPT), as follows:

RetailInc, a large retail chain, and [SWCorp] 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.

[SWCorp] 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 a disclaimer: *"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."* suggesting that through this prompting technique, the model displays behavior that indicates the content generated is *intended* for collaboration with others.

While the outputs provided are from one session, they are representative of typical, repeatable patterns we have seen from multiple tests. Given these results, we conclude that ChatGPT was able to generate content specific to the context of the session, of reasonable quality and creativity, in a manner that is useful for 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 technique, we have demonstrated that ChatGPT has the capability to provide useful creative input in live ideation 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.

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 and co-creative environments to enhance communication around ideation and creativity. It is also recommended that future research address potential risk of utilizing generative co-creative AI agents in these contexts, 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 we 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. We suggest more research is done considering the risks, benefits, and potential future applications of similar frameworks (both useful and harmful) so that we may better understand the future this technology and the framework implies, so we may shape the future use of these models in co-creative contexts to be useful, safe, and ultumately 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/

Received 20 February 2023; revised 12 March 2009; accepted 5 June 2009