

Architecting Novel Interactions With Generative AI Models

Michael S. Bernstein
Stanford University
Stanford, CA, USA
msb@cs.stanford.edu

Joon Sung Park
Stanford University
Stanford, CA, USA
joonspk@stanford.edu

Meredith Ringel Morris
Google DeepMind
Seattle, WA, USA
merrie@google.com

Saleema Amershi
Microsoft Research
Seattle, WA, USA
samershi@microsoft.com

Lydia Chilton
Columbia University
New York, NY, USA
chilton@cs.columbia.edu

Mitchell L. Gordon
MIT and University of Washington
Cambridge, MA, USA
mgord@cs.stanford.edu

ABSTRACT

The new generation of generative AI models offers interactive opportunities that may fulfill long-standing aspirations in human-computer interaction and open doors to new forms of interaction that we have yet to imagine. The UIST community has a unique vantage point that can lead to critical contributions in envisioning a future of interactive computing that appropriately leverages the power of these new generative AI models. However, we are only just beginning to understand the research area that exists at the intersection of interaction and generative AI. By bringing together members of the UIST community interested in this intersection, we seek to initiate discussions on the potential of generative AI in architecting new forms of interactions. Key topics of interest include the exploration of novel categories of interactions made possible by generative AI, the development of methods for enabling more powerful and direct user control of generative AI, and the identification of model and architecture requirements for generative AI in interaction literature. The workshop will foster community building and produce concrete deliverables, including a research agenda, model/architecture requirements, and a simulated debate generated by a generative agent architecture.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI); Interaction design**; • **Computing methodologies** → **Artificial intelligence; Natural language generation.**

KEYWORDS

HCI, AI, Generative AI, Human-AI Interaction

ACM Reference Format:

Michael S. Bernstein, Joon Sung Park, Meredith Ringel Morris, Saleema Amershi, Lydia Chilton, and Mitchell L. Gordon. 2018. Architecting Novel Interactions With Generative AI Models. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 3 pages. <https://doi.org/XXXXXXX.XXXXXXX>

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.

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00
<https://doi.org/XXXXXXX.XXXXXXX>

1 INTRODUCTION

Generative AI has made waves. Models ranging from ChatGPT [4], LaMDA [7], Stable Diffusion [3], and Make-A-Video [6] are now familiar to many and used for generating content across various modalities on the web. Their recent prominence owes much to their seemingly emergent capabilities that encode rich context about us and the world that we inhabit, and synthesize that context in non-trivial manners when prompted to do so in natural language [8]. This richness of the new generation of generative AI models and their low entry threshold offer significant interactive opportunities between humans and computational tools [1, 2]. We posit that some of these opportunities will fulfill long-standing aspirations in human-computer interaction, such as the creation of sophisticated model human processors, others will open doors to new forms of interaction that we have yet to imagine, and still others will create both frustrations and harms to people.

In the face of these opportunities, and the technical and societal challenges they inevitably bring, we have to ask: how ought we to shape a future of interactive computing powered by generative AI to have a lasting positive impact in augmenting human capabilities? Various academic and open-source communities have begun exploring the intersection between human factors and generative AI, with a focus on benchmarking the out-of-the-box capabilities of these new models. However, it is the UIST and broader HCI communities, who bear a long tradition of formulating innovative and sometimes unexpected modes of interaction with intelligent systems, who have a unique advantage in envisioning how generative AI models can empower new forms of human-computer interaction.

This workshop aims to bring together the UIST community members who are interested in using generative AI to develop innovative forms of interaction. Our objective is to initiate a discussion that will establish the groundwork for future contributions in this field. We specifically seek researchers who are exploring the following:

- (1) Usage of generative AI: What are the novel categories of interactions that can be enabled by generative AI, and how do they contribute to the existing literature on interaction?
- (2) Controllability: How can we enable more powerful and direct user control of generative AI models and their outputs?
- (3) Model and architecture requirements: What assumptions or capabilities of generative AI models and their associated architectures do we need to consider in order to formulate new forms of interaction and establish reproducible evaluation techniques?

In addition to community building and the proposed activities, we will produce three concrete deliverables: 1) a research agenda for the field, 2) a set of model/architecture requirements for generative AI models used in interaction literature, and 3) a simulated debate between the participants on the aforementioned topics generated with the generative agent architecture [5].

2 WORKSHOP PLANS AND GOALS

Our one-day workshop will include in-person activities on the day of the workshop, as well as remote activities conducted in preparation for the workshop.

2.1 Preparatory Activities

2.1.1 Application form. Potential participants will fill out an online form where they will be asked to provide links to their past or current work related to the topics addressed in the workshop, or to write a paragraph explaining their interests and current work relevant to the workshop. The selection process will be based on their unique expertise at this intersection and the diverse perspectives they can bring to the workshop. We will advertise the opportunity for participation not only within the UIST community but also within broader AI venues.

2.1.2 Survey. The participants will be asked to fill out an optional 30-minute online survey that explores their positions on the research topics to be discussed during the workshop. The organizers will collect the survey responses in order to create a generative agent [5] as part of one of the workshop activities that argues on behalf of each participant, reflecting their perspectives on the field. These generative agents will be deployed in a simulated debate, where they will act as proxy debaters representing their respective participants. The simulated debate will be played during a workshop session, serving as a catalyst for discussions on the current potentials and limitations of generative AI techniques.

2.2 Workshop Activities

2.2.1 Lightning talks. The first morning session of the workshop will be a 2-minute lightning talk from each participant. They will be asked to introduce themselves and present their interests or a project that they are currently working on. During these talks, participants may discuss a range of topics, such as interactions and systems they have constructed using generative AI, the challenges they encountered, or ideas for future research. To maintain momentum and engagement, we will adopt an ignite-style format, where slides will automatically advance every 2 minutes.

2.2.2 Morning breakout session. The second morning session will consist of a breakout session, where participants will form subgroups focusing on two of the topics mentioned in the introduction. They will explore interesting opportunities and challenges within each area. Here are potential aspects that each subgroup may cover:

- (1) Reciprocal impact [2]:
 - What are the categories of interactive applications that could benefit from generative AI?
 - How might interactive applications lead to changes in the architecture of generative AI models?
- (2) Controllability

- How can we ensure that generative AI models accurately reflect the intentions of human users?
- How can we create intuitive interfaces that empower users to control and guide the output of these models?

At the conclusion of the breakout session, we will reconvene as a whole group to share the discussions that took place within each of the subgroups. Building upon the insights gained from the subgroup discussions, we will explore the technical model and architecture requirements necessary for creating the desired interactions and their evaluations, striving to address key questions such as: How do different tasks interact with model errors? What level of model latency is deemed acceptable for various tasks? In which areas is it crucial to have replicable model output?

2.2.3 Afternoon debate. The afternoon session will feature a debate involving the generative agents acting as proxies of participants. While we hope this will be humorous and enjoyable in its own right, the hope is that hands-on experience trying to shape your own generative agent to represent you correctly will highlight the challenges with control of generative models. This simulation will delve into the opportunities, as well as the technical and societal challenges associated with building interactions using generative AI. The organizers will create these agents based on the participants' survey responses and pre-simulate the debate. During the simulation, participants will be encouraged to read aloud their generative agents' lines, and try to adjust the prompts given to their model in order to improve its debate performance. Subsequently, we will reconvene as a group to engage in a discussion centered around the interactive opportunities and challenges that the simulated debate has stimulated.

2.2.4 Research agenda. In the penultimate session of the workshop, we will collaborate to create a research agenda that draws inspiration from the workshop. This agenda will encompass various topics, including identifying immediate research questions and long-term goals for the field. We will deliberate on the direction in which we are heading and strategize on how to reach our goals. We will also explore challenges that need to be addressed and determine which, if resolved, would have transformative effects on our field. Finally, we will examine the temporary attributes of current generative AI models that are unlikely to stand the test of time.

2.2.5 Keynote. The final session will feature a keynote from Will Wright, renowned game designer and creator of The Sims and SimCity, and Lauren Elliott, co-designer of the Where In The World Is Carmen Sandiego series of games. Leveraging their expertise in designing immersive open worlds, Wright and Elliott will explore the practical implications of incorporating generative AI into interactive experiences. They will discuss how these AI models can enhance user engagement, foster creativity, and facilitate collaborative storytelling. The speakers have confirmed their attendance.

2.3 Follow-up

We will also include summaries or opinion pieces from the participants on our workshop website.

REFERENCES

- [1] Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, and et al. 2022. On the Opportunities and Risks of Foundation Models. arXiv:2108.07258 [cs.LG]
- [2] Meredith Ringel Morris, Carrie J. Cai, Jess Holbrook, Chinmay Kulkarni, and Michael Terry. 2023. The Design Space of Generative Models. arXiv:2304.10547 [cs.AI]
- [3] Stable Diffusion Online. 2023. *Stable Diffusion*. <https://stablediffusionweb.com/>
- [4] OpenAI. 2023. ChatGPT – Release Notes. <https://web.archive.org/web/20230314123821/https://openai.com/blog/chatgpt-release-notes/>.
- [5] Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. arXiv:2304.03442 [cs.HC]
- [6] Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, Songyang Zhang, Qiyuan Hu, Harry Yang, Oron Ashual, Oran Gafni, Devi Parikh, Sonal Gupta, and Yaniv Taigman. 2022. Make-A-Video: Text-to-Video Generation without Text-Video Data. arXiv:2209.14792 [cs.CV]
- [7] Romal Thoppilan, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin, Taylor Bos, Leslie Baker, Yu Du, YaGuang Li, Hongrae Lee, Huaixiu Steven Zheng, Amin Ghafouri, Marcelo Menegali, Yanping Huang, Maxim Krikun, Dmitry Lepikhin, James Qin, Dehao Chen, Yuanzhong Xu, Zhifeng Chen, Adam Roberts, Maarten Bosma, Yanqi Zhou, Chung-Ching Chang, Igor Krivokon, Will Rusch, Marc Pickett, Kathleen S. Meier-Hellstern, Meredith Ringel Morris, Tulsee Doshi, Renelito Delos Santos, Toju Duke, Johnny Soraker, Ben Zevenbergen, Vinodkumar Prabhakaran, Mark Diaz, Ben Hutchinson, Kristen Olson, Alejandra Molina, Erin Hoffman-John, Josh Lee, Lora Aroyo, Ravi Rajakumar, Alena Butryna, Matthew Lamm, Viktoriya Kuzmina, Joe Fenton, Aaron Cohen, Rachel Bernstein, Ray Kurzweil, Blaise Agüera y Arcas, Claire Cui, Marian Croak, Ed H. Chi, and Quoc Le. 2022. LaMDA: Language Models for Dialog Applications. *CoRR* abs/2201.08239 (2022). arXiv:2201.08239 <https://arxiv.org/abs/2201.08239>
- [8] Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, Ed H. Chi, Tatsunori Hashimoto, Oriol Vinyals, Percy Liang, Jeff Dean, and William Fedus. 2022. Emergent Abilities of Large Language Models. arXiv:2206.07682 [cs.CL]