Published: 06 October 2025

Citation in BibTeX format

CHItaly 2025: CHItaly 2025: 16th Biannual Conference of the Italian SIGCHI Chapter
October 6 - 10, 2025
Salerno, Italy

SHORT-PAPER

Prompting Realities: Exploring the Potentials of Prompting for Tangible Artifacts

MAHAN MEHRVARZ, Delft University of Technology, Delft, Zuid-Holland, Netherlands

DAVE MURRAY-RUST, Delft University of Technology, Delft, Zuid-Holland, Netherlands

HIMANSHU VERMA, Delft University of Technology, Delft, Zuid-Holland, Netherlands

CHItaly '25: Proceedings of the 16th Biannual Conference of the Italian SIGCHI Chapter (October 2025)
https://doi.org/10.1145/3750069.3750089
ISBN: 9798400721021

Prompting Realities: Exploring the Potentials of Prompting for Tangible Artifacts

Mahan Mehrvarz
AI Futures Lab
Delft University of Technology
Delft, Netherlands
m.mehrvarz@tudelft.nl

Dave Murray-Rust
AI Futures Lab
Delft University of Technology
Delft, Netherlands
d.s.murray-rust@tudelft.nl

Himanshu Verma
Knowledge and Intelligence Design
Delft University of Technology
Delft, Netherlands
h.verma@tudelft.nl

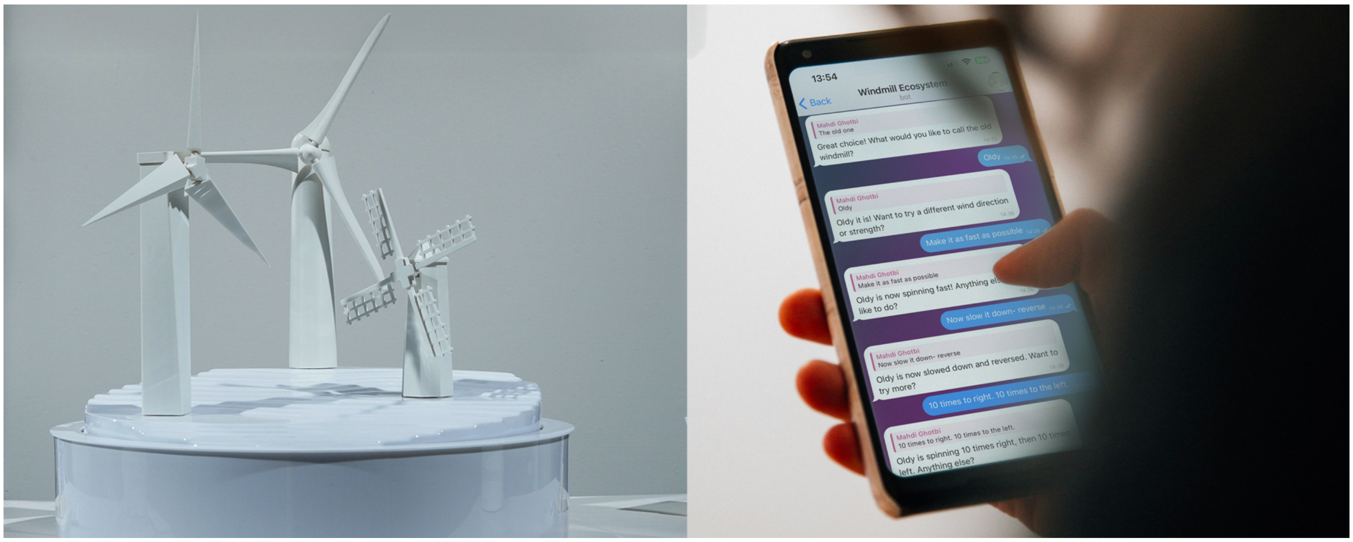

Figure 1: Users use the chat interface to spin the windmills at different speeds and in different directions.

Abstract

Designing meaningful tangible and embodied interactions remains challenging due to their situated nature, complex user needs, and the limited programming skills of many users as well as designers. We developed an interaction model where users and LLMs co-perform tangible actions through prompt engineering beyond deterministic logic of commercial smart systems. In this model AI systems interpret natural language descriptions of environmental context, internalize technical functionalities and spatial cues, and translate these into tangible actions. We encapsulated the interaction model within a LLM-enabled tangible artifact as a HCI provotype and conducted an initial exploratory study around it. Our preliminary findings point to opportunities in refinement and reappropriation of such systems over the use period as well as challenges in adapting deictic spatial references.

CHItaly 2025, Salerno, Italy

ACM ISBN 979-8-4007-2102-1/25/10
https://doi.org/10.1145/3750069.3750089

CCS Concepts

• **Human-centered computing** → *Interaction design theory, concepts and paradigms*; **Natural language interfaces**; **Natural language interfaces**; • **Computer systems organization** → Embedded software.

Keywords

Tangible interaction, embedded interaction, large language models, prompting, conversational AI, Human-AI collaboration

ACM Reference Format:
Mahan Mehrvarz, Dave Murray-Rust, and Himanshu Verma. 2025. Prompting Realities: Exploring the Potentials of Prompting for Tangible Artifacts. In *CHItaly 2025: 16th Biannual Conference of the Italian SIGCHI Chapter (CHItaly 2025), October 06–10, 2025, Salerno, Italy*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3750069.3750089

1 Introduction and background

The rapid adoption of Generative AI, particularly following the emergence of tools like ChatGPT, has brought attention to the possibilities for integration of Large Language Models (LLMs) into the fabric of everyday life. In particular, by offering naturalistic interaction through *prompting*, [3], LLMs can serve as interfaces to computational systems, being given access to web browsers [10, 18], API calls [23, 28] or command lines [31] to allow users to describe

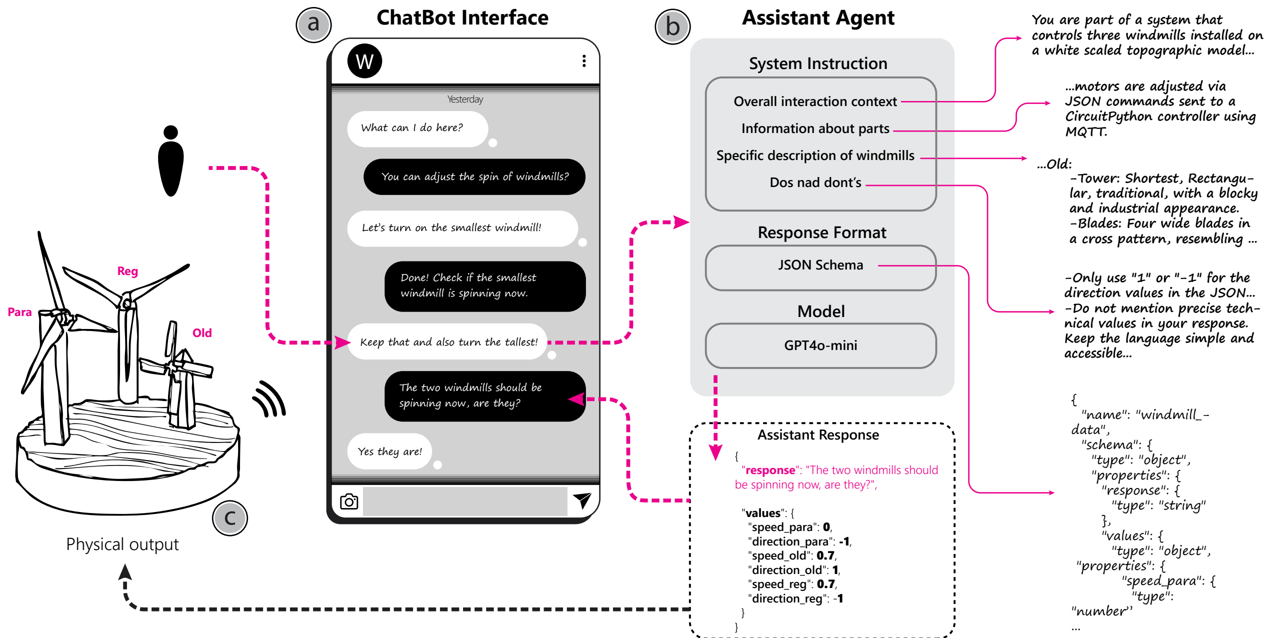

Figure 2: The system consists of three primary components: (a) a chat interface built through a Telegram bot that captures user input and displays system responses natural language, (b) an AI assistant that has a system instruction describing the setup as well as an output structure of the JSON commands it should produce; and (c) a diorama of three windmills powered by a electronic microcontroller with motors to simulate spinning which can receive JSON commands (over MQTT) to control its function.

what they would like to happen and have the model take care of the computation. There are parallels here with *end user programming* [22] where users are given simplified interfaces to design behaviors for technical systems. This can be seen with complex technology such as robotics [26], where exploring possibilities requires specific hardware, functional libraries, and fixed laboratory environments [15], as well as in a simplified form in voice assistant systems to carry out home automation tasks [8]. While Brich et al. [8] found rule based approaches too simple for important tasks, and process-based approaches challenging for end users [8], LLMs offer an alternative, *declarative* approach, where users can describe what they would like to happen, and the system make it so, potentially allowing for reappropriation of functionality and a fluid approach to personalization and customization in response to their complex, situated needs. **We were curious to investigate opportunities and challenges of ways in which users engage with such LLM-enabled tangible artifacts (RQ)**.

To explore this space, we developed a pipeline where an LLM is given a description of an artifact's physical appearance, configuration and context, along with technical details enabling it to establish a correspondence between computational variables and real-world actions. Unlike commercial systems such as *SmartThings* or *Alexa*, our pipeline supports improvisation, ambiguity, and redefining interaction rules on the fly. Users describe the desired tangible action, and the LLM sends out the relevant commands to the hardware. This work is centered in HCI for AI practices, i.e. the development of techniques for creating AI powered interactions [2, 14, 20, 21], rather than approaches that use tangibility for explanation and understanding [e.g. 9, 13] or that use AI as part of the design process [29]. We contribute a *research artifact* [30] in close connection to Boer and Donovan [5]'s framing of such artefacts as *provotype*, in order to stimulate discourse [32] around new possibilities for LLM mediated interaction with tangible digital systems. We explored what kinds of prompt experiments the users carry out, their assumption of artifact's understanding of their prompts, and how they are able to modify the system through conversational repair and rule setting.

2 Prompting Realities: an LLM-based experiential pipeline

To explore this space, we developed a pipeline where users can control a physical artifact through conversational interactions with an LLM. We utilized a Research through Design approach [11, 25, 32] to refine and develop the software, hardware, and physical part of the pipeline through prototyping, starting with two key criteria: (a) an LLM mediates between the user and a physical artifact (b) the system provides enough openness to encourage non-technical users to improvise and explore through prompting. Through design iteration and unstructured reflective sessions we added two more criteria to provide better affordances for prompt experimentation: (c) the prototype should maintain a conversational memory to support customizing interactions and (d) the system accompanies textual response in alignment with its tangible feedback to enrich the conversational experience. Our overall intention was that the system would allow users to reappropriate, manipulate, and alter the functionalities of physical and tangible artifacts through natural language descriptions and LLM-based prompting techniques.

2.1 Provotype

Over the course of the project, we have experimented with various objects to control, for example tactile sensor creating patterns of action and a LED strips creating color patterns. Finally, we deliberately used a diorama (Figure 1) of a set of windmills in an out-of-context situation in which the windmills spin upon request, instead of being driven by wind. This encouraged users to see the windmills as both familiar and unfamiliar artifacts opening the space for reinterpretation and non-functional use [16]. We adapted this *defamiliarization* [4, 27] in our provotype—a prototype that provokes reflection and speculation rather than solving a known problem[5]—to encourage alternative interpretations of LLM-powered everyday interactions distanced from the experience of everyday smart devices. This strategic ambiguity [12] allowed users to focus on the open speculations around exploring new interaction and relations with the artifact like "calling the windmills names" or "asking them to dance together" without being constrained by preconceptions of how voice-controlled automation are commonly proceeded.

2.2 Implementation in-brief

The users send their desired prompt through a chatbot using their own mobile devices which then is processed via an LLM. The LLM generates both a natural language response—returning to the chat interface—and a JSON message—going to the diorama (Figure 2). This actuates the windmills appropriately based on user's descriptive instruction.

Users can type instructions such as 'make all the windmills spin fast' and the LLM will produce the relevant JSON commands to achieve the desired output as well as giving the user a feedback through the chat interface. We used a 700-word-long description as the *system instruction* (full version in supplementary materials). This consisted of i) context *"You are part of a system that controls three windmills."*; ii) information about other parts of the system: *"motors are adjusted via JSON commands sent to a CircuitPython controller using MQTT."*; iii) overall information about the prototype: *"The whole thing looks like a topographic scaled model"*; iv) names and descriptions for the windmills and; v) a list of *dos* and *don'ts*. This configuration creates an iterative, tangible loop from the user's prompt to the resulting physical actuation—shifting the role of system designers from direct programming to describing and setting initial system behavior.

3 Exploratory pilot study

We conducted an initial exploratory study (Ethics Review number: 4788) investigating the prototype's potential to probe the opportunities and challenges of this interaction model (interview guide, system logs, and transcriptions are available in supplementary documents). We conducted a try-out sessions with 6 participants from the first author's professional network focused on early-stage experiential engagement. Therefore, all tasks were followed by a consistent set of open-ended questions to capture the participants' thoughts and reflections. The study had three distinct sections, with gradually moving from familiarization with the system to increasing opportunities for more open exploration.

Firstly, to familiarize the participants with the process of giving natural language instructions in their own ways, we showed them a pre-recorded video of specific windmill movements and asked them to replicate it through the chatbot interface. This video was chose to resolve the issue of giving the users a verbal instruction about the task risking the chance of leading their wording and natural language articulation. Second, to explore how conversational memory influence prompt experimentation, participants were asked to create a 'trigger word' that intrigue certain system behavior. This explored whether users could map conversational tokens to physical states, enabling improvisational rule-setting on-the-go. Finally, we asked participants to physically modify the prototype by creating a river using blue tape followed by editing the description in the *system instruction* to adapt the new element. Participants then were asked to use the river as a spatial reference in their prompting. This activity probed the opportunities of higher agency of participants by modifying the contextual awareness in the system in the initial setup layer.

We collected i) audio recordings of responses to open-ended questions; ii) system logs; iii) photographs of modifications made in the diorama and the modified system instruction. We reflectively looked at user engagement logs and interview transcripts, using a Reflexive thematic analysis approach [6, 7] to short list interesting moments of interactions. We performed a brief data analysis session to extract "unanticipated possibilities" [24] that our experiential pipeline can provide for future inquiries on this area of research.

4 Discussion of results and implementation

This section explores insights from our study to identify emergent interaction patterns between users and the representative prototype, as well as opportunities and challenges that LLMs can bring into the tangible and embodied interaction design space.

General responses. Overall, the participants appreciated the fluidity of this approach, particularly the declarative aspect: *"I was surprised at how enjoyable it was to have it work according to my instructions."* (P3) This leads to an ease of exploration, of *"setting up rules, testing them, breaking them, and figuring out what made it tick"* (P4), and reduced frustration: *"if a chatbot gives me a wrong answer, I get annoyed. But if the windmill does the wrong thing, I feel like I should adjust my command"* (P2).

Conversational memory and fluid rules. The conversational memory enabled users to establish rules through natural language and recall them later. In the second part of the study, we asked users to assign 'trigger words' to windmill behaviors, testing whether the system could retain these associations across conversational turns. This shaped expectations of the AI's ability to store functional rules or a perceived learning which allowed a fluid, lighthearted approach to rule-setting: "*When I say "day" in a sentence, you should do the behavior [previously set]. All windmills are now stopped. The day is beautiful. Day detected! Windmills are now spinning as set. Okay stop. Windmills are stopped now. Today started on a positive note. Day wasn't mentioned, so no changes were made.*

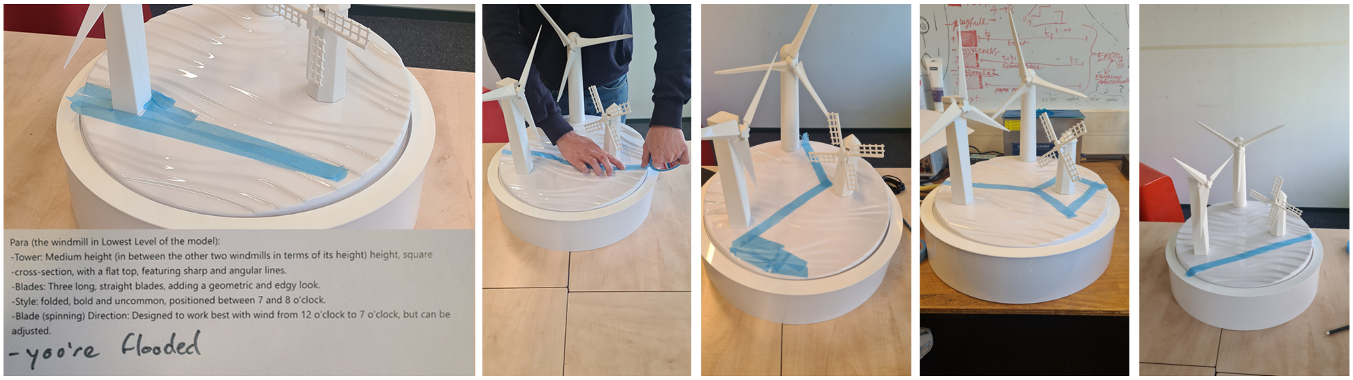

Figure 3: Incorporating a new element into the physical prototype through physical and descriptive modification of prototype

Okay, how about: Until this day, I feel a part of the studio. *Day detected! Windmills are spinning as set.* *Perfect!* "

This flexibility extended to redefining references within the system. For instance when P5 said: "*Stop the medium one.*", intending windmill 'Para' and the system spun 'Old', he replied: "*No, that was the small one!*", and the system adapted its response accordingly, treating 'Para' as the 'medium' windmill regardless of what was originally established in the system instruction (the 'old' windmill was installed in a medium part of the topographic model, which made the choice of 'medium' unexpected for the model). For P5, it was *"more engaging [than prompting with ChatGPT]. Because it was a physical system and [they] could immediately see whether [the] commands worked."* (P5)

From rules to co-performing tangible interactions. The combination of conversational memory with a *"tangible intermediary"* (P3) enabled participants to utilize incremental adjustments through intuitive measurements. For example, P6 tried to reduce the speed of a windmill: "*The tallest windmill is now spinning faster clockwise.* *Too fast, a little slow."* which shows a surprisingly flexible form of *co-performance* [17]. P2 referred to this as well:

> *"it felt more like a negotiation [. . .] it would just do something, and then either I would correct it or not. Where at home, if I say 'Hey Google, turn on the room light,' then it would just turn it on. And I would never say 'turn it on to 75% brightness' or 'dim it a little bit.'"* (P2)

The tangible actuation of windmills reinforced understanding through material change rather than abstract texts. This highlighted deviations from desired outcomes, leading some participants to rationalize inconsistencies, expectations and performance. The physical actuation seemed to reinforce participants sense that the model was following some sort of logic *"because it's moving in the real world."* (P5), even though the logic was absent and the result is stochastically generated. Participants engaged in iterative prompting *"to figure out what kind of words actually triggered actions"* (P6) through significant conversational repair to cover a lack of underlying understanding behind performance: *"I realized it doesn't take height that well. So I had to think a lot—how can I make it very simple and very direct, something which is not confusing, something which does not have another meaning..."* (p2)

Collaborative deictic references. The third part of the study asked participants to incorporate a river into the diorama and reference it in their descriptive prompts. This brought focus to the use of *deictic reference*, or referring to things by their relative positions. This is a common concern with tangible systems, as it fits naturally with human descriptive instructions within an physical context[1]: "*Turn on the windmill that is closest to the river."* (P2) This is a challenging area for the LLM, as it has only a very limited linguistic description of the relative positions, and needs to make some heroic leaps of inference. Users adopted different strategies to reference locations, from simple "*Turn on the windmill that is underneath the river."* (P2) to more complex: "*Turn on the windmill which is closest to the merging spot of the river."* (P4) or "*Turn on the windmill that is surrounded by a river."* (P6) Here users often ran to unexpected system behavior, inconsistent responses, or limitations in spatial reasoning. This aligns with prior research on human mental models of AI systems, where users construct their own understanding of underlying rules based on observed patterns[19]. The relational notions (e.g., close, far, behind) were often problematic. For instance P2 tried to enforce his spatial reasoning "*I already told you in the past that the old one or the most right-hand one is closest to the river."* There were also moments of repair: " *I think the one in the back is the furthest away from the river.* *Understood. Moving the correct windmill now."*

The ability to repair these pretended reasoning allowed for a relatively fluid interaction, as the system adjusted to expectations, although the limitations of the LLM's spatial reasoning were quite clear. Some users struggled with the spatial reasoning themselves, feeling that the *"most challenging part of incorporating the river [was] creating this mental model of space that would be interpretable for the AI."* (P3) They also reached for extremely complex commands, e.g., "*turn on the one that is the middle distance away from the entrance or exit of the rivers"* (P4) indicating both that they felt comfortable in setting up desired outcomes that would have been difficult to specify more formally, and easy it is to get lost in spatial references.

From deixis to physical reasoning. There was a natural shift in several interactions from using deictic references to bringing in physical and spatial reasoning, e.g. “*There is a storm on the ocean.* *All windmills will spin clockwise at maximum speed.*” (P5) or “*Turn on the* driest *mill*” (P1). Interestingly, some multi-turn reasoning examples worked correctly, for example: “ *all the windmills broke down.* *All windmills will stop now.* *they are repaired, but there's no wind.* *All windmills will remain stopped without wind.* ” (P5) In the case of unexpected system behavior, some participants were able to investigate, and to some extent repair, the disconnect in reasoning: “*there's no water in the lake.* *The lowest windmill will stop now.* *water in the lake doesn't influence wind and windmills* *there's no water in the lake* *No change; water doesn't affect windmills.* ” (P5) There were also moments of unexpectedly accurate response. One user modified the Assistant's system instructions by adding the phrase *“You are flooded!”* to description of the "Para" windmill. They later tested this modification by asking the prototype to “*Turn off the driest mill.*” P1 later reflected on this: *“when I told it to shut off the driest windmill, it turned the one furthest away from the one that is flooded [. . .] probably it has an idea.”*

5 Conclusion and future research

This preliminary work explored how LLMs can enable tangible interaction through natural language, revealing both opportunities and challenges with LLM-enabled artifacts. Our contribution is the research artifact as elaborated in Section 2, as well as the articulation of how conversational improvisation, ambiguity, and repair can serve as design materials for future LLM-based tangible interactions. Our findings suggest that while users successfully leveraged prompting to establish and recall rules leading to tangible actuation, inconsistencies in deictic references highlight fundamental gaps between human spatial reasoning and AI interpretation.

These insights present multiple avenues for future research. First, future work could investigate how LLM-enabled tangible systems can balance user expectations of memory persistence with the probabilistic behavior of LLMs, potentially through intermediary rule-based mechanisms and opportunities for refinement on-the-go. Second, the challenges observed in deictic articulation suggest a potential for incorporating spatial modeling or real-time sensory data to enhance spatial reasoning of LLMs as well as implementing Vision Language Models to constantly enhance system's context awareness.

These insights invite the HCI community to explore new interaction paradigms that embrace uncertainty, co-performance, and improvisation in AI-mediated tangible systems shaped by natural language prompting.

Acknowledgments

This publication was supported by the AI Futures Lab on Rights and Justice which is part of the TU Delft AI Labs program.

References

[1] Patti Bao and Darren Gergle. 2009. What's "This" You Say? The Use of Local References on Distant Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1029–1032. doi:10.1145/1518701.1518858

[2] Jesse Josua Benjamin, Arne Berger, Nick Merrill, and James Pierce. 2021. Machine Learning Uncertainty as a Design Material: A Post-Phenomenological Inquiry. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3411764.3445481

[3] Luca Beurer-Kellner, Marc Fischer, and Martin Vechev. 2023. Prompting Is Programming: A Query Language for Large Language Models. *LMQL as described in Prompting Is Programming: A Query Language for Large Language Models* 7, PLDI (June 2023), 186:1946–186:1969. doi:10.1145/3591300

[4] Julian Bleecker. 2009. Design Fiction: A Short Essay on Design, Science, Fact and Fiction.

[5] Laurens Boer and Jared Donovan. 2012. Provotypes for participatory innovation. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*. Association for Computing Machinery, New York, NY, USA, 388–397. doi:10.1145/2317956.2318014

[6] Virginia Braun and Victoria Clarke. 2019. Reflecting on Reflexive Thematic Analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (Aug. 2019), 589–597. doi:10.1080/2159676X.2019.1628806

[7] Virginia Braun and Victoria Clarke. 2021. One Size Fits All? What Counts as Quality Practice in (Reflexive) Thematic Analysis? *Qualitative Research in Psychology* 18, 3 (July 2021), 328–352. doi:10.1080/14780887.2020.1769238

[8] Julia Brich, Marcel Walch, Michael Rietzler, Michael Weber, and Florian Schaub. 2017. Exploring End User Programming Needs in Home Automation. *ACM Trans. Comput.-Hum. Interact.* 24, 2 (April 2017), 11:1–11:35. doi:10.1145/3057858

[9] Ashley Colley, Kaisa Väänänen, and Jonna Häkkilä. 2022. Tangible Explainable AI - an Initial Conceptual Framework. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia (MUM '22)*. Association for Computing Machinery, New York, NY, USA, 22–27. doi:10.1145/3568444.3568456

[10] Yang Deng, An Zhang, Yankai Lin, Xu Chen, Ji-Rong Wen, and Tat-Seng Chua. 2024. Large Language Model Powered Agents in the Web. In *Companion Proceedings of the ACM Web Conference 2024 (WWW '24)*. Association for Computing Machinery, New York, NY, USA, 1242–1245. doi:10.1145/3589335.3641240

[11] Abigail C. Durrant, John Vines, Jayne Wallace, and Joyce S. R. Yee. 2017. Research Through Design: Twenty-First Century Makers and Materialities. *Design Issues* 33, 3 (July 2017), 3–10. doi:10.1162/DESI-a-00447

[12] William W. Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a Resource for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. Association for Computing Machinery, New York, NY, USA, 233–240. doi:10.1145/642611.642653

[13] Maliheh Ghajargar and Jeffrey Bardzell. 2023. Making AI Understandable by Making it Tangible: Exploring the Design Space with Ten Concept Cards. In *Proceedings of the 34th Australian Conference on Human-Computer Interaction (OzCHI '22)*. Association for Computing Machinery, New York, NY, USA, 74–80. doi:10.1145/3572921.3572942

[14] David Philip Green, Joseph Lindley, Zach Mason, and Paul Coulton. 2022. A Design-Led Exploration of Material Interactions Between Machine Learning and Digital Portraiture. In *[ ] With Design: Reinventing Design Modes*. Springer, Singapore, Springer, Singapore, 3268–3283. doi:10.1007/978-981-19-4472-7_211

[15] Ulas Berk Karli, Juo-Tung Chen, Victor Nikhil Antony, and Chien-Ming Huang. 2024. Alchemist: LLM-Aided End-User Development of Robot Applications. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24)*. Association for Computing Machinery, New York, NY, USA, 361–370. doi:10.1145/3610977.3634969

[16] David Kirby. 2010. The Future is Now: Diegetic Prototypes and the Role of Popular Films in Generating Real-world Technological Development. *Social Studies of Science* 40, 1 (Feb. 2010), 41–70. doi:10.1177/0306312709338325 Publisher: SAGE Publications Ltd.

[17] Lenneke Kuijer and Elisa Giaccardi. 2018. Co-Performance: Conceptualizing the Role of Artificial Agency in the Design of Everyday Life. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3173699

[18] Hanyu Lai, Xiao Liu, Iat Long Iong, Shuntian Yao, Yuxuan Chen, Pengbo Shen, Hao Yu, Hanchen Zhang, Xiaohan Zhang, Yuxiao Dong, and Jie Tang. 2024. AutoWebGLM: A Large Language Model-based Web Navigating Agent. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)*. Association for Computing Machinery, New York, NY, USA, 5295–5306. doi:10.1145/3637528.3671620

[19] Michael Xieyang Liu, Advait Sarkar, Carina Negreanu, Benjamin Zorn, Jack Williams, Neil Toronto, and Andrew D. Gordon. 2023. “What It Wants Me To Say”: Bridging the Abstraction Gap Between End-User Programmers and Code-Generating Large Language Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–31. doi:10.1145/3544548.3580817

[20] Dave Murray-Rust, Maria Luce Lupetti, and Iohanna Nicenboim. 2024. Metaphor Gardening: Experiential engagements for designing AI interactions. In *Proceedings of DRS2024: Boston* (2024-06-23). Design Research Society, Boston, USA. doi:10.21606/drs.2024.376 Accessed: 2025-07-30.

[21] D. Murray-Rust, M. L. Lupetti, I. Nicenboim, and W. van der Hoog. 2024. Grasping AI: experiential exercises for designers. *AI & Society* 39, 4 (2024), 2891–2911. doi:10.1007/s00146-023-01794-y Accessed: 2025-07-30.

[22] Brad A. Myers, Amy J. Ko, Thomas D. LaToza, and YoungSeok Yoon. 2016. Programmers Are Users Too: Human-Centered Methods for Improving Programming Tools. *Computer* 49, 7 (July 2016), 44–52. doi:10.1109/MC.2016.200 Conference Name: Computer.

[23] Ying Sheng, Sudeep Gandhe, Bhargav Kanagal, Nick Edmonds, Zachary Fisher, Sandeep Tata, and Aarush Selvan. 2024. Measuring an LLM's Proficiency at Using APIs: A Query Generation Strategy. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)*. Association for Computing Machinery, New York, NY, USA, 5680–5689. doi:10.1145/3637528.3671592

[24] Hazel Smith. 2009. *Practice-led Research, Research-led Practice in the Creative Arts*. Edinburgh University Press. Google-Books-ID: mBCrBgAAQBAJ.

[25] Pieter Jan Stappers and Elisa Giaccardi. 2017. Research through Design. In *The Encyclopedia of Human-Computer Interaction*. The Interaction Design Foundation.

[26] Sai Vemprala, Rogerio Bonatti, Arthur Bucker, and Ashish Kapoor. 2023. ChatGPT for Robotics: Design Principles and Model Abilities. (Feb. 2023). https://www.microsoft.com/en-us/research/publication/chatgpt-for-robotics-design-principles-and-model-abilities/

[27] Ron Wakkary, Doenja Oogjes, Sabrina Hauser, Henry Lin, Cheng Cao, Leo Ma, and Tijs Duel. 2017. Morse Things: A Design Inquiry into the Gap Between Things and Us. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. Association for Computing Machinery, New York, NY, USA, 503–514. doi:10.1145/3064663.3064734

[28] Hongru Wang, Yujia Qin, Yankai Lin, Jeff Z. Pan, and Kam-Fai Wong. 2024. Empowering Large Language Models: Tool Learning for Real-World Interaction. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24)*. Association for Computing Machinery, New York, NY, USA, 2983–2986. doi:10.1145/3626772.3661381

[29] Justin D. Weisz, Jessica He, Michael Muller, Gabriela Hoefer, Rachel Miles, and Werner Geyer. 2024. Design Principles for Generative AI Applications. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3613904.3642466

[30] Jacob O. Wobbrock and Julie A. Kientz. 2016. Research contributions in human-computer interaction. *Interactions* 23, 3 (April 2016), 38–44. doi:10.1145/2907069

[31] Neng Zhang, Chao Liu, Xin Xia, Christoph Treude, Ying Zou, David Lo, and Zibin Zheng. 2022. ShellFusion: Answer Generation for Shell Programming Tasks via Knowledge Fusion. In *Proceedings of the 44th International Conference on Software Engineering (ICSE '22)*. Association for Computing Machinery, New York, NY, USA, 1970–1981. doi:10.1145/3510003.3510131

[32] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, New York, NY, USA, 493–502. doi:10.1145/1240624.1240704