Talking to Your Surroundings: Enabling Rich Human-Object Conversations Through AI and Augmented Reality

Fabrice Matulic
Preferred Networks, Japan
fmatulic@preferred.jp

ABSTRACT

Humans naturally tend to anthropomorphise objects, yet current interactions with AI agents and smart devices remain largely functional and disembodied. This position paper argues for the potential of combining Augmented Reality and multimodal AI to bridge this gap and enable richer and more meaningful human-object relationships. I propose visually embodied agents in AR, possessing unique personalities derived from the physical objects they represent and exhibiting awareness of their surrounding context. These situated, world-grounded personas can offer experiences that are not merely transactional and utilitarian but also relatable, engaging, and potentially emotionally supportive. I discuss the key technological enablers and potential applications of this approach to transform passive objects into active participants in our daily lives and enhance our connection with our surroundings.

1 INTRODUCTION

Humans possess a well-documented tendency to anthropomorphise inanimate objects, from naming vehicles, scolding malfunctioning appliances, to chatting with houseplants [9]. Far from being mere quirks, these behaviours often serve deeper psychosocial functions, helping people forge connections, navigate loneliness, and enhance emotional well-being [4, 7, 8]. This innate human capacity to perceive agency and personality in objects suggests a largely untapped potential for richer human-object relationships.

While voice assistants like Siri or Alexa and, more recently, conversational systems powered by Large Language Models (LLMs) [2, 5] allow for increasingly natural language interactions with devices, these exchanges remain predominantly functional or transactional. Current AI agents, typically disembodied voices or screen-based chatbots, lack the richness and emotional nuance of human-like conversation and possess limited grounding in, or direct perception of, the physical world they inhabit alongside the user. There exists a significant gap between our innate desire to relate to objects and the functional nature of current human-device interaction.

In this position paper, I argue that the convergence of Augmented Reality (AR) and advanced AI presents a unique opportunity to move beyond these limitations, enabling richer, more meaningful interactions with the everyday objects that surround us. I posit that by imbuing everyday physical objects with AI-driven, interactive AR characters (conceptualised in Figure 1), we can facilitate conversations that are not merely transactional but also more relatable, entertaining, and potentially emotionally supportive. This vision involves creating situated agents whose personalities and dialogues are directly linked to the objects they represent, allowing users to "talk with" their belongings in a meaningful way. Such an approach

Figure 1: Concept image representing AR characters overlaid on everyday objects that engage in rich human-like conversations with users.

aligns with more-than-human design philosophies, which recognise the interactive potential and agency of non-human entities [6]. We are actively developing a prototype system for smartphones and AR headsets [3] to explore this vision and its implications for how we relate to the objects around us.

2 THE POTENTIAL OF AI AND AR FOR HUMAN-OBJECT CONVERSATIONS

The synergy between real-time multimodal AI and AR enables fundamentally new ways to interact with the physical world, moving beyond simple voice commands towards rich, situated conversations with everyday objects. By integrating AI-powered vision, language, and speech processing within an AR framework, we can create interactions that are not only contextually relevant but also deeply personal and engaging.

A key element in realising this potential is the *visual embodiment* of AI agents directly onto physical objects via AR. Unlike disembodied voice assistants offering purely auditory feedback, providing objects with expressive visual characters, i.e. faces that react and emote, significantly enhances the interaction. This visual presence leverages our natural inclination to engage with face-like stimuli, making the interaction feel more direct and personal. Seeing a character "on" the object reinforces the connection between the AI personality and its physical anchor, while visual expressions, driven by AI, provide vital non-verbal cues that enrich communication and strengthen the perception of the character's persona and affective state.

This combination of AI and embodied AR enables a wide range of interesting scenarios:

- **Context-Aware Companionship:** Imagine a houseplant character visually assessing its own condition and proactively commenting ("I'm feeling a bit thirsty today!") or responding knowledgeably to care questions. Similarly, a coffee maker character, noticing increased usage, might thoughtfully inquire about the user's fatigue, offering moments of connection akin to a concerned friend and transforming passive objects into active participants in daily life.
- **Enhanced Utility and Information Access:** Instead of sterile status reports from IoT devices, an AR character representing a smart oven could playfully announce task completion or guide users through functions with personality-rich dialogue. A backpack character, aware of its contents and context, could offer timely reminders, making mundane interactions more intuitive and enjoyable by integrating object-specific knowledge with engaging personas.
- **Novel Social, Entertainment, and Emotional Support:** AR characters can introduce new layers to social situations, like playful banter from personified food items during a family meal, or enabling readers to converse with book characters emerging from the pages. Characters capable of empathetic conversation and responding to user emotions could also foster a sense of connection and potentially mitigate loneliness, tapping into a growing trend where generative AI is increasingly used for therapy and companionship (the top use case of GenAI in early 2025 according to a Harvard Business Review study [10]). This could be particularly valuable for individuals experiencing social isolation. Regular, positive interactions with familiar, personified objects might cultivate a comforting sense of presence within people's personal environments.

These examples illustrate how moving beyond disembodied, purely functional interactions towards visually embodied, personality-rich, and context-aware AR characters can fundamentally reshape our relationship with the objects and spaces around us, making technology feel more integrated, intuitive, and humane.

3 OUR WORK

We are currently developing a prototype that leverages AR and AI to transform everyday objects into interactive, talking characters [3]. Deployed on both AR headsets and smartphones, our system combines object tracking with real-time multimodal AI to enable dynamic, contextually aware conversations between users and AR characters, as well as among the characters themselves. In addition to the user's speech, the system captures and interprets various input sources, including AR device pose, visual features of objects and the environment, and user interactions through hand and eye tracking. This information is used to generate detailed context prompts that accompany speech input for a real-time LLM, which produces dialogue scripts for the characters, complete with mood tags that drive the animation of the characters' facial expressions. A text-to-speech module then lends a voice to the characters, bringing them to life.

A core aspect of our approach is the automatic generation of unique personalities tailored to each object. Upon first identifying an object via visual analysis, a vision language model (VLM) infers suitable personality traits, behavioural styles, and conversational tones based on the object's appearance and inferred function. For instance, a cactus might be given a "prickly" personality, while a cake might be assigned a "sweet" persona. This automatically derived personality profile guides the LLM to produce conversations with a captivating charm, making them more engaging and relatable, as the characters' responses reflect their individual traits and the characteristics of the objects they represent.

Furthermore, the system integrates dynamic environmental and interaction awareness to ensure characters feel truly situated. It processes visual scene changes (e.g. lights turning off) and user interactions (e.g. moving, touching, or gazing at an object) via dedicated modules. This contextual information is relayed to the LLM through the prompt system, enabling characters to react appropriately and dynamically. For instance, a character might express surprise if its object is suddenly moved or comment relevantly based on the user's focus of attention, making the interactions feel more grounded and responsive to the unfolding physical reality. This ongoing development serves as a platform for investigating the design and experience of these novel human-object conversations.

4 CONCLUSION AND FUTURE DIRECTIONS

I believe that the integration of AI and AR technologies presents a compelling opportunity to transform human-object interaction. Moving beyond purely functional and disembodied conversational agents, this AR-AI synergy allows us to imbue everyday objects with interactive, context-aware personas, which enable experiences that are more engaging, meaningful, and emotionally resonant. By creating visually embodied characters grounded in the physical world, we can design interactions that align more closely with our natural tendency to anthropomorphise and connect with our surroundings.

Significant research questions remain regarding the long-term effects and societal implications of such an approach. Future work includes longitudinal studies conducted within natural home environments to evaluate the impact of daily interactions with personified objects on user well-being, feelings of connection or loneliness, and overall integration into daily routines. Understanding how these relationships evolve over time and assessing the capabilities of AI models to maintain coherent, long-term interactions based on memory are important next steps.

Furthermore, the creative potential is vast. The enduring appeal of narratives where inanimate objects come alive, as seen in imaginative works like "Toy Story"and "The Brave Little Toaster", or the humorous exchanges in comedy shows like "Annoying Orange" [1], underscores a deep-seated fascination with personified objects. Exploring how our approach could enable new forms of interactive entertainment, character-driven storytelling, or educational content presents exciting avenues for future exploration.

I envision a future where the objects around us evolve into active participants in our lives, contributing to our well-being and fostering a deeper connection with our surroundings. While this vision may seem futuristic, I believe that the rapid advancements in AI and AR technologies are steadily bringing it closer to reality.

REFERENCES

[1] Dane Boedigheimer. [n. d.]. The Annoying Orange. https://www.youtube.com/channel/UCi-5OZ2tYuwMLIcEyOsbdRA. Accessed: 2024-12-03.

[2] Mustafa Doga Dogan, Eric J Gonzalez, Karan Ahuja, Ruofei Du, Andrea Colaço, Johnny Lee, Mar Gonzalez-Franco, and David Kim. 2024. Augmented Object Intelligence with XR-Objects. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 19, 15 pages. https://doi.org/10.1145/3654777.3676379

[3] Nozomu Iwai and Fabrice Matulic. 2025. Bringing Everyday Objects to Life in Augmented Reality with AI-Powered Talking Characters. In *Extended Abstracts of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '25)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3706599.3719978

[4] Valerie K. Jones, Michael Hanus, Changmin Yan, Marcia Y. Shade, Julie Blaskewicz Boron, and Rafael Maschieri Bicudo. 2021. Reducing Loneliness Among Aging Adults: The Roles of Personal Voice Assistants and Anthropomorphic Interactions. *Frontiers in Public Health* 9 (2021). https://doi.org/10.3389/fpubh.2021.750736

[5] Evan King, Haoxiang Yu, Sangsu Lee, and Christine Julien. 2024. Sasha: Creative Goal-Oriented Reasoning in Smart Homes with Large Language Models. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 12 (March 2024), 38 pages. https://doi.org/10.1145/3643505

[6] Anuradha Reddy, A. Baki Kocaballi, Iohanna Nicenboim, Marie Louise Juul Søndergaard, Maria Luce Lupetti, Cayla Key, Chris Speed, Dan Lockton, Elisa Giaccardi, Francisca Grommé, Holly Robbins, Namrata Primlani, Paulina Yurman, Shanti Sumartojo, Thao Phan, Viktor Bedö, and Yolande Strengers. 2021. Making Everyday Things Talk: Speculative Conversations into the Future of Voice Interfaces at Home. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 23, 16 pages. https://doi.org/10.1145/3411763.3450390

[7] Kristen Rogers. 2024. Why do people give feelings to inanimate objects? What experts say. *CNN* (7 September 2024). https://edition.cnn.com/2024/09/07/health/empathize-inanimate-objects-anthropomorphize-wellness/index.html Accessed December 3, 2024.

[8] Shane Saunderson and Isabella Kuo. 2022. Why We See Humanity in Everyday Objects. (6 April 2022). https://shanesaunderson.com/2022/04/06/why-we-see-humanity-in-everyday-objects/ Accessed on December 03, 2024.

[9] Karen Stollznow. 2014. *Talking Objects*. Palgrave Macmillan UK, London, 207–216. https://doi.org/10.1057/9781137404862_21

[10] Marc Zao-Sanders. 2025. How People Are Really Using Gen AI in 2025. *Harvard Business Review* (April 2025). https://hbr.org/2025/04/how-people-are-really-using-gen-ai-in-2025