A Vision for Room-scale AI Interaction

Carl Oechsner[1], Jan Leusmann[1], Xuedong Zhang[1], Thomas Weber[1] and Sven Mayer[1]

[1]LMU Munich, Munich, Germany

Abstract

Interactive systems often burden users with micromanagement, requiring them to oversee and control every detail of the interaction. This limits scalability and efficiency in interacting with environments. We address this by presenting a vision for room-scale AI, where interactive environments are enriched with AI functionality to facilitate natural, context-aware interaction. Leveraging multi-agent systems, tasks can be divided into smaller components, enabling AI agent-powered devices to proactively assess situations, support users, and adapt to the environment as needed. This paradigm shift transforms the environment into an active collaborator. We present the potential of room-scale AI through use cases and highlight its broader implications. Finally, we outline key challenges and research directions to advance this transformative concept, aiming to reimagine how users interact with and are supported by intelligent environments.

Keywords

human-ai interaction, human-robot interaction, room-scale ai, llm, agents

1. Introduction

Artificial Intelligence (AI) increasingly extends beyond the purely digital realm into the physical world, enabled, for example, by the wide availability of smart home devices featuring a range of sensors and actuators. The data that these devices collect allow AI systems to construct a model of the physical context, analyze how it changes, and act upon this information. For instance, recent advances like Large Language Models (LLMs) enable AI to combine contextual and user-specific information – such as intent – to suggest, plan, or execute complex processes and even complete whole tasks for the user. However, most current systems are still rooted in a paradigm where users must oversee tasks and provide explicit input and control. Thus, one emerging challenge is creating environments that move beyond this paradigm to independently assess situations, support users dynamically, and proactively adapt without requiring constant user intervention.

Recent advancements in user interface design with and for AI, such as conversational systems, have made interaction with intelligent systems more natural and accessible across many platforms, like smartphones and wearables. While these systems excel at providing natural user interaction, they typically still rely on explicit user input and are limited in scale and scope to individual devices. In contrast, using sensing data from multiple environmental sources in combination with AI to create a model of the environment – a *room-scale AI system* – enables dynamic, context-aware, and multi-modal engagement. Current implementations of smart environments, such as those centered around devices like Amazon Echo or Google Home in a private space, simplify interaction by centralizing communication. However, these systems typically interact with devices in the environment in isolation, focusing solely on their individual functions without understanding the broader implications of their use. This disconnected approach limits their capabilities and generally reduces the potential efficiency and fluidity of interaction. This model of the environment also hinders its ability to act collaboratively and autonomously, respond to contextual factors, or anticipate user needs effectively.

To address these limitations, we propose a vision for *room-scale AI*, where environments become active

Joint Proceedings of the ACM IUI Workshops 2025, March 24-27, 2025, Cagliari, Italy

c.oechsner@lmu.de (C. Oechsner); jan.leusmann@ifi.lmu.de (J. Leusmann); xuedong.zhang@ifi.lmu.de (X. Zhang); thomas.weber@ifi.lmu.de (T. Weber); info@sven-mayer.com (S. Mayer)
https://leusmann.io/ (J. Leusmann); https://sven-mayer.com (S. Mayer)
0000-0002-5901-6811 (C. Oechsner); 0000-0001-9700-5868 (J. Leusmann); 0009-0002-6361-1717 (X. Zhang); 0000-0002-6894-605X (T. Weber); 0000-0001-5462-8782 (S. Mayer)

agents that facilitate proactive, context-aware collaboration. Using sensors and actuators, coupled with AI-inference methods like computer vision and natural language processing, room-scale AI environments can autonomously assess and modify their surroundings to support users. While *room-scale AI* has to keep one centrally managed communication channel as the user interface, leveraging multi-agent systems, AI agent-powered devices can divide complex tasks into smaller, manageable components, with AI-powered devices specializing in specific subtasks. For example, these environments could include users' cognitive-affective states [1], social context [2], personality [3], or curiosity [4] to enhance interaction. Moreover, this proactive capability allows the system to preempt undesirable or dangerous situations, adjust to user preferences, and coordinate interactions across devices. With the internal environment model, it is crucial to design interactions that balance proactive support with user trust and account for privacy concerns, for instance, through privacy communication patterns [5].

In this work, we outline a vision for room-scale AI interaction across various exemplary contexts, including households, workshops, emergency rooms, and factories. Implementing such systems presents technical, social, and ethical challenges. A key focus is enabling environments to act proactively by dynamically assessing situations, anticipating user needs, and adapting in real-time. This involves seamless integration of sensors and actuators, maintaining state awareness to track environmental changes, and handling ambiguous or incomplete information. Proactive systems must also manage real-time replanning, provide alternatives during failures, and offer alternatives when the users' goals can not be achieved. Additionally, multi-user interactions and privacy concerns require carefully balancing user support with ethical design principles. Addressing these challenges will allow room-scale AI to transform environments into intelligent, anticipatory collaborators, enhancing efficiency, safety, and user satisfaction across diverse applications.

2. Room-scale AI Interaction

We envision a paradigm shift from users needing to interact with multiple individual devices to a single intelligent room-scale AI agent. While centralized agents interfacing smart environments have already been proposed to direct the personality of the room [6], recent advancements in LLMs now offer the potential to go beyond a controlling instance and create agents that consolidate information from the environment's smart devices, sensors, and actuators, combining it with contextual knowledge to assist users proactively. LLM-based agents have demonstrated their capability of generating natural and context-aware responses [7], planning correct sequences of actions [8], and selecting appropriate tools or combinations thereof to achieve complex goals. These capabilities make LLM-based agents particularly suitable for room-scale AI systems [9].

To realize this vision, the room-scale AI agent must access all devices in the environment, e.g., through APIs like Home Assistant, and monitor and control their capabilities, status, and sensor data. This

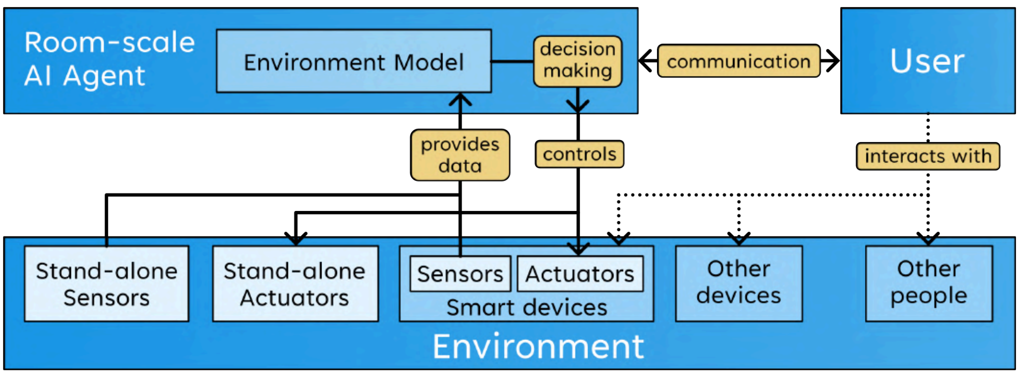

Figure 1: A room-scale AI system gathers data from the various devices in the environment. It forms a model of the environment through interaction with the user, prior data, and contextual information. Based on the model, it can decide when and how to independently and proactively control the environment to support the user best.

situational awareness enables the agent to form a model of the environment, allowing it to anticipate user needs, plan contextually relevant actions, make decisions independently, and act proactively. Here, proactivity can manifest in two ways: implicit proactivity, where the agent seamlessly assists the user without requiring explicit prompts, e.g., by altering the environment, or explicit proactivity, where the agent directly engages with the user to inquire about information for better assistance or to inform the user about important updates or decisions. For example, the agent could dynamically replan tasks, adjust settings based on user preferences, or even preemptively warn users of potential issues, such as a device malfunction. Based on sensor data, it can suggest actions tailored to current user activity. Entities requiring user action communicate their needs contextually, minimizing overhead by engaging the user at opportune moments. For example, while using the sink, the system may indicate nearby plants that need watering, as suggested by Bittner et al. [10]. Unlike traditional systems requiring explicit user input, room-scale AI fosters seamless, anticipatory interaction by adapting to users' behaviors and environmental changes.

A multi-agent architecture can further enhance this system. By leveraging LLM-based multi-agents, the strengths of specialized agents can be combined to improve robustness, autonomous problem-solving, and collaborative efficiency [11, 12, 13]. These agents can divide tasks into manageable components, coordinate their actions, and handle complex situations collaboratively, ensuring that the environment responds proactively to both immediate and long-term needs.

This approach has the potential to simplify user interaction radically. Users no longer need to remember which device to interact with or how to operate it; instead, the room-scale AI agent acts as a single, unified interface that proactively engages with users. By leveraging contextual information from sensors and devices, the agent can anticipate user needs, offer assistance before being asked, and adapt dynamically to changing conditions. This can reduce mental load and the amount of overhead for the user while increasing efficiency. This represents a significant step forward in creating intelligent, adaptive, and user-friendly environments (see Figure 1 for an example architecture).

3. Use Cases

In the following, we will outline four use cases and scenarios for room-scale AI that demonstrate how it can be beneficial in different contexts. Table 1 summarizes how these use cases represent a cross-section across potential contexts with different properties.

3.1. Everyday Household Activities

Targeting individual end users, one objective can be to simplify and enhance common household activities, such as cooking [14]. There is already a wide selection of smart home appliances for this domain, but their current usage is often rudimentary. Users mostly interact with smart appliances individually and explicitly. A room-scale AI environment could incorporate all these devices into one agent that benefits from a more comprehensive and connected view of the room.

Table 1
We selected four exemplary use cases that represent a diverse range of potential scenarios in which room-scale AI can provide benefits.

Use Case	Stakes	# People	Task Predictability	Internal Env. Consistency	External Env. Consistency
Everyday household Activities	Low	1-3	Moderate	Moderate	Low
Workshop	Medium	1-3	Low	Low	Low
Emergency Room	High	2-10	Moderate	High	High
Factory	Low	>10	High	High	Moderate

Actuators, Sensors, and Infrastructure Such an environment model can build on many existing actuators (e.g., ovens, vacuum robots, lights) and sensors (e.g., cameras, temperature sensors, motion detectors) already integrated into commercially available appliances, making this use case very accessible. Furthermore, these devices are typically connected via home networking systems and offer APIs, which ease integration.

Interaction This diversity in sensing and actuating technology allows a wide range of multi-modal interactions for explicit proactivity. For example, verbal commands can be beneficial when manual interaction is impossible, e.g., because the user is occupied or has dirty hands. Many of the activities in a kitchen also involve steps where the AI agent can proactively perform tasks inferred from the environment's state and the user's goal without needing additional explicit instructions [15]. To keep the user informed, existing appliances offer some simple means that can be further enhanced, for example, by explaining actions in-situ through augmented reality (AR) visualizations [16, 17].

Scenario Consider a user preparing dinner. After scanning the fridge's contents, an AI system could suggest potential recipes through explicit proactivity and engage the user in a conversation about available options based on their diet. Once a decision is made, the preparation progress is monitored using cameras and other sensors. Through this, it can detect when the user has chopped ingredients and is ready to start cooking, for which it can proactively preheat the oven. Similarly, when a pot is about to boil over, the system can lower the temperature without the user's explicit command. This implicit proactivity allows the system to respond to the different tasks with adequate caution needed to handle the objects [18]. Finally, it could inform other household members when it determines that the meal is nearing completion. This process eliminates the need for explicit configuration, as the AI adapts to the user's actions and the environment's state.

(a) Kitchen

(b) Workshop

(c) Emergency Room

(d) Factory

Figure 2: Visuals of Room-Scale AI Use Cases

Benefits This approach significantly reduces complexity and burden for users. Instead of setting up complex rules, which quickly becomes unfeasible given the diverse range of household tasks, or manually operating devices, users can collaborate with the AI, relying on the system's context-aware intelligence to understand the situation at hand as a whole and assist with tasks seamlessly, integrating data from a wide range of household appliances without requiring the user to make all decisions.

3.2. Workshop

Workshops are a use case that lies at the boundary between private and professional use. This environment is characterized by a diverse set of specialized tools and materials. The activities can range from simple, repetitive tasks, like sorting fasteners, to complex creative projects, like wood carving or engraving. Thus, people using these spaces can range from novices to experienced professionals.

Actuators, Sensors, and Infrastructure In contrast to domestic settings, there are not yet as many smart devices widely available for workshops. However, with Industry 4.0, professional workshop environments are increasingly equipped with smart devices, which might propagate to personal workshops soon. One important use case for smart devices in the workshop is the regulation of safety, e.g., smoke-dust sensors, CO detectors, cameras, air filters, ventilation, or cleaning robots. Combined with more digitally controlled tools, this offers the potential for a connected environment and room-scale AI.

Interaction Similar to the household scenario, users may not be able to interact with an assistive system in a workshop directly because they may be occupied otherwise. In addition, the noise level in a workshop can be such that verbal interaction is not viable. Giving the environment the ability to perform tasks without explicit user requests proactively can mitigate this, as it allows the system to perform, for example, context-specific, safety-critical tasks without requiring explicit user input. The diversity of tasks in this space can also lead to ambiguity in verbal instructions, gestures, etc. A room-scale AI system can combine ambiguous, explicit user input with context information to clear up and help achieve the user's goals.

Scenario Imagine a worker in a woodworking shop building a bookcase. Displaying explicit proactivity, the system can suggest ways to execute the project based on available materials and tools. Knowing the user's intent, a robot can prepare the space and organize tools just in time, aligning with the project's progression and user preferences. As the worker mills, cuts, and sands, the AI supports by moving materials, handling repetitive tasks, and ensuring safety through implicit and explicit proactivity, like adjusting air filtration or reminding the worker to wear a mask. It autonomously removes obstacles, mitigates safety risks, and warns of potential injuries. After project completion, the system cleans the space, gathers tool data for predictive maintenance, and orders consumables, ensuring the workshop is ready for the next task.

Benefits In a workshop setting, room-scale AI can support tasks that require precision and coordination. Computer vision and sensors can monitor the work to avoid costly errors, e.g., by optimizing material usage and inventory. Proactively reordering supplies, preparing tools, or taking over complete subtasks can greatly increase productivity and reduce downtime. Since explicit interaction can be challenging, e.g., due to the noise level or while lifting heavy objects, a system capable of proactively determining when support is necessary can transform the workshop into a dynamic, cooperative environment that adapts to the user's needs while improving safety and efficiency.

3.3. Emergency Room

An operating room is an environment to provide essential and life-saving medical care where thoroughness and knowing your next steps are key. While teams of highly trained professionals have well-defined procedures, they also need to be able to adapt to emergency circumstances. For this, the staff relies on functional medical tools, equipment, and consumables available in their designated places.

Actuators, Sensors, and Infrastructure In an ER, medical monitoring systems track patients' vital signs, like heart rate and respiratory rate, with alarms and visualizations alerting the medical team. The patients' beds are equipped with sensors to monitor occupancy or detect falls and can be adjusted in height and position for ergonomic support. Electronic health record (EHR) systems centralize patient data, including medical history, test results, and treatment plans, aiding decision-making. Equipment trackers provide real-time availability and location. However, a lack of system interconnection often leads to organizational overhead and potential procedural delays.

Interaction As medical professionals need full attention on life-saving measures, their time for additional interaction in emergency rooms is limited and may require supporting staff. As explicit touch interaction is challenging since equipment needs to remain sterile, and voice instructions are not viable due to noise levels, or because they may upset patients, this setting can greatly benefit from implicit interaction and systems that reduce the mental load and overhead for the medical staff by performing tasks proactively.

Scenario Since time can be a critical factor, scenarios in the ER can benefit from the interconnection of agents between multiple locations: already during the transport to the hospital, data is recorded in the ambulance, which can then inform the environmental model in the ER and decisions on how to set it up, e.g., by preparing medication or notifying specialists. As the patient is treated, room-scale AI can prepare tools and mitigate safety risks. Similar to previous scenarios, if the situation changes, the system is flexible enough to adapt without distracting the staff by requiring reconfiguration. Using a system that already consolidates information about the patient, conducts procedures, and administers medication can also reduce documentation overhead through automation, giving the medical staff more time to focus on the patients.

Benefits Having a comprehensive model of the ER itself and beyond, room-scale AI helps medical professionals by allowing them to focus on patient care while letting the system take care of supporting activities. Like other use cases, room-scale AI can aid the medical staff and increase their effectiveness by taking over trivial but potentially distracting tasks within the ER and increase safety by alerting staff to critical changes in real-time. Through its knowledge of adjacent spaces like the ambulance or other ER patients, the system can help prioritize or detect synergies. In addition, it is available to the medical staff if they explicitly require information about patients or previous activities in their treatment.

3.4. Factory

A factory is a highly collaborative space focused on transforming raw materials into finished products through interconnected processes to deliver goods to the market. This transformation relies on the coordinated efforts of various roles, including production planners, machine operators, and quality inspectors. The complex flow of information and multi-person collaboration are central to human-factory interaction. We envision a room-scale AI factory equipped with a Digital Twin, supporting key processes such as resource allocation, work planning, and task scheduling. Additionally, intelligent information processing capabilities would tailor information delivery to workers in different roles, presenting it clearly while avoiding information overload [19].

Actuators, Sensors, and Infrastructure A factory has already integrated various equipment, such as industrial robots, conveyor systems, autonomous guided vehicles (AGVs) or autonomous mobile robots (AMRs) for internal logistics, and environmental control devices such as air conditioners or ventilation systems to maintain optimal conditions for production. In a room-scale AI factory, in addition to these basic devices, various Internet of Things (IoT) sensors collecting real-time data are essential for constructing a Digital Twin, such as RGB and depth cameras or motion and positioning sensors (such as IMUs and LiDAR).

Interaction As factories usually involve a larger group of people and high noise levels, verbal communication might not always be practical. Thus, the room-scale AI should use mainly visual communication modalities to interact with users, e.g., projection-based support [20]. AR might also be a good solution, as it enables each user direct and personal feedback [21, 22]. To ensure worker safety, the system can integrate safety alerts. For example, it can use lights, sounds, or vibrations to warn workers of potential dangers. Additionally, the system can leverage wearable devices like smartwatches or biosensors to monitor workers' health status, immediately issuing alerts and taking protective measures, if necessary, to ensure the workers' well-being.

Scenario Consider a scenario where a production line is running, and suddenly, a machine malfunctions. The room-scale AI system can immediately locate the faulty equipment and inform the responsible maintenance staff. Once they arrive at the site, the system analyzes the possible causes based on sensor data and historical records, provides targeted repair suggestions, and predicts downtime based on the type of malfunction and the active dialogue with the repair crew. During the repair process, personnel assignments are adjusted, and workers are reallocated from the affected production line to other tasks based on the current production schedule. Once the repair is completed, the system integrates the latest production information. It redistributes workers and resources and re-enables the production line to resume normal operations.

Benefits Room-scale AI enhances efficiency, quality control, and worker safety in manufacturing. Digital Twins offer floor managers a cohesive view of operations and enable querying for deeper insights via conversational interfaces. AI can autonomously detect product quality issues or workflow inefficiencies and adjust machine parameters accordingly. It can also mitigate risks by activating warning lights or floor markings when workers approach hazardous areas.

For collaborative human-robot tasks, AI dynamically adapts robot behavior to workers' movements, reducing accident risks. Predictive maintenance monitors equipment health and schedules repairs to minimize downtime. With multimodal interfaces like AR glasses, workers can access instructions and diagnostics directly, streamlining workflows and transforming factories into intelligent, optimized ecosystems.

4. Challenges

The implementation of room-scale AI faces several challenges, including physical constraints, technical limitations, human factors, and ethical considerations. Below, we discuss key issues and considerations for a successful implementation.

A room-scale agent is meant as an addition and improvement. Thus, additional devices or sensors must not obstruct regular space usage or require additional attention from the user, which aligns with Mark Weiser's early vision of ubiquitous computing [23]. This is particularly important in critical domains like the ER but also applies to simpler household use. However, different circumstances impose different priorities. For example, the ER will emphasize reliability more, whereas home use will focus on convenience and user acceptance.

To provide proactive support, a room-scale AI must be capable of handling unexpected state changes in the environment. This goes beyond merely distinguishing between visually similar objects [24, 25, 26]. For example, environmental changes can occur independently of user or system actions, e.g., fluids may evaporate, and fruit can perish. Thus, these systems must maintain a consistent model of the environment, continuously updating it to account for predictable but also unexpected changes. Context awareness through sensor fusion has long been one of the main challenges for applications in ubiquitous environments [27]. The interpretation of sensor data becomes even more complex when conflicting information arises from multiple sources, such as noise or discrepancies between sensors. This is particularly critical in robotic environments where actions based on erroneous interpretations could pose risks to the user. A proactive system can mitigate such a situation by requesting clarification from the user. Similarly, in settings with overlapping conversations or loud machinery, a system must be able to deal with this on its own or by initiating explicit interaction with the user. This also applies to multi-user environments where instructions and feedback to the system may be ambivalent. Resolving this may require a social or hierarchical understanding of the persons involved.

Even with the best conflict resolution, errors can happen, and systems and sensors can fail. In this case, the room agent should not simply stop working but proactively provide alternative solutions. For example, a smart home agent could use the stove or a toaster if the oven is defective. Especially in critical domains like the ER, disruptions must be minimized while a fallback solution is available at any time. Issues must be clearly and transparently communicated to allow the user to deal with the problem.

Privacy is another area where clear communication is key. As a room-scale AI system relies on a lot of potentially sensitive and intimate data, informing the user when and how data is recorded, stored, or processed is an important factor towards user acceptance and trust [5].

While we outlined similarities to challenges of traditional ubiquitous systems, like sensor fusion or unobtrusiveness, a unique challenge to robotic room-scale AI is to balance autonomy and user trust. Since it can proactively intervene and alter the environment, its level of autonomy must be carefully balanced with the users' preferences while enabling them to override the room agent's decisions. Achieving user trust is crucial for acceptance and involves, among others, ensuring predictability and respecting social norms of interaction [28, 29].

Finally, not every device in an environment will be easily accessible. Particularly, specialized settings like hospitals may rely on domain-specific hardware that might lack the necessary connectivity or data standards. Thus, an AI agent that should have a comprehensive model of the environment will need ways to integrate those, for example, by adding independent sensors or simulating regular usage using robotic arms.

While these use cases are used to showcase potential, they can only serve as a starting point. To sharpen these visions, the next step is to engage with respective domain users and experts to identify potential constraints and understand their needs and expectations from a user perspective.

5. Summary and Outlook

In this paper, we outlined the concept of *room-scale AI*, which presents a paradigmatic change in how users interact with their environment. Where currently most interactions require explicit input from the user, we envision an agent that maintains a model of the environment and can also affect and alter it. It can achieve this by proactively engaging in explicit interactions with users or making decisions based on implicit factors like the environment's state or the user's intent.

To illustrate the practical applications of room-scale AI, we present four distinct use cases, showcasing how this concept can offer unique benefits across various domains and cater to different user groups. These use cases highlight the potential of such an environment to enhance user experiences, streamline processes, and provide context-aware assistance. Based on these use cases, we identify key challenges that need to be addressed to make this vision a reality.

Declaration on Generative AI

During the preparation of this work, the author(s) used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword, Improve writing style. Further, the author(s) used ChatGPT for Figure 2 in order to: Generate images. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] M. Kraus, D. Betancourt, W. Minker, Does it affect you? social and learning implications of using cognitive-affective state recognition for proactive human-robot tutoring, in: 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), IEEE, 2023, pp. 928–935. doi:10.1109/RO-MAN57019.2023.10309574.

[2] M. Kraus, N. Wagner, R. Riekenbrauck, W. Minker, Improving proactive dialog agents using socially-aware reinforcement learning, in: Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization, UMAP '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 146–155. doi:10.1145/3565472.3595611.

[3] L. Christoforakos, S. Diefenbach, D. Ullrich, Designing robots with personality, in: Meaningful Futures with Robots, Chapman and Hall/CRC, 2022, pp. 70–78.

[4] J. Leusmann, C. Wang, S. Mayer, M. Gienger, A. Schmidt, How to design interactive physical learning robots that are fun to teach: If you are curious, i will show you!, IEEE Pervasive Computing 23 (2024) 1536–1268. doi:10.1109/MPRV.2024.3516432.

[5] M. Windl, J. Leusmann, A. Schmidt, S. S. Feger, S. Mayer, Privacy communication patterns for domestic robots, in: Tenth Symposium on Usable Privacy and Security (SOUPS 2024), USENIX Association, Philadelphia, PA, USA, 2024. URL: https://www.usenix.org/conference/soups2024/presentation/windl.

[6] S. Diefenbach, A. Butz, D. Ullrich, Intelligence comes from within—personality as a ui paradigm for smart spaces, Designs 4 (2020). doi:10.3390/designs4030018.

[7] S. Gao, J. Dwivedi-Yu, P. Yu, X. E. Tan, R. Pasunuru, O. Golovneva, K. Sinha, A. Celikyilmaz, A. Bosselut, T. Wang, Efficient tool use with chain-of-abstraction reasoning, 2025. URL: https://arxiv.org/abs/2401.17464. arXiv:2401.17464.

[8] F. Joublin, A. Ceravola, P. Smirnov, F. Ocker, J. Deigmoeller, A. Belardinelli, C. Wang, S. Hasler, D. Tanneberg, M. Gienger, CoPAL: Corrective Planning of Robot Actions with Large Language Models, 2023. doi:10.48550/arXiv.2310.07263.

[9] L. Liao, G. H. Yang, C. Shah, Proactive conversational agents, in: Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 1244–1247. doi:10.1145/3539597.3572724.

[10] B. Bittner, I. Aslan, C. T. Dang, E. André, Of Smarthomes, IoT Plants, and Implicit Interaction Design, in: Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction, ACM, Tempe Arizona USA, 2019, pp. 145–154. URL: https://dl.acm.org/doi/10.1145/3294109.3295618. doi:10.1145/3294109.3295618.

[11] J. He, C. Treude, D. Lo, Llm-based multi-agent systems for software engineering: Literature review, vision and the road ahead, 2024. URL: https://arxiv.org/abs/2404.04834. arXiv:2404.04834.

[12] S. Rasal, Llm harmony: Multi-agent communication for problem solving, 2024. arXiv:2401.01312.

[13] Y. Talebirad, A. Nadiri, Multi-agent collaboration: Harnessing the power of intelligent llm agents, 2023. URL: https://arxiv.org/abs/2306.03314. arXiv:2306.03314.

[14] C. Oechsner, S. Mayer, A. Butz, Challenges and opportunities of cooperative robots as cooking appliances, in: Proceedings of the 2022 Workshop on Engaging with Automation, AutomationXP22, 2022. URL: https://sven-mayer.com/wp-content/uploads/2022/05/oechsner2022challenges.pdf.

[15] M. Kraus, N. Wagner, W. Minker, A. Agrawal, A. Schmidt, P. K. Prasad, W. Ertel, Kurt: A household assistance robot capable of proactive dialogue, in: 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), IEEE, New York, NY, USA, 2022, pp. 855–859. doi:10.1109/HRI53351.2022.9889357.

[16] C. Wang, A. Belardinelli, S. Hasler, T. Stouraitis, D. Tanneberg, M. Gienger, Explainable human-robot training and cooperation with augmented reality, in: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, CHI EA '23, Association for Computing Machinery, New York, NY, USA, 2023. doi:10.1145/3544549.3583889.

[17] A. Belardinelli, C. Wang, M. Gienger, Explainable human-robot interaction for imitation learning in augmented reality, in: C. Piazza, P. Capsi-Morales, L. Figueredo, M. Keppler, H. Schütze (Eds.),

Human-Friendly Robotics 2023, Springer Nature Switzerland, Cham, 2024, pp. 94–109.
[18] J. Leusmann, C. Oechsner, J. Prinz, R. Welsch, S. Mayer, A database for kitchen objects: Investigating danger perception in the context of human-robot interaction, in: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, CHI EA '23, Association for Computing Machinery, New York, NY, USA, 2023. doi:10.1145/3544549.3585884.
[19] D. Toure, R. Welsch, S. Mayer, The future of proxemic interaction in smart factories., in: AutomationXP@ CHI, 2021. URL: https://sven-mayer.com/wp-content/uploads/2021/05/toure2021future.pdf.
[20] M. Funk, S. Mayer, A. Schmidt, Using in-situ projection to support cognitively impaired workers at the workplace, in: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility, ASSETS '15, Association for Computing Machinery, New York, NY, USA, 2015, p. 185–192. doi:10.1145/2700648.2809853.
[21] V. Paelke, Augmented reality in the smart factory: Supporting workers in an industry 4.0. environment, in: Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA), 2014, pp. 1–4. doi:10.1109/ETFA.2014.7005252.
[22] H. Subakti, J.-R. Jiang, Indoor augmented reality using deep learning for industry 4.0 smart factories, in: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), volume 02, 2018, pp. 63–68. doi:10.1109/COMPSAC.2018.10204.
[23] M. Weiser, J. S. Brown, The coming age of calm technology, in: Beyond calculation: The next fifty years of computing, Springer, 1997, pp. 75–85.
[24] S. Mayer, G. Laput, C. Harrison, Enhancing mobile voice assistants with worldgaze, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 1–10. URL: https://doi.org/10.1145/3313831.3376479. doi:10.1145/3313831.3376479.
[25] R. Schweigert, V. Schwind, S. Mayer, Eyepointing: A gaze-based selection technique, in: Proceedings of Mensch Und Computer 2019, MuC '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 719–723. URL: https://doi.org/10.1145/3340764.3344897. doi:10.1145/3340764.3344897.
[26] K. Ahuja, A. Kong, M. Goel, C. Harrison, Direction-of-voice (dov) estimation for intuitive speech interaction with smart devices ecosystems, in: Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology, UIST '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 1121–1131. URL: https://doi.org/10.1145/3379337.3415588. doi:10.1145/3379337.3415588.
[27] G. D. Abowd, E. D. Mynatt, Charting past, present, and future research in ubiquitous computing, ACM Transactions on Computer-Human Interaction 7 (2000) 29–58. URL: https://dl.acm.org/doi/10.1145/344949.344988. doi:10.1145/344949.344988.
[28] J. Mumm, B. Mutlu, Human-robot proxemics: Physical and psychological distancing in human-robot interaction, in: Proceedings of the 6th International Conference on Human-robot Interaction, ACM, Lausanne Switzerland, 2011, pp. 331–338. doi:10.1145/1957656.1957786.
[29] S. Daronnat, L. Azzopardi, M. Halvey, M. Dubiel, Impact of agent reliability and predictability on trust in real time human-agent collaboration, in: Proceedings of the 8th International Conference on Human-Agent Interaction, 2020, pp. 131–139.