A new framework to design explainable AI for augmented reality applications

May 8 2023, by Ingrid Fadelli

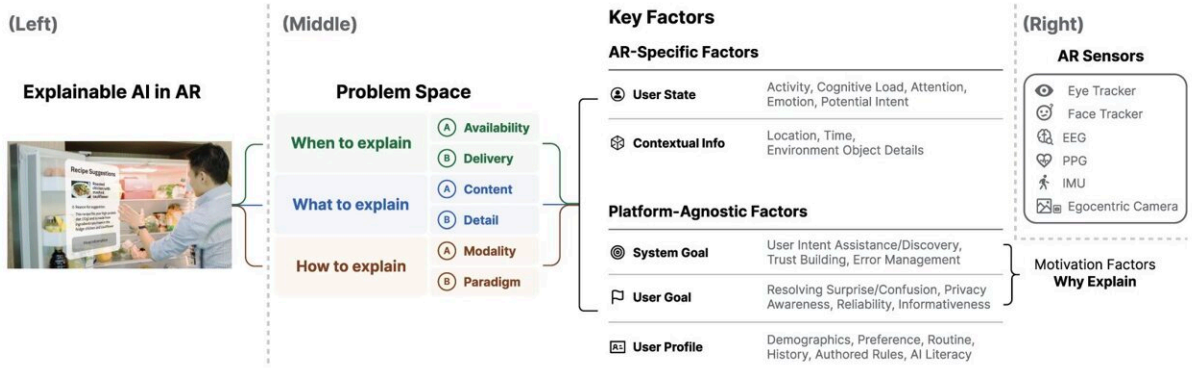


Figure 1: An Overview of XAIR Framework. (Left) An example of the AR interface with explanations. (Middle) The main structure of XAIR: the problem space and the key factors. (Right) Sensors that are integrated into AR.

An overview of XAIR. Credit: Xu et al.

While artificial intelligence (AI) and machine learning tools are now commonly used to enhance technological applications, the underpinnings of many of these tools are hard to decipher. This is because most of them are based on 'black box' models, models that analyze data and learn to make predictions about it but that do not share the processes behind these predictions with human users.

Researchers at Meta Reality Labs recently created XAIR, a framework that could help developers to make the processes underpinning the predictions of AI easier to understand. This framework, introduced in a

paper presented and published as part of the *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, is specifically designed to create explainable AI (XAI) systems that can be applied in various augmented reality (AR) settings.

"As black-box models are increasingly being employed in daily life, we are having more and more concerns about humans misusing AI and losing control," Xuhai Xu and Anna Yu, two of the researchers who carried out the study, told Tech Xplore. "This has led to the need to make algorithms easier to understand, leading to the surge of XAI. Existing studies have found that XAI can help end-users resolve confusion and build trust. Therefore, industrial practitioners have tried to use XAI to improve user experiences."

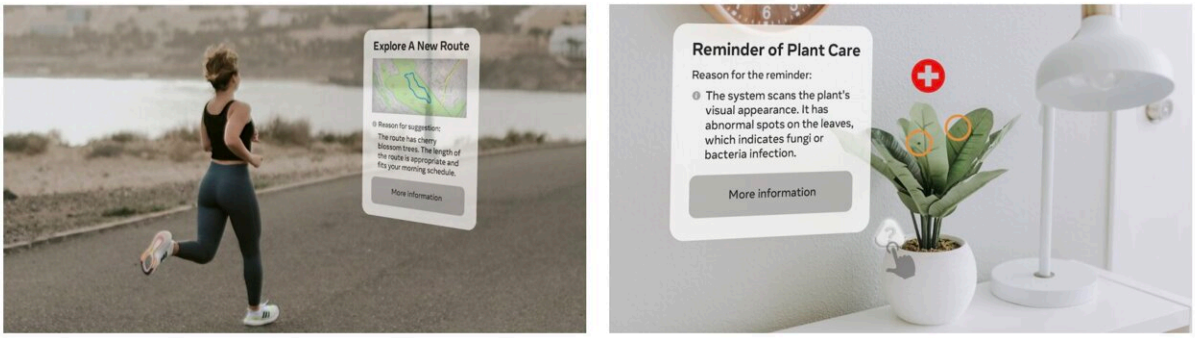
AR technology allows users to view a modified version of their surrounding environment, which integrates digital elements, sounds and/or visual enhancements. These 'digitally enhanced' versions of reality can be viewed through head-mounted displays, goggles, other wearable gear and even simply through the smartphone screen.

Recently, some researchers have been exploring the use of AI to enhance AR applications, for instance making them more responsive to changes in a user's environment or allowing them to analyze and make predictions about specific objects. Xu, Yu and their colleagues set out to create a framework that could make the results of these AI tools for AR applications easier to understand, thus increasing users' trust in them.

"As AI models will be needed for context-aware, everyday AR, XAI will also be essential because end-users will interact with all kinds of AI outcomes," Xu and Yu said. "XAI can be useful in many ways, such as making intelligent AR behavior interpretable, resolving confusion or surprise for unexpected AI outcomes, promoting privacy awareness, and building trust. Given the importance of XAI for AR, we aim to answer

the research question about the right way to create effective XAI experiences for AR in everyday scenarios."

The team at Meta created the XAIR framework hoping that it would facilitate the design of XAI for AR applications. Their framework essentially addresses three open questions: when, what and how? Answers to these questions can be used to provide more effective explanations for AI predictions in AR scenarios. In addition to helping developers to create AI that can answer these three questions, XAIR outlines a series of key guidelines for researchers and developers working on XAI for AR applications.



(a) Scenario 1: Route Suggestion when Jogging

(b) Scenario 2: Plant Fertilization Reminder.

Two good examples of the design outcome using our framework. Credit: Xu et al.

"We identified five key factors based on a large-scale literature review," Xu and Yu explained. "These factors determine the design of the 'when, what, how' aspects, including two AR-specific factors—user state and contextual information, and three non-AR-specific factors—system goal, user goal, and user profile."

Essentially, to use the team's framework, developers first need to address these five factors, pin-pointing contextual information about their users, their users state, the goal of their overall system, as well as the goal and profile of potential users. Once they did this, they can simply refer to the XAIR framework to adapt and polish their XAI systems for AR applications.

"As AI is becoming more and more powerful, in the foreseeable future, we can anticipate AI to help automatically identify at least a subset of these five factors," Xu and Yu said. "So that the framework can be turned into an automated or self-automated tool to help designers improve their design of XAI in AR."

As part of their study, Xu, Yu and their colleagues summarized the findings of over 100 studies rooted in different fields to identify important dimensions that one should consider when developing XAI for AR uses, answering when, what and how questions. Subsequently, the researchers conducted a large-scale survey involving over 500 users and held a workshop with 12 experts in the field. The survey responses and views shared by experts during the workshop offered valuable insight that guided their development of XAIR.

"XAIR is the first framework for XAI design in AR scenarios, also including guidelines to support designers in their design thinking process," Xu and Yu said. "The results of design workshops with 10 designers indicated that XAIR could provide meaningful and insightful creativity support for designers. We further implemented a real-time AR system based on one design and tested it with 12 end-users."

To evaluate the value of their framework, the researchers used it to create a real XAI system and tested it in real-time in a series of AR scenarios. They found that users perceived this system as both transparent and trustworthy, suggesting that their framework had

effectively guided its development. In the future, the XAIR framework could be used to create a variety of AI systems to enhance AR applications, which can explain their predictions and are thus perceived as more trustworthy by users.

"In our next studies, we plan to explore the automation of the design framework, the creation of personalized XAI experience in AR over time, and the enabling of users to provide feedback to further improve the system, etc.," Xu and Yu added. "The XAIR framework creates the foundation and guidelines for us to explore XAI interaction in future AR systems. The recent explosion of generative AI is also very exciting, and we would be interested in exploring how this trend can influence future XAI design in AR."

More information: Xuhai Xu et al, XAIR: A Framework of Explainable AI in Augmented Reality, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023). DOI: 10.1145/3544548.3581500