

MEDIA RELEASE

RELEASE DATE: 9 October 2026

Helping children question AI. An interactive voice-based chatbot that explores trust, accuracy and advice

An interactive voice-based chatbot developed by researchers invites children to explore what happens when AI chatbots agree with their ideas and helps them consider the risks when artificial intelligence prioritises agreement over accuracy.

As children increasingly turn to artificial intelligence for information, advice and support, researchers are exploring how the way chatbots respond can shape young people's understanding of these technologies.

A new voice AI chatbot, developed through collaboration between researchers at the [ARC Centre of Excellence for Automated Decision-Making and Society](#) (ADM+S), the [ARC Centre of Excellence for the Digital Child](#) and the [QUT Gen AI Lab](#), invites users to compare two different chatbot behaviours by switching between an "always agree" mode and a mode designed to respond more carefully and thoughtfully.

The experience explores 'AI sycophancy', a known problem in which AI systems flatter or agree with users, prioritising likeability over accuracy, critical thinking and sound judgement.

The interactive experience is based on the voice-based chatbot developed for research with children, extending on the pilot project [Making AI Friends? Enhancing Primary Years AI Literacy with a Voice AI Chatbot Experience](#), which explored how children can develop critical capabilities for understanding AI.

Dr Henry Fraser, a QUT researcher and ADM+S Associate Investigator, said that AI safety and governance needs bottom-up approaches, alongside top-down regulation and enforcement.

"Children's critical AI capabilities are an essential part of AI safety and governance. Regulation, like Australia's proposed digital duty of care, is important, but it can't fix everything. What we can do is enhance children's preparedness for and resilience to AI risks in a way that takes their experiences and perspectives seriously" Dr Fraser said.

The project comes amid growing attention to how young people use AI for personal advice and support. [Recent reporting](#) on research involving Adelaide University and the Australian Federal Police found that 35 per cent of children who had experienced online sexual victimisation either turned to an AI chatbot for help or told no one.

The findings underline the importance of helping young people to understand the limitations of AI systems and recognise when they need support from a trusted person.

Voice interfaces are increasingly appearing in toys and other smart devices used by children and families, making AI literacy an important part of understanding the digital environments children encounter.

The interactive experience encourages users to reflect on the differences between chatbot responses, why a system might be designed to agree with its users, and whether agreement should be mistaken for trustworthy advice.

Professor Tama Leaver, a Chief Investigator in the Digital Child, emphasised the importance of these interventions in building children's AI literacy and sense of agency.

"With the right support, young children have very clear ideas about what AI should be able to do for them, and with their data, and what would be unacceptable," said Professor Leaver.

"Encouraging these literacies helps them grow as citizens in our inescapably digital world."

The project reflects the importance of involving children in conversations about the technologies increasingly shaping their lives. By exploring different chatbot behaviours and discussing what better responses might look like, children can contribute their perspectives to how AI is understood and designed.

The research underpinning the interactive chatbot will be presented at the Association of Internet Researchers (AoIR) Annual Conference in Mexico in October, contributing to international discussions about children's experiences with AI and the importance of developing AI literacy from an early age.

The researchers hope this work will support conversations among children, families and educators about both the possibilities and limitations of conversational AI, including the importance of questioning responses, recognising when AI may be wrong, and seeking help from a trusted person when needed.

Try the interactive voice-enabled chatbot and spot the sycophant on the [GenAI Arcade website](#)

The chatbot is intended for supervised educational use. Children should ask a trusted adult before using it and should not follow chatbot advice without checking with a trusted adult.

Read more about the research [Should I hide my mum's slippers? A design-based pilot study of children's critical engagement with AI sycophancy using a custom AI chatbot](#) by Henry L Fraser, Jean Burgess, Kristy Corser, Susan Danby, Tama Leaver, William He, Katherine Nickels, Suzanne Srdarov and Irina Mils-Homens Figueira Dos Santos Silva.

Media contact:

Kathy Nickels, Communications and Engagement Manager
ARC Centre of Excellence for Automated Decision-Making and Society

Mobile: 0433 431 550
katherine.nickels@gut.edu.au





ABOUT THE ARC CENTRE FOR EXCELLENCE FOR AUTOMATED DECISION-MAKING AND SOCIETY

The ARC Centre of Excellence for Automated Decision-Making and Society (ADM+S) is a cross-disciplinary, national research centre, which aims to create the knowledge and strategies necessary for responsible, ethical, and inclusive automated decision-making (ADM).

Funded by the Australian Research Council from 2020 to 2027, ADM+S is hosted at RMIT in Melbourne, Australia, with nodes located at seven other Australian universities, and partners around the world. The Centre brings together leading researchers in the humanities, social and technological sciences in an international industry, research and civil society network.

Our Centre aims to contribute to the mitigation of the social and economic risks in the development and implementation of ADM and artificial intelligence (AI), and to improve outcomes and efficiencies in four key focus areas where these technologies are already well advanced: news and media, transport and mobility, health care, and social services.

Available for interview

[Distinguished Professor Jean Burgess](#), Associate Director, ARC Centre of Excellence for Automated Decision-Making and Society at QUT.

[Dr Henry Fraser](#), Research Fellow, ARC Centre of Excellence for Automated Decision-Making and Society at QUT.