

# From stage to society: An arts-based approach to open research and responsible AI

Poster Presentation | 04. May 2026 | FOR2026 Conference, Munich

## Overview: Rationale and Objectives of the MoralPLai Project

Generative AI, especially chatbots powered by large language models, enables **individuals to increasingly rely on automation to support their decisions** by answering questions, offering information or even providing concrete moral advice and emotional support in high-stakes decision contexts (Aharoni et al., 2024; Meng & Dai, 2021). For example, ChatGPT shows promise in improving triage accuracy in emergency departments (Kaboudi et al., 2024) and can serve as a personal therapist by delivering mental health support to individuals in real-time (Alanezi, 2024). Although AI chatbots may help address shortages in medical care and alleviate feelings of loneliness (Syed, 2024), their use also entails substantial risks, as demonstrated in court cases that connect chatbot interactions to suicide incidents (e.g., Yousif, 2025). Considering the profound societal impact of AI chatbots, **it is important to inform the public and involve them in relevant ongoing scientific research**, an imperative underscored by open science agendas (Ramachandran et al., 2021) and recognized by regulatory initiatives (e.g., Pact for Research and Innovation IV (2021–2030); EU AI Act). However, **traditional science communication practices face various challenges, such as the often difficult-to-understand technical language** in academic publications, which can reduce the public's interest in scientific work (Dietrich et al., 2024). To make scientific debates accessible beyond the academic community, **innovative methods and mediums are needed through which civil society can stay informed about up-to-date research and engage with scientists** (Weber et al., 2021). In this endeavor, arts – an important reference for social knowledge and inclusion – can become a key enabler to ensure human-centric, participatory discussions around the design of pertinent technological innovations (Guryanova et al., 2019). More specifically, arts-based methodologies such as research-based theatre, ethnodrama or the playwright-approach “combine research and theater to create novel opportunities for inquiry and knowledge translation” (Nichols et al., 2023; p.1).

Adapting this alternative format to open research, the *MoralPLai* Project uses a **creative, arts-based approach to explore and communicate AI Ethics research**. At its core is *The Third Voice*, a **research-based theatre play that examines the ethical and emotional challenges of relying on AI chatbots for decision-making**. Based on expert interviews, the play portrays a doctor and her daughter who both turn to a chatbot for guidance in a morally complex situation, highlighting the potential benefits and risks of AI chatbots as moral dialogue partners. By engaging the audience in accessible ways, the performance aims to promote AI literacy and gather public insights that can guide the design of human-centered AI. This poster outlines the MoralPLai Project, summarizing: (1) the narrative of *The Third Voice* and its debut in 2025, (2) the utilized methodologies for audience engagement, and (3) the corresponding feedback findings. Overall, the MoralPLai Project aims to lay the groundwork for **transforming scientific communication and inquiry in AI Ethics, drawing on humanities and cultural sciences as vital methodologies for bridging academia and society and fostering AI literacy**.

## The Third Voice: A Research-Based Theatre Production

The narrative of *The Third Voice* was developed on the basis of scientific research, including reviews of literature on AI chatbots and moral guidance, regulatory documents, and corporate responsible AI guidelines, as well as original semi-structured interviews with academic experts.

*The Third Voice* tells the story of a **doctor who goes on trial after relying on an AI chatbot**, AITHONA, to advise a terminally ill patient considering whether to stop treatment. Simultaneously, her **lonely teenage daughter turns to the same chatbot for emotional support**. The play examines key issues such as echo chambers and feedback loops, framing effects, moral deskilling, the manipulation of content filters, the diffusion of responsibility, and AI hallucinations. It also highlights potential benefits, including the ability of AI chatbots to provide quick, synthesized information, offer support at any time when human help is unavailable, and encourage reflection.

Moving between scenes in a hospital, courtroom, and family home, the play explores how **AI chatbots can both support and weaken human ethical decision-making**, depending on how carefully and critically they are used.

*The Third Voice* premiered on May 22, 2025, at the **Amerikahaus in Munich**, followed by a panel discussion and an open Q&A session to further engage the audience with the play's themes.



The Third Voice Debut, 2025, Amerikahaus Munich  
Photo Credits: Hanna Genschler

## Findings: Key Insights from Audience Feedback

87%

agree that the play helped them **better understand** the potential consequences and diverse perspectives of using AI chatbots as moral dialogue partners or advisors.

93%

agree that the play succeeded in **raising awareness** of the complexities of using AI chatbots for moral guidance and emotional support.



### Reflections on the Theatre Play

- The play was described as intense, **emotionally engaging**, and timely.
- Many noted that the **characters felt highly relatable**, facilitating connections to real-world issues and enabling **engagement with diverse perspectives**.
- The play sparked curiosity about AI and moral decision-making, **motivating audiences to explore relevant research further** and share insights with others beyond the performance.



### Reflections on the Format of the Event

- Theatre was seen as a **powerful format for exploring complex topics** and provoking meaningful, **inclusive discussion around scientific insights**.
- Suggestions included maximizing impact through **targeted outreach** and **supporting materials** (e.g., readings, glossaries).



### Identified Challenges & Risks

- Informational **manipulation**, ideological entrenchment
- Controlling **access for minors**
- Bias** in AI chatbots and complexity to detecting bias
- Inability for companies to control user prompts, **jailbreaking**
- Lack of transparency** about how chatbots make decisions
- Over-reliance** and erosion of critical thinking



### Takeaways for Personal Use

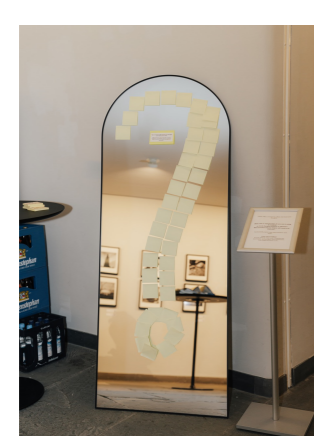
- Request chatbots to ask questions** instead of providing answers
- Request chatbots to **provide diverse perspectives**
- Don't use AI for moral guidance, **human guidance** instead (or at least complementary)
- Stay critical, **verify information**, and reflect on responses

Icon Credits: Generated using ChatGPT (GPT-5.3)

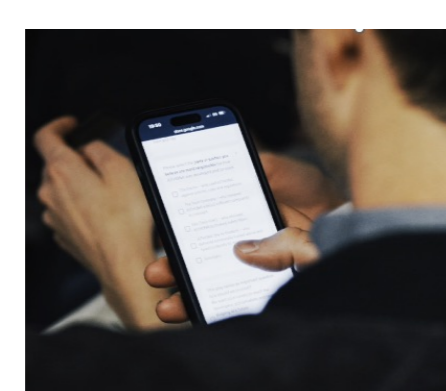
## Methods for Capturing Audience Responses

May 2025

December 2025



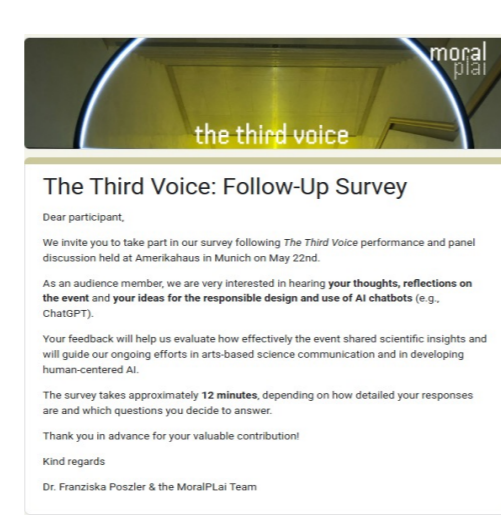
Post-it notes, in foyer at debut event



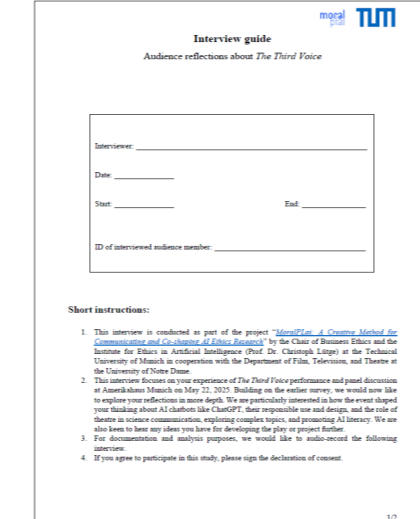
Post-performance poll, right after play (n=141)



Team debriefing session, one day after event



Post-performance survey, a few weeks after event (n=44)



Post-performance interviews, a few months after event (n=21)

### References

- Aharoni, E., Fernandez, E., Brady, D. J., Alexander, C., Grier, M., Queen, K., ... & Orrego, V. (2024). Attributions toward artificial agents in a modified Moral Turing Test. *Scientific reports*, 14(1), 8498.
- Alanezi, F. (2024). Assessing the effectiveness of ChatGPT in delivering mental health support: a qualitative study. *Journal of interdisciplinary healthcare*, 48(1-2).
- Dietrich, A., Lopez, R., & Spelt, L. (2024). Die Wissenschaftsbotschaft: einen freien Vortrag über die Risiken von Wissenschaftskommunikation. *The Inequality Mind*.
- Chapman, A. V., & Bracken, V. (2019). Transformation of audience expectations in the digital era: towards an arts, social and health perspective. In *International Scientific Conference: Digital Transformation of the Economy: Challenges, Trends, New Opportunities* (pp. 47-53). Cham: Springer International Publishing.
- Kaboudi, N., Fardipour, S., Ebrahimi, M. S., Fayazbakhsh, F., Jahanmoussalorani, N., Ghaderi, S., ... & Shariati, R. (2024). Diagnostic accuracy of ChatGPT for patient triage: a systematic review and Meta-Analysis. *Archives of Academic Emergency Medicine*, 12(1), e60.
- Meng, J., & Dai, Y. (2021). Emotional support from AI chatbots: Should a supportive partner and listener be left? *Journal of Corporate-Marketing Communication*, 27(4), 207-222.
- Nichols, J., Cox, S. M., & Cuddeback, M. (2023). World tips and guidelines: Identifying and mitigating ethical lessons in research-based theater. *Qualitative Inquiry*, 29(2), 257-266.
- Ramachandran, R., Bagdasarian, K., & Marathe, K. (2021). From open data to open science. *Earth and Space Science*, 8(5), e20200401762.
- Spelt, L. A. (2024). The role of AI in addressing loneliness among adults in the United States. *International Journal of Engineering, Technology Research & Management (IJETRM)*, 8(04), 404-421.
- Weber, C., Allen, S., & Neubauer, N. (2021). Scaling learning to support scientists to engage with the public in non-traditional venues. *Journal of Science Communication*, 20(4).
- Yousif, N. (2025). Parents of teenager who took his own life sue OpenAI. *BBC News*. Available at: <https://www.bbc.com/news/technology-67222222>

**Funding**  
The MoralPLai project team gratefully acknowledges financial support from the Notre Dame-IBM Tech Ethics Lab, the Friedrich-Schiedel-Fellowship Program at the TUM Think Tank and the TUM School of Social Sciences and Technology, the TUM Global Incentive Fund, the Tuition Substitution Funds, the Klaus Tschira Foundation gGmbH, TUM Center for Culture and Arts (CCA) as well as the TUM Global Visiting Professor Program.

**Technical University of Munich**  
TUM School of Social Sciences and Technology  
Institute for Ethics in Artificial Intelligence  
Principal Investigator & Project Lead: Dr. Franziska Poszler ([franziska.poszler@tum.de](mailto:franziska.poszler@tum.de))  
For more information about the MoralPLai Project, please visit: <https://www.ieai.tum.de/research/moralplai/>