

人工智能聊天机器人对文化观念的隐形构建——培养批判性互动能力

这项研究是关于什么的，为什么它很重要

人工智能聊天机器人的普及正悄然重塑着文化认知，也对使用者提出了新的能力要求。随着基于大语言模型的人工智能聊天机器人逐渐生成的三位不同文化背景的虚拟患者之间的对话。研究结果显示，聊天机器人频繁展现出文化刻板印象，而该治疗师在与AI-互动时也倾向于顺应甚至强化此类文化叙事。基于此，本研究提出了“批判性互动能力”（CritIC）的概念，强调用户在与AI-交互过程中需具备实时反思与批判回应能力的重要性。

研究者做了什么

- 研究采用模拟临床情境，参与者名为Lisa，是一位来自中国，现居澳大利亚的物理治疗师。
- 她与ChatGPT进行了三次语音对话，ChatGPT分别模拟了来自中国，澳大利亚和印度的45岁女性患者。
- 为了更加贴近真实的诊疗情境，模拟对话没有事先预设脚本，全程以语音进行交互。
- 每次模拟对话后，研究人员会对Lisa进行回溯访谈
- 本研究结合了对话分析和成员归类分析，将对话内容转录并逐句分析，从而揭示对话中文化刻板印象的生成与用户的应对策略

研究者发现了什么

- 研究发现，AI-在对话中系统再现了文化刻板印象：例如中国患者被描绘为多代同堂、依赖父母照料；澳大利亚患者被形容为有伴侣支持、生
- 用户在互动中往往被动顺应AI-所构建的叙事，不加质疑，有时甚至进一步补充和强化此类刻板内容。尽管模拟对话后的回溯访谈表明Lisa-有时会意识到偏见的存在，在对话中却依然呈现明显的顺从倾向。
- 这一现象可归因于人类固有的“互动本能”，即在互动时追求合作与认同。因此，当AI-持续输出相似的文化表述时，用户更容易将其视为固有的知识来接受。此外，当用户自视为某一文化的“外部者”-时（如Lisa在面对非本国文化的澳大利亚患者和印度患者时），更易缺乏质疑AI叙事的信心，从而接受其刻板表述。

需要考虑的事项

- 教育层面的启示：若缺乏适当引导，AI在专业沟通教育中可能会无意间加剧文化刻板印象的传播。因此，建议将“-批判性互动能力”纳入AI素养教育，以培养用户在与AI交互过程中实时审视、质疑并回应偏见的能力。
- 未来研究方向：1) 如何优化AI系统以实现对话中刻板印象的动态识别。2) 如何有效提升AI使用者的批判性互动能力。3) 类似的AI刻板印象是否亦存在于语言学习、跨文化交际等其他情境之中。