

ABSURDIST CHATBOTS

In an era dominated by technological solutionism, artificial intelligence (AI) is often seen as a quick fix for complex problems. Inspired by the surrealist notion of pataphysics, this graduation project employs a novel, critical design approach to address the imaginary problem of trustworthy AI and its solutions in the context of generative chatbots.

This project presents three pataphysical chatbots, each addressing a specific requirement—transparency, human control, or fairness—currently implemented in AI systems through technical mechanisms intended to ensure trust. Each chatbot is designed to deliberately exaggerate the mechanism behind the assumed requirement, aiming to help users understand the absurdist nature of the chatbot and, ultimately, provoke critical reflection on the pursuit of trustworthy AI.

Through an iterative design process of prototyping, testing, and reflecting upon pataphysical chatbots, progress was made toward the final pata-design, which was evaluated in an online experiment. A key insight emerged: while the pataphysical chatbots succeeded in conveying their absurdist nature to users, critical reflection on the pursuit of trustworthy AI did not occur. However, additional findings suggest a potential link between understanding the absurdity and the occurrence of this particular type of critical reflection.

Ultimately, this project contributes to the critical design practice and in particular the field of pataphysical design, by demonstrating how pataphysical chatbots can effectively challenge both the prevailing methods for achieving trustworthy AI as the dominant problem-solution paradigm of AI solutionism. Moreover, the pataphysical design approach proposed in this project offers a new direction for current pataphysical design practices, emphasizing the creation of engaging pataphysical experiences that aim to provoke specific user responses to the premises of the imaginary problem.