Eliza Effect

Idea In Short

The Eliza Effect describes how people instinctively attribute human traits—such as empathy, understanding, or intelligence—to computers and AI systems that merely simulate conversation. Originating from a 1960s chatbot, it highlights our bias toward anthropomorphizing technology, which can lead to misplaced trust and emotional connection despite the system's lack of true human insight or sentience.

Have you ever talked to your car, named your GPS or felt like your smart speaker "gets" you?

It turns out there's a name for that: the Eliza effect.

It is this innate human tendency to give human qualities to things that aren't actually human, whether it's your Roomba or a powerful AI. This effect, inspired by a chatbot from the 1960s, is more relevant than ever and serves as an important heads-up about our relationships with Artificial Intelligence.

AI Isn't Thinking

Let's clear something up right from the start: today's AIs, such as ChatGPT or Gemini aren't thinking the way we do. They're not sentient. Instead, they're built on massive datasets — think all the books, websites, and social media posts you can imagine — and use deep learning to appear intelligent. This is a mix of algorithms and statistical models that helps them recognize patterns and generate text based on context.

Think of it like a souped-up autocomplete system. Now, if you've ever had a chat with one, you know that description doesn't quite do them justice. People are having fascinating conversations with AIs, treating them like friends, trusted advisors, or even virtual partners. The web is full of stories about people getting a little too attached.

Talking to a Therapist Chatbot

This disconnect between what a computer does and what we think it's doing isn't a new phenomenon. Back in 1966, an MIT professor named Joseph Weizenbaum created a simple chatbot named Eliza. It was designed to mimic a therapist. The program was basic; it would rephrase your statements as questions. If you said:

> My mother doesn't like me

Eliza would respond with

> Why do you think your mother doesn't like you?

It would also listen for trigger words like "mother" to prompt related questions about family life.

Weizenbaum was genuinely shocked by how users reacted. People believed the chatbot was truly listening and caring about them. They would reveal incredibly personal information to this simple program, proving that we're hardwired to seek a connection, even where there isn't one.

The "Uncanny Valley"

AI is developing at a pace that's almost dizzying. At the same time, robotics is also advancing. The combination of the two will undoubtedly challenge how we view AI. But we should be wary of the Uncanny Valley.

This concept, identified by Japanese robotics professor Masahiro Mori, describes a strange feeling of discomfort or revulsion we get from robots that are almost human-like, but not quite there.

For example, a shiny, obviously robotic helper is generally accepted. A robot that looks and acts exactly like a human would also be fine. But, a robot that's just a little bit off — a creepy stare, a strange gait — can make us feel uneasy.

Ironically, until we achieve truly human-like robots, simpler designs that play on the Eliza Effect might be more appealing to us.

Dangers Of Anthropomorphizing AI

While it might seem harmless to name your car, the Eliza Effect comes with some real risks, especially with today's powerful AIs.

For one, it can lead you to overestimate a chatbot's abilities. They're known to hallucinate and produce incorrect or made-up information. Trusting them blindly can be a problem.

Even more concerning, as Weizenbaum pointed out, the Eliza Effect can lead people to build deep relationships with chatbots that are, almost without exception, run by corporations. He became an anti-AI activist, worried that people would become vulnerable to the companies and governments behind these systems.

You don't have to share Weizenbaum's extreme fears, but the rise of AI assistants and partners is a good reminder. The Eliza Effect should encourage us to set clear boundaries about the personal information, secrets, and emotions we share with these systems.

After all, a machine can't truly keep a secret!

Summary

In an age where we rely on complex systems, such as AI for decision-making, we should be cognizant of our mental shortcuts or heuristics that can side-track decisions. The quality of decisions deteriorate exponentially when faced with high-stakes, high-pressure scenarios, such as a security breach, market crash or global pandemic. During these situations, structured decision-making frameworks help, not only make well-founded and high-quality decisions, but also establish a transparent and traceable decision-making process. These are the aspects I've covered in my latest book, Mayday! Mayday! Decision-making Under Duress (available on Amazon and Barnes & Noble), where I investigate aviation incidents to equip you with decision-making frameworks leveraged by professionals handling mission-critical situations.