



&lt; Magazine: 29 July 2023

*Webb Keane and Scott J. Shapiro*

# Deus ex machina: the dangers of AI godbots

From magazine issue: 29 July 2023



Photo-illustration: Coral Hoeren (iStock)



Text



Comments



Something weird is happening in the world of AI. On Jesus-ai.com, you can pose questions to an artificially intelligent Jesus: 'Ask Jesus AI about any verses in the Bible, law, love, life, truth!' The app Delphi, named after the Greek oracle, claims to solve your ethical dilemmas. Several bots take on the identity

of Krishna to answer your questions about what a good Hindu should do. Meanwhile, a church in Nuremberg recently used ChatGPT in its liturgy – the bot, represented by the avatar of a bearded man, preached that worshippers should not fear death.

Elon Musk put his finger on it: AI is starting to look ‘godlike’. The historian Yuval Noah Harari seems to agree, warning that AI will create new religions. Indeed the temptation to treat AI as connecting us to something superhuman, even divine, seems irresistible.

New godbots are coming online – and we should be afraid. They raise two serious concerns: first, they are powerful tools that bad actors can use to victimise others; second, even when well-intentioned, these bots trick users into surrendering their autonomy and delegating ethical questions to others.

*The temptation to treat AI as connecting us to something superhuman seems irresistible*

It’s important to understand what is novel about godbots. Bots such as ChatGPT which employ a relatively new technology, LLM (Large Language Models), are trained on enormous amounts of data and capable of performing astonishing tasks. At the same time, these apps are not new in their ability to satisfy the human desire for answers in times of uncertainty. They exploit our tendency to impute divinity to inexplicable processes by speaking in certainties. Our response to AI is strikingly similar, therefore, to how humans have always reacted to the power and inexplicability of divine and godlike figures – and, more specifically, to the ways we try to get the gods to talk to us.

Let’s begin with a central feature of the new AI – its ‘unexplainability’. Algorithms trained by machine learning can give surprising and unpredictable answers to our questions. Even though we built the algorithm, we don’t know how it works. It seems as if the thing has a mind of its own.

But the problem, more precisely, is that we can’t explain how the algorithm gets the results it does and that any explanation we can give is at least as long and as complicated as the algorithm we are trying to explain. Like a map that’s so detailed it ends up the same size as the territory it’s supposed to depict, repeating the algorithm

simply doesn't help us grasp where we are. We haven't got a proper explanation. As a result, the workings of the device seem in-effable, uninterpretable, inscrutable.

When such ineffable workings produce surprising results, it seems like magic. When the workings are also incorporeal and omni-scient, it all starts to look a lot like something divine. God's reason is indescribable – mysterious and beyond human comprehension. By these measures, the AI in GPT-4 seems godlike. We may never be able to explain why it gives the answers it does. The AI is also body-less, an abstract mathematical entity. And if not utterly omniscient, the algorithm has access to more information than any human could ever know.

GPT-4 seems to open a conduit to something truly potent. And so we want it to answer our hardest questions. One of the most widespread techniques human societies have used for seeking answers from a god is divination, trying, for example, to read the movements of birds or casting lots.

Traditionalists on the Indonesian island of Sumba practise several kinds of divination, such as reading the entrails of animals and certain operations similar to casting lots. But they don't do this for ordinary technical or agricultural questions. They turn to divination when they face uncertain and troubling circumstances, especially moral or political ones. And they don't try to explain how divination works – it remains opaque, which is surely part of its power.

Divination was common in the ancient world. Cicero, a skeptic, worried that divination would be used by charlatans to manipulate users and trick them into doing their bidding. It is not hard to see how the same dynamic will play out with godbots. Malicious actors need only insert code that makes the godbot respond in the way that the malicious actor wants. On some existing apps, Krishna has already advised killing unbelievers and supporting India's ruling party. What's to stop con artists from demanding tithes or promoting criminal acts? Or, as one Chatbot has done, telling users to leave their spouses? When God asks you to do something, you don't say no.

It's easy to think there is a gulf between our scientific, secular present and the benighted, superstitious past. But in fact, those who are not religious may still retain, in secular form, a yearning for magic under the stress of uncertainty or loss of control. When the magic consists in interpreting the words of an all-knowing, unimaginable and disembodied device – something like GPT-4 – it can be like talking to a god.

AI chatbots tap into our desire for magical thinking by speaking in certainties. Even though their large language models employ sophisticated statistics to guess the most likely response to a prompt, the bot replies as if there is just one answer. They imply there's nothing more to discuss. Bots won't tell you their sources or offer links inviting you to consider alternatives. They will not or cannot explain their reasoning. Where once the gods spoke through entrails, now there's an app.

What can we do about this? Like divination, AI seems independent of its creators. But, like divination, it is not. Behind the hype, self-learning programs depend on human input. Even traditional diviners didn't take the signs at face value, they interpreted them. Their answers had input from human beings, even if disguised. When AI scrapes the web, it reflects back to us what *we* have put there. Our apps should show this. They should be required to present some of the evidence relevant to their decisions. This way users can see the artificial intelligence is drawing on *human* intelligence. Bots should not speak in absolutes and spurious certainties. They should make clear they are only giving probabilities. They should not speak like divine beings that magically *know* the answer.

AI is made by humans and can be constrained by humans. We should not give AI divine authority over human dilemmas. Humans are morally accountable for their actions. As agonising as some ethical issues are, we cannot outsource them to snippets of code. AI will only become godlike if we make it so.

WRITTEN BY

*Webb Keane and Scott J. Shapiro*

Webb Keane is a professor of anthropology at the University of Michigan. Scott J. Shapiro is a professor of law and philosophy at Yale Law School and the founding director of Yale Cybersecurity Lab.



Comments



TOPICS [Society](#) / [AI](#) / [God](#) / [Religion](#) / [Science](#)

Most popular