# AI models pass the Turing Test, showcasing advancements in conversational technology



Advancements in artificial intelligence (AI) have reached a notable milestone, with a recent study from the University of California in San Diego indicating that two prominent AI models have successfully passed the Turing Test, a measure of machine intelligence formulated by British mathematician and cryptanalyst Alan Turing in 1950. This study highlights the increasing sophistication of AI and its ability to closely mimic human conversational patterns.

The research, led by Cameron Jones from UC San Diego’s Language and Cognition Lab, involved testing four large language models (LLMs): GPT-4.5, LLaMa 3.1 (developed by Meta AI), GPT-4o, and a 1960s-era chatbot called ELIZA. Participants included 126 undergraduate students from UC San Diego and 158 individuals sourced from an online data pool via Prolific. They engaged in five-minute online conversations with both a human and an AI without knowing which was which—essentially acting as interrogators attempting to determine the identity of their interlocutors.

Results indicated that when prompted to assume a human-like persona, GPT-4.5 was identified as human 73 per cent of the time, significantly surpassing the random chance of 50 per cent. LLaMa 3.1 achieved a similar level of success when adopting a human-like persona, being judged as human 56 per cent of the time. However, older models like GPT-4o and ELIZA failed to pass the test, with success rates of only 23 per cent and 21 per cent, respectively.

Jones noted in a post on X that the study provides the first empirical evidence that artificial systems can pass a standard three-party Turing Test, stating, “If interrogators are not able to reliably distinguish between a human and a machine, then the machine is said to have passed.” He also remarked that the artificial models performed substantially better when given a prompt to act human-like compared to their efforts without such guidance, achieving only a 36 per cent win rate in that scenario.

The findings prompt discussions about the implications of AI's growing capability to emulate human interaction. Experts suggest that while these AI models can approximate human conversation convincingly, they do so not necessarily due to intrinsic intelligence but rather due to their capacity for emotional mimicry. John Nosta, founder of the innovation think tank Nosta Lab, posited, “This wasn’t a failure of AI detection. It was a triumph of artificial empathy.”

The implications of this research extend beyond academic interest; they highlight potential concerns over job automation and the possibility of AI being employed in social engineering attacks. With AI systems evidently capable of conducting brief interactions indistinguishable from humans, there are rising anxieties about their application in sensitive areas of society.

This latest development in AI evaluation comes against the backdrop of an ongoing dialogue about the intersection of technology and society, particularly as models continue to evolve and improve year by year, raising crucial questions about the future role of AI in daily life and its impact on our interactions with both machines and each other. The study is currently in a preprint stage and awaits peer review, as researchers continue to explore the implications of these findings.

Source: [Noah Wire Services](https://www.noahwire.com)

## References

* <https://arxiv.org/html/2503.23674v1> - This link supports the claims about the study where large language models, such as GPT-4.5 and LLaMa-3.1, passed the Turing Test by convincingly mimicking human conversations. It details the methodology and results of the testing.
* <https://english.mathrubhumi.com/features/technology/gpt-ai-pass-turning-test-humans-robot-1.10484831> - This article corroborates the findings of the study, highlighting that models like GPT-4.5 successfully passed a version of the Turing Test by being identified as human more often than chance or even real humans.
* <https://en.wikipedia.org/wiki/Turing_test> - This Wikipedia page explains the Turing Test, originally proposed by Alan Turing in 1950, which is used to assess a machine's ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human.
* <https://www.openai.com/> - This is the official website of OpenAI, the developer of GPT models, which provides context on the company's role in advancing AI capabilities like those discussed in the study.
* <https://www.meta.com/> - This link is to the official Meta AI website, relevant because it is the developer of the LLaMa model, another successful AI in the Turing Test mentioned in the study.
* [https://static1.squarespace.com/static/5f72a312c37e21261a2e9a64/t/62e257e9d3bb2d1b27988a7c/1659343505043/LaMa+Technical+Report.pdf](https://static1.squarespace.com/static/5f72a312c37e21261a2e9a64/t/62e257e9d3bb2d1b27988a7c/1659343505043/LaMa%2BTechnical%2BReport.pdf) - This link provides technical information about the LLaMa AI model developed by Meta AI, offering insight into the capabilities of such models.
* <https://nypost.com/2025/04/04/tech/terrifying-study-reveals-ai-robots-have-passed-turing-test-and-are-now-indistinguishable-from-humans-scientists-say/> - Please view link - unable to able to access data