

Chatbots are Learning to Speak Physics

Researchers are teaching AI models the language of particle physics, giving them the ability to compose novel theories. But human input remains essential.

By **Eleanor DeGoffau**

Just two years ago, the most advanced AI chatbots were failing at basic arithmetic. Now they are nearly fluent in the language of mathematics, generating innovative solutions to decades-old math problems. This rapid advancement has caught the attention of particle physicists who are developing interactive AI models that can communicate in the language of physics. In a series of recent studies, researchers have shown that physics-fluent AI can analyze data, construct theories, and recover particle properties.

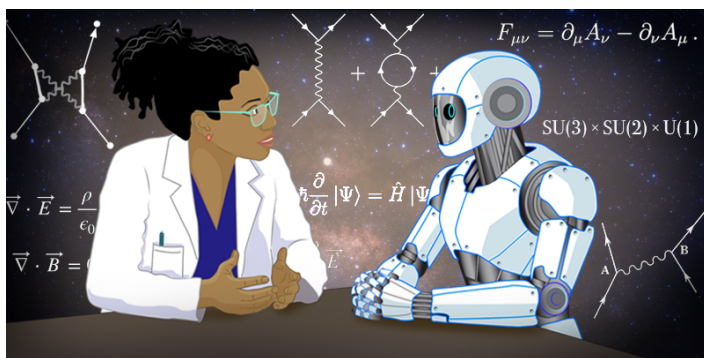
Chatbots can talk thanks to large language models (LLMs)—AI computer systems that are trained on large quantities of written content, allowing them to generate text using patterns observed in the training input. In the case of physics, researchers are exploring what happens if this input includes a physics knowledge base—such as experimental data and basic

theoretical principles. The physics-trained LLMs have been asked to solve problems and interpret observations, and in response they have produced novel computational methods and discovered hidden information. Although the progress is striking, the researchers stress that considerable human oversight is still needed.

Particle physicists, like many researchers, have been using generative AI for general-purpose tasks, such as manuscript editing and literature searches. But the community is now recognizing that recent advances in AI's physics-reasoning capabilities could potentially transform research [1]. Cosmologist Moritz Münchmeyer and his team at the University of Wisconsin–Madison witnessed this rapid growth firsthand [2]. In early 2025, the researchers began posing exam-style physics questions to a variety of chatbots to test their skills at physical reasoning. The questions ranged in difficulty from undergraduate to research level and included problems asking AI to calculate trajectories of rockets or simulate the distribution of galaxies.

The results were initially mixed: The top performing LLM, OpenAI's o1, only obtained the right answer 64% of the time. This score dropped to 18% in the research-level category. But when Münchmeyer and colleagues ran the same tests just over one year later, the LLMs showed dramatic improvements. OpenAI's GPT-5.5 had a total score of 90%, with a 50% success rate for research-level problems.

The problem-solving capabilities of LLMs could boost productivity in research. Given a list of research objectives, an LLM can now review the literature, analyze particle-collider



Looking for a boost to their productivity, physicists are instructing AI in the basics of physics-speak.

Credit: APS/Carin Cain

data, develop mathematical models, and write a scientific paper.

David Shih, a theoretical particle physicist at Rutgers University in New Jersey, wanted to find out how much he could rely on Claude Opus 4.6 as a research partner. Under Shih's supervision, Claude wrote code and trained machine-learning algorithms for complex mathematical tasks. "It was exhilarating and frustrating," Shih says. "It kept making mistakes, but it's insanely fast." The errors meant that Shih had to constantly check the output. But the results were impressive. In one case Claude was asked to simplify the loop integrals from Feynman diagrams, and it uncovered a new method that simplified complex integrals with less computer-memory requirements than current methods [3].

Other physicists, meanwhile, are using LLMs as interdisciplinary collaborators. Chiara Bissolotti, a nuclear physicist at Argonne National Laboratory in Illinois, and her colleagues developed a proof-of-concept chatbot named ArgoLOOM [4]. The chatbot uses an LLM model, GPT-4o, that has been given access to a curated base of papers and computer codes in three physics domains: cosmology, high-energy physics, and nuclear physics. This set of resources gives ArgoLOOM field-specific knowledge that generic chatbots lack. A cosmologist could, for example, chat with ArgoLOOM and ask it to produce a model of galaxy formation with inputs, such as hypothetical particles or novel nuclear interactions, coming from other domains.

Bissolotti doesn't expect ArgoLOOM to discover new physics, but it could help researchers along the way. "Maybe you get a new idea seeing how ArgoLOOM interpreted your request," she says. "It's a support for discovery, and it's a support for the human researcher."

While most physicists are making modifications to off-the-shelf LLM and chat-interface combinations like ChatGPT, some researchers are training their own LLMs from scratch. Inspired by ChatGPT's ability to write poetry, a team from Brown University in Rhode Island wondered if AI could write its own "physics poems." In other words, could an LLM combine basic physics principles into new explanatory models, just as a poet arranges words into creative verses?

The team, which included theoretical physicist Stephon

Alexander and PhD candidate Cooper Niu, trained the Autonomous Lagrangian Building and Exploration with RL-trained Transformer—ALBERT, for short—on the building blocks of quantum field theory, such as gauge symmetries and the behavior of fermions. Then, through reinforcement learning algorithms, they taught ALBERT to generate—on its own—physically consistent quantum theories on the basis of these fundamentals [5].

Finally, the team gave ALBERT experimental data from the now-decommissioned Large Electron-Positron Collider, along with physics knowledge available before 1990. They asked ALBERT to infer the properties of particles that were not discovered at the time. In under one hour, ALBERT successfully predicted the mass and other properties of the top quark, which was observed for the first time in 1995.

"There's a vast space of possible theories, trillions upon trillions," Alexander says. ALBERT can efficiently sift through this large space and find the best candidates that match the data. "We believe that what we have done is quite novel," he says.

Even as physicists are proud of the LLMs they've taught to "speak" physics, they're doubtful that AI will function without human input. "I think the LLMs, even if they continue to improve at their current rate, won't be able to replicate some of the things that make us human, and so we'll still have something to add," Shih says. Common sense and physical intuition are skills he's not sure they'll ever fully acquire. "I'd be surprised if they came up with anything really useful completely on their own. But, with a human directing them, maybe the sky's the limit."

Eleanor DeGoffau is a freelance science journalist based in Michigan.

REFERENCES

1. M. Naddaf, "How are researchers using AI? Survey reveals pros and cons for science," *Nature* (2025).
2. D. J. H. Chung *et al.*, "Theoretical physics benchmark (TPBench)—A dataset and study of AI reasoning capabilities in theoretical physics," *Mach. Learn.: Sci. Technol.* **6**, 030505 (2025).
3. D. Shih, "Learning to unscramble Feynman loop integrals with

- SAILIR,” [arXiv:2604.05034](#).
4. S. D. Bakshi *et al.*, “ArgoLOOM: Agentic AI for fundamental physics from quarks to cosmos,” [arXiv:2510.02426](#).
 5. S. Alexander *et al.*, “Autonomous discovery of particle physics theories from experimental data,” [arXiv:2603.28935](#).