

Across Campus, Princeton Explores the Use of Large AI Models for Research

Allison Gasparini

A crowd of 200 students, faculty and staff gathered in the Friend center auditorium on Nov. 28, for the Princeton Language and Intelligence initiative's first symposium since [launching in September](#).

"The goal today is to showcase how researchers are using or plan to use large AI models to further their research," said Sanjeev Arora, Princeton's Charles C. Fitzmorris Professor in Computer Science and Director of Princeton Language and Intelligence initiative, in his welcoming to the crowd who filled out the lecture hall. "The other goal," he added, "is that several of today's presenters are looking for collaborators." He stressed that PLI is devoted to promoting collaboration across disciplines as well as enhancing the broader community's fundamental understanding of AI.

A Multidisciplinary Approach

At the symposium, Princeton faculty members representing 8 different departments gave ten presentations exploring how large language models enhance or inform their own research. ChatGPT is just one example of a large language model, systems which "learn" by being fed trillions of words accessed around the internet.

Each meant to last just five minutes long, the lightning talks provided the audience a glimmer of how AI models shape research around campus.

Jason Ren, professor of civil and environmental engineering and the Andlinger Center for Energy and the Environment described how he headed a special issue of the journal Environmental Science & Technology, which focused on the use of data science to solve issues of sustainability. "We got an overwhelming response," said Ren. "It was in 70 years of ES&T history the largest special issue ever."

Environment and sustainability research covers a wide range of problems. Ren's collaborator Junjie Zhu, associate research scholar, said they're looking to large language models to help understand these issues – particularly those topics which have yet to accumulate much data. "Among these thousands of topics...is it possible to collect a variety of data together?" he asked, "And to develop a pretrained model that knows this area pretty well?"

From the Department of Psychology, Associate Chair Adele Goldberg wants to know what large language models can teach us about the way that humans comprehend language. "I think there are really deep and important parallels between human language and large language models," she said.

Mengdi Wang, associate professor of electrical and computer engineering and the Center for Statistics and Machine Learning, said that she and her collaborators recently "got really excited

about language models for scientific discovery.” She explained that AI models could be used to understand certain functional elements of RNA, the messenger molecule, which carries genetic information to our DNA.

Following the lightning talks, the symposium attendees gathered for a poster session featuring research using AI models from 24 graduate and post-doctoral researchers.

“One of the great things about Princeton is the amazing number of colleagues and ideas floating around campus,” said Arora.