

# Data Science and Analysis Semianr

FELIX YE

## BEYOND “CHATGPT, PLEASE PROVE THIS THEOREM”: HARNESSES FOR MATHEMATICAL RESEARCH

Tuesday, October 27, 2026  
3:00 P.M. in Massry B008

**ABSTRACT.** In July 2026, the key theorem in Shanmu Jin’s proof of Crouzeix’s conjecture came out of a sixteen-hour run of GPT-5.6 Sol. Jin did not just ask for a proof. He gave the model a prompt that organized it into a team: try different approaches, attack every candidate proof, and return to a dead end only with a new idea. Current models are strong on research-level benchmarks. But a proof takes hours, has no answer key, and goes through many failed attempts. The software that organizes this work around a model is called a harness. I explain how language models work, what a harness adds, and how my own systems carry out these ideas. The examples include a counterexample from a recent paper, and the talk ends with a workflow you can use on your own problems. No machine learning background is assumed.