LLMs as Tools

17-313 Spring 2024

Foundations of Software Engineering

https://cmu-313.github.io

Michael HIlton

Administrivia

P4 Checkpoint 1 moved to Thursday

LLMs as Tools

Developers are using LLMs as a part of their workflow to generate code

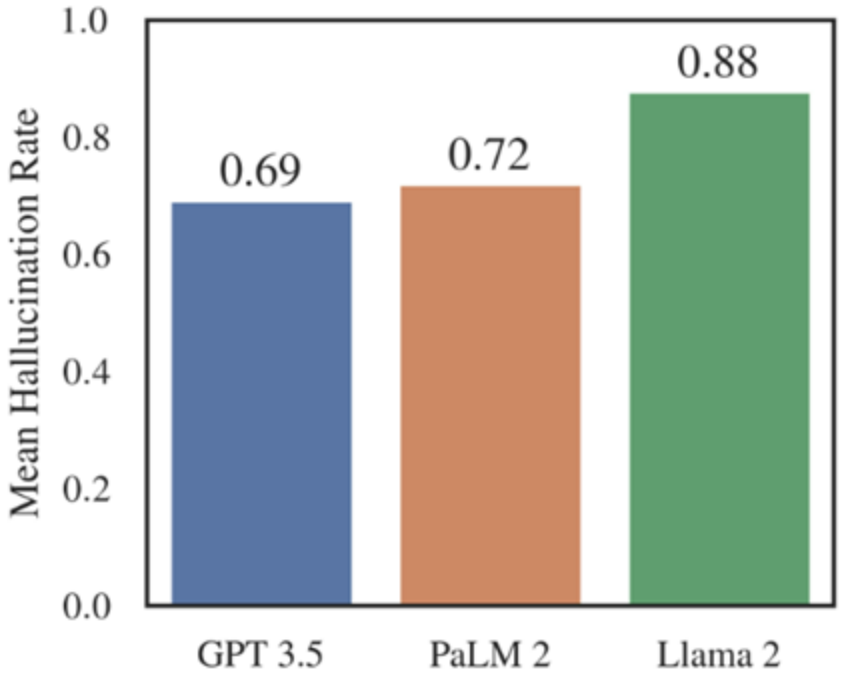
They can do a lot, but not everything

Is an LLM right for your problem?

The Stochastic parrot

LLMs still struggle with large context windows and detail oriented spaces even with it's many techniques to improve performance in those areas

- "Chat-GPT Lawyer": lawyer who submitted a legal report largely created by Chat-GPT
  - responses described as filled with "bogus judicial decisions , bogus quotes, and bogus internal citations."
- Example of a larger problem of hallucinations in detail oriented tasks



Legal hallucination rates across three popular LLMs.

Challenges with LLMS?

Today’s Activity

Fill out Survey

Demo

Activity

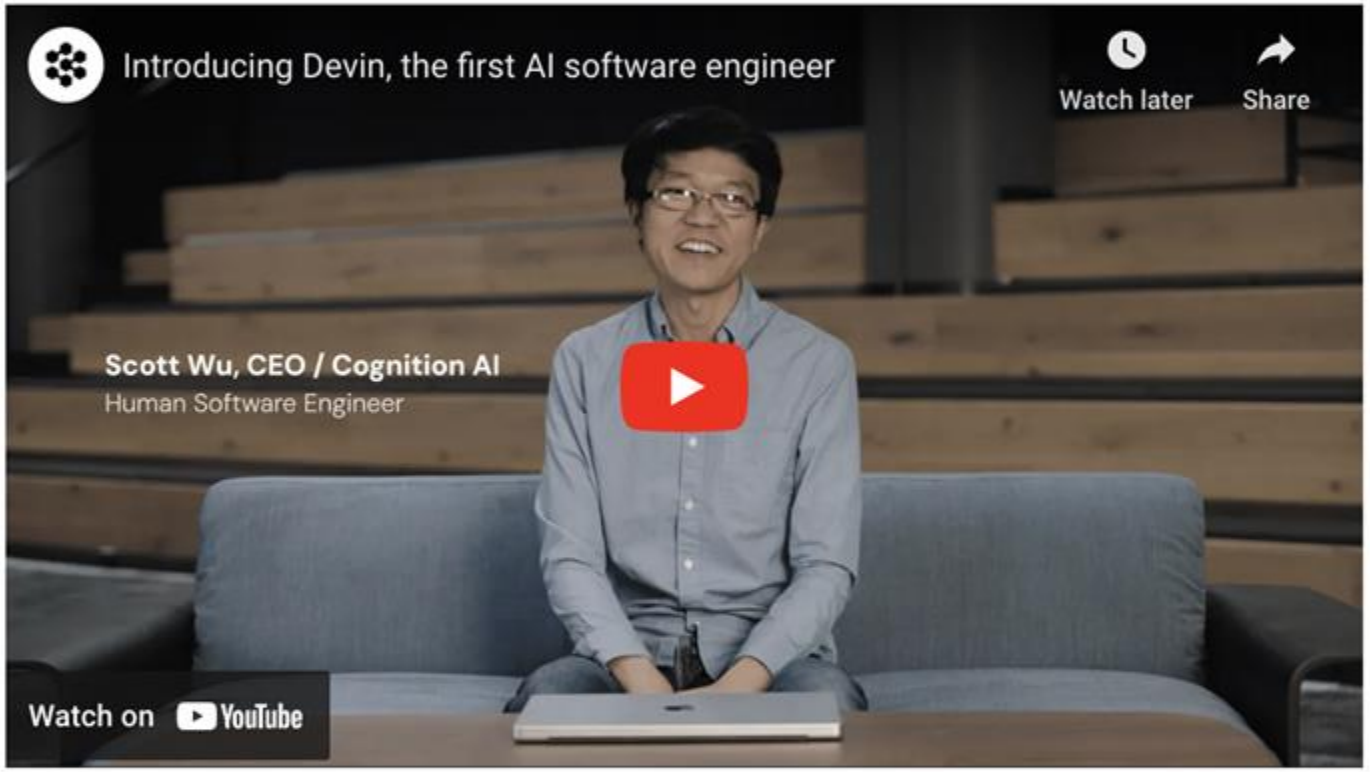
Final Survey

Discussion

We've raised a $21 million Series-A led by Founders Fund. Learn more here.

Cognition

About us Blog



March 12th, 2024 | Written by Scott Wu

Introducing Devin, the first AI software engineer

And setting a new state of the art on the SWE-bench coding benchmark

S3D
Carnegie Mellon University

Cognition

Open Positions (3)

Engineering

Machine Learning Researcher

Engineering • New York City or San Francisco Bay Area • Full time

Software Engineer

Engineering • New York City or San Francisco Bay Area • Full time

General

General Application

General • New York City or San Francisco Bay Area • Full time