

Teaching with Generative AI — Fall 2024

Polk Wagner

Faculty Retreat | September 19, 2024



Penn Carey Law
UNIVERSITY of PENNSYLVANIA

Agenda

The Generative AI Landscape (for Penn Law Faculty)

Some Ways to Use AI to Help Teaching — Now

Discussion & Feedback

Next Steps on Pedagogy Conversations



AI for Penn Law Faculty — Fall 2024

Basic Principles for the AI World

Use these AI models!

- drop in a paragraph and ask to re-write
- ask for information on a topic
- use them for 'weird stuff'
- Anything routine or straightforward
- iterate your inquiries

Use only the advanced or professional versions!

- usually \$20 / \$30 a month
- seems like soon we will have institutional accounts

Other Important Topics

There are **lots** of legal and ethical questions associated with LLMs:

- ethics of the way they were created
- ethics of the way they are constructed / used
- distributional concerns
- bias inherent in the models / bias in the training / bias in the guiderails
- copyright and right of publicity issues
- evidence and authentication of 'reality'
- amazingly efficient and effective tools for crime

These are important conversations and we should participate in them. But not today.

Two Major LLM Models of Relevance to Us

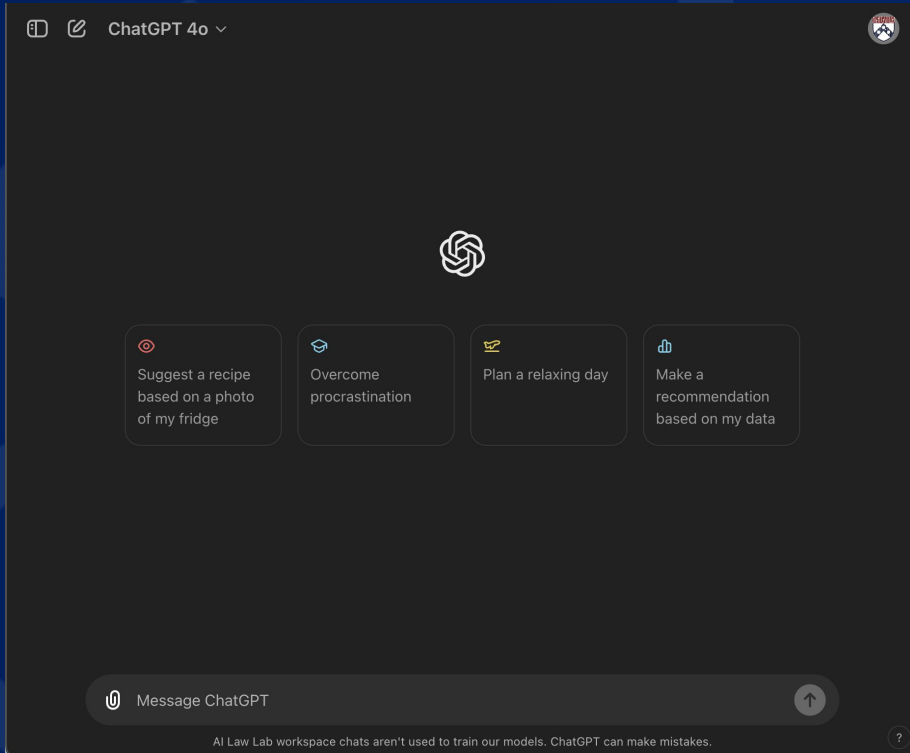
ChatGPT 4o / o1 / o3*

- the most capable across the most domains
- very multimodal (text, images, PDFs, spreadsheets, web access)
- offers the most options for customization and extensions
- good mobile app with voice

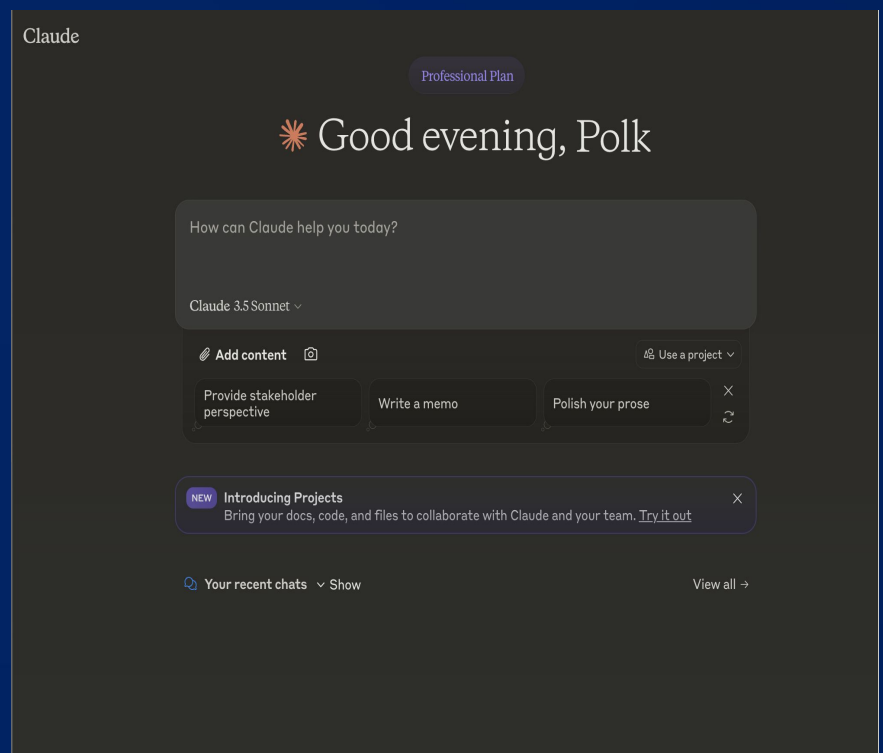
Claude 3.6/3.7 'Sonnet'

- virtually the same power as GPT 4o in most contexts
- typically a better writer at the first pass
- starting to develop innovative tools
- this model is built for safety, ethics
- excellent mobile app

Two Major Models of Relevance to Us



The screenshot shows the ChatGPT 4o interface. At the top left, there is a menu icon, a pencil icon, and the text "ChatGPT 4o" with a dropdown arrow. In the top right corner, there is a small profile icon. The main area features the OpenAI logo in the center. Below the logo are four rounded rectangular buttons with icons and text: "Suggest a recipe based on a photo of my fridge" (with a target icon), "Overcome procrastination" (with a graduation cap icon), "Plan a relaxing day" (with a calendar icon), and "Make a recommendation based on my data" (with a bar chart icon). At the bottom, there is a text input field with the placeholder "Message ChatGPT" and a send button with an upward arrow. A footer note at the bottom reads: "AI Law Lab workspace chats aren't used to train our models. ChatGPT can make mistakes." with a question mark icon to its right.



The screenshot shows the Claude interface. At the top left, the word "Claude" is displayed. In the top right corner, there is a "Professional Plan" badge. The main heading is "Good evening, Polk" with a starburst icon. Below this is a large text input field with the placeholder "How can Claude help you today?". Underneath the input field, it says "Claude 3.5 Sonnet" with a dropdown arrow. There is an "Add content" button with a plus icon and a "Use a project" dropdown menu. Below these are three buttons: "Provide stakeholder perspective", "Write a memo", and "Polish your prose", each with a refresh icon. A "NEW" badge is next to the "Introducing Projects" section, which includes the text "Bring your docs, code, and files to collaborate with Claude and your team. [Try it out](#)". At the bottom, there is a "Your recent chats" section with a "Show" button and a "View all" link with a right arrow.

Other Generative AI Tools of Note

Grammarly - Penn Faculty & Staff Access

- focused on writing
- seems good, but limited scope
- less utility for faculty

Lexis AI

- caselaw searching
- brief assistance
- utility is marginal so far

Westlaw / CoCounsel

- most advanced general-purpose legal-specific AI tool
- GPT 4 underneath the hood
- But ... the general purpose GPT now seems better, faster, with more tools

What you need to know now (1)

ChatGPT is probably the most flexible and strongest tool to use now.

Claude is also very good—better in some use cases—but a bit more limited.

These models are much better than they were in May 2024.

Chat GPT has launched a preview of a new model - “o1”

- the key difference is ‘reasoning’ - iterative analysis
- this is really important for some of what we do, like exams
 - in Spring 2024, GPT 4 could create B- (maybe B) answers on my exams
 - GPT o1 preview (this week): enormously better, definitely in the B to B+ range



Post-apocalyptic education

What comes after the Homework Apocalypse

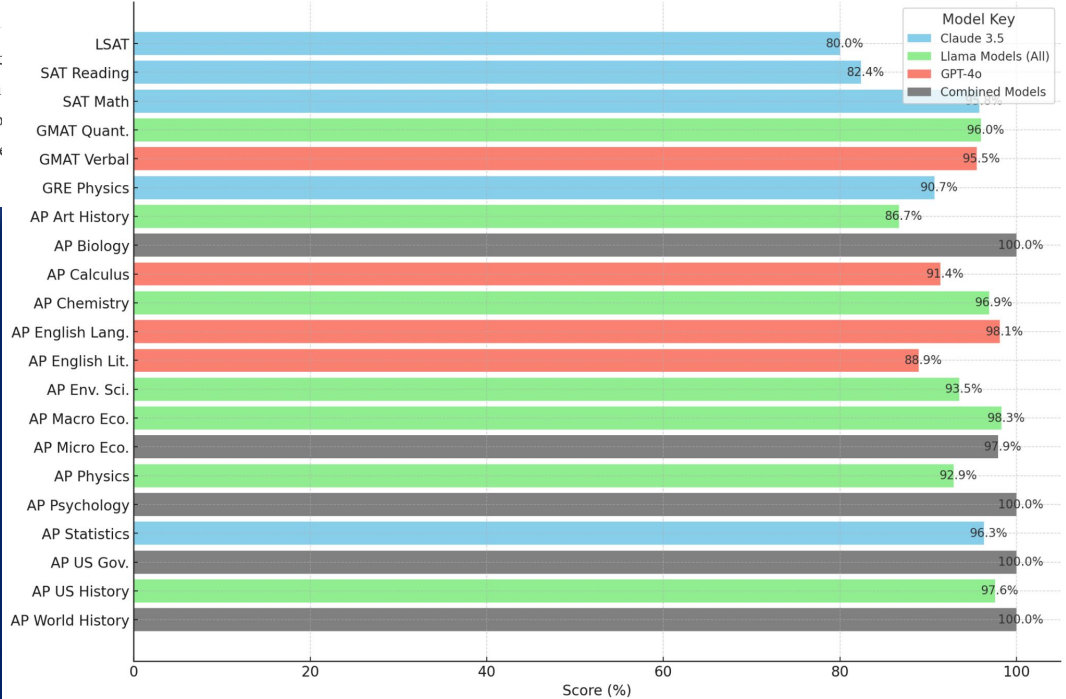


ETHAN MOLLICK
AUG 30, 2024

485 likes, 50 comments, 46 shares

Last summer, I wrote about [the Homework Apocalypse](#), the c could complete most traditional homework assignments, ren learning tools and assessment measures. My prophecy has co ace most tests. Yet remarkably little has changed as a result, c nearly universal among students.

AI Model Performance Across Various Standardized Tests

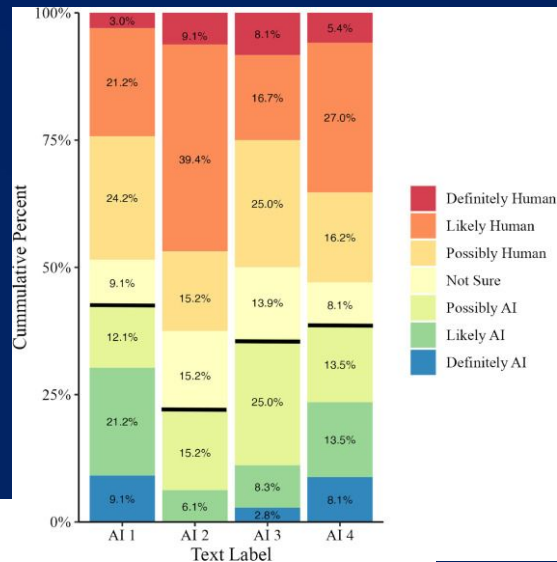


What You Need to Know Now (2)

There are no reliable tools to detect AI-generated writing. **None.**

Even expert humans cannot reliably detect AI writing.

The AI models cannot detect themselves. (95% wrong!)

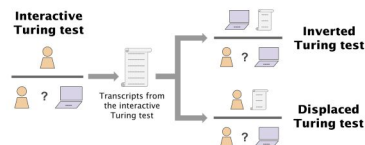


GPT-4 is judged more human than humans in displaced and inverted Turing tests

Ishika Rathi and Sydney Taylor and Benjamin K. Bergen and Cameron R. Jones
Department of Cognitive Science, UC San Diego
9500 Gilman Dr, San Diego, CA, USA
{irathi, cameron}@ucsd.edu

Abstract

Everyday AI detection requires differentiating between people and AI in informal, online conversations. In many cases, people will not interact directly with AI systems but instead read conversations between AI systems and other



Now What?

We do need to think hard about how to teach — and evaluate — students in different modes.

- that's the topic of our October "Pedagogy Conversation"

For exam-givers:

- seal off the exam conditions [?]
- ask questions that AIs will have trouble with
- flip the script and give AI-generated answers and ask for revisions and critique

For paper courses:

- assign outlines and rough drafts, verbal meetings and discussions
- note that LLMs are not that great at long-form text (yet)

Now What?

Bigger picture, these tools will become (are already?) fundamental to the creation of communication — and therefore legal practice.

They can radically reduce the “work” associated with a lot of the things we ask students to do.

- **But that work is the point!**
- **Our job is to find ways to create learning - regardless**



Using AI to Help You Teach - Today

Levels of Teaching with AI

Level 1

Use AI to create images for slides

Use AI to help script or outline a class

Use AI to create slides

Use AI to create hypotheticals

Level 2

Use AI to create practice problems

Use AI as a Virtual TA

Use AI to evaluate ('grade') essays

Levels of Teaching with AI

Level 3

Create AI agents to

- act as counterparty
- act as a judge
- act as a client

Level 4

....



AI Teaching Support - Demo

{ link }

Next Up

October (a Thursday lunch - TBA):

A conversation on rethinking assessments.

November (a panel):

How to have hard conversations in class.

December:

Holiday break

January:

Tips and Tricks / New Pedagogical Efforts

Meanwhile: Consider how you want to grow your teaching.

Talk with me about how I can help.