The Future of AI in Education and the Role of ChatGPT

by
Dr. Alfredo Capozucca

Wrap-up Event: Whether We Welcome It or Not, ChatGPT is in the classroom.
Thursday, 29 February 2024, Belval, Luxembourg.

Deep Blue (1997)

Creator: Stan Honda. Source: WIRED.

Creator: James the photographer. Source: Wikipedia.

AlphaGo (2016)

AlphaGo seals 4-1 victory over Go grandmaster Lee Sedol

DeepMind's artificial intelligence astonishes fans to defeat human opponent and offers evidence computer software has mastered a major challenge

📷 The world's top Go player, Lee Sedol, lost the final game of the Google DeepMind challenge match. Photograph: Yonhap/Reuters
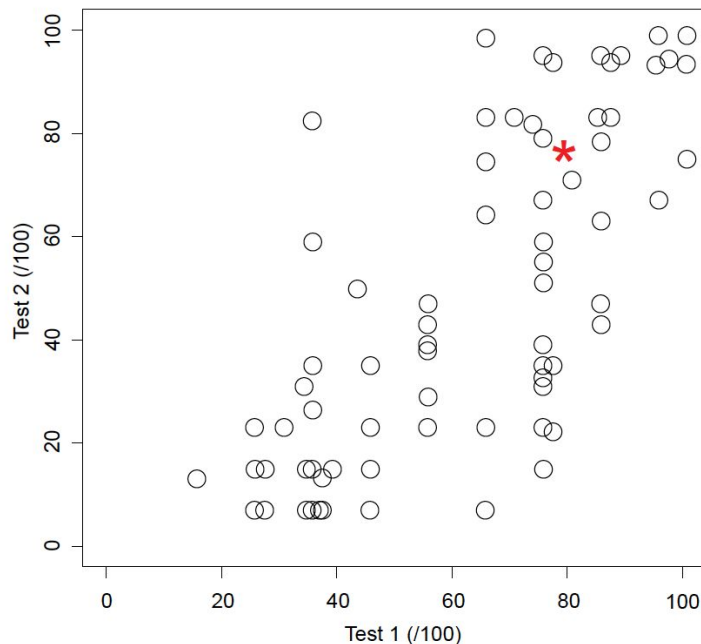
Source: The Guardian.

Codex (2021)

Scores of 71 students enrolled in the CS1 course in 2020 at the University of Auckland.

The performance of the responses generated by Codex is marked with a red asterisk.

Averaging both Test 1 and Test 2 performance, Codex's score is in position **17 when ranked alongside the 71 students' scores**: i.e. top quartile of class performance.

GPT-3.5 (2022)

Professional license exam in US is a as a precondition for law practice

Seven years of post-secondary education, including three years at an accredited law school.

Most test-takers also undergo **weeks to months** of further, exam-specific preparation.

One in five test-takers still score under the rate required to pass the exam on their first try.

GPT-3.5 performs at a passing rate for both Evidence and Torts.

GPT Takes the Bar Exam

Michael Bommarito II [1,2,3], Daniel Martin Katz[1,2,3,*]

1 Illinois Tech - Chicago Kent College of Law (Chicago, IL USA)
2 Bucerius Law School (Hamburg, Germany)
3 CodeX - The Stanford Center for Legal Informatics (Stanford, CA USA)

* Corresponding Author: dkatz3@kentlaw.iit.edu

Abstract

Nearly all jurisdictions in the United States require a professional license exam, commonly referred to as "the Bar Exam," as a precondition for law practice. To even sit for the exam, most jurisdictions require that an applicant completes at least seven years of post-secondary education, including three years at an accredited law school. In addition, most test-takers also undergo weeks to months of further, exam-specific preparation. Despite this significant investment of time and capital, approximately one in five test-takers still score under the rate required to pass the exam on their first try. In the face of a complex task that requires such depth of knowledge, what, then, should we expect of the state of the art in "AI?" In this research, we document our experimental evaluation of the performance of OpenAI's TEXT-DAVINCI-003 model, often-referred to as GPT-3.5, on the multistate multiple choice (MBE) section of the exam. While we find no benefit in fine-tuning over GPT-3.5's zero-shot performance at the scale of our training data, we do find that hyperparameter optimization and prompt engineering positively impacted GPT-3.5's zero-shot performance. For best prompt and parameters, GPT-3.5 achieves a headline correct rate of 50.3% on a complete NCBE MBE practice exam, significantly in excess of the 25% baseline guessing rate, and performs at a passing rate for both Evidence and Torts. GPT-3.5's ranking of responses is also highly-correlated with correctness; its top two and top three choices are correct 71% and 88% of the time, respectively, indicating very strong non-entailment performance. While our ability to interpret these results is limited by nascent scientific understanding of LLMs and the proprietary nature of GPT, we believe that these results strongly suggest that an LLM will pass the MBE component of the Bar Exam in the near future.

Source: Dec 2022, https://arxiv.org/abs/2212.14402

ChatGPT (November 30, 2022)

Photograph: Pavlo Gonchar/Sopa Images/Rex/Shutterstock

- 100 million users just two months after launching.

- TikTok took about nine months.

- Instagram more than two years.

- Facebook, took around four and a half years.

- Twitter took over five years,

Source: Feb 3, 2023, The Guardian and Nov 6, 2023, The Verge.

Reactions

New York City schools ban AI chatbot that writes essays and answers prompts

ChatGPT tool will be forbidden across all devices and networks in public schools over 'concerns about negative impacts on learning'
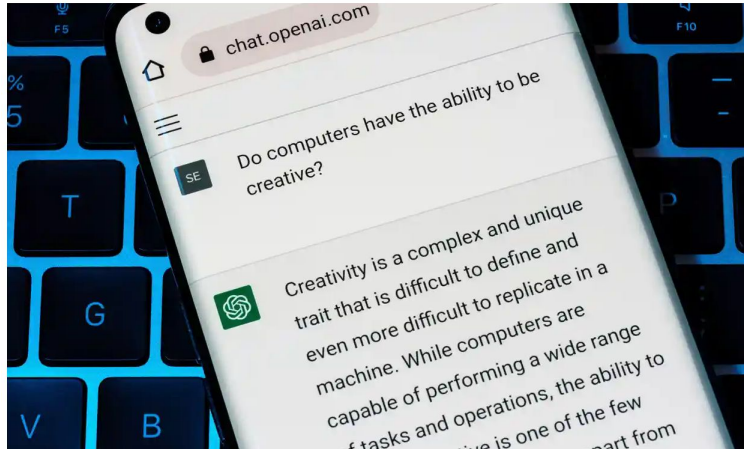
Source: The Guardian, 6 **Jan** 2023. Link

Alarmed by AI chatbots, universities start revamping how they teach

At schools including George Washington University in Washington, D.C., Rutgers University in New Brunswick, New Jersey, and Appalachian State University in Boone, North Carolina, professors are phasing out take-home, open-book assignments — which became a dominant method of assessment in the pandemic but now seem vulnerable to chatbots. They are instead opting for in-class assignments, handwritten papers, group work and oral exams.

Source: The Japanese Times, 18 Jan 2023. Link

Reactions (2)

Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

Signatures
33708

Add your signature

Signatories

Yoshua Bengio, Founder and Scientific Director at Mila, Turing Prize winner and professor at University of Montreal

Stuart Russell, Berkeley, Professor of Computer Science, director of the Center for Intelligent Systems, and co-author of the standard textbook "Artificial Intelligence: a Modern Approach"

Elon Musk, CEO of SpaceX, Tesla & Twitter

Steve Wozniak, Co-founder, Apple

Yuval Noah Harari, Author and Professor, Hebrew University of Jerusalem.

Source: Future of Life Institute, 22 **March** 2023. Link

ChatGPT banned in Italy over privacy concerns

🕑 1 April 2023

GETTY IMAGES

OpenAI launched ChatGPT last November

Source: BBC, 1 **April** 2023. Link

Reaction's results

- In less than one month [ChatGPT was accessible again in Italy](#).

- [New York City Public Schools Reverses ChatGPT Ban](#) in May 2023.

- Pause on AI? Instead [Development Sped Up](#)

  - Creation of [xAI company](#)

  - ChatGPT-4 and Turbo (OpenAI)

  - Gemini (Google)

  - LlaMA (Meta)

  - Sora (OpenAI)

  - ...

AI assistants are here to stay.

AI assistants in education

- 1,200 UK undergraduates.

- 53% have used generative AI to help them prepare assessments

- Just 3% admitted to copying and pasting unedited AI-generated text into their assessments.

- 35% do not know how often it produces made-up facts

More than half of UK undergraduates say they use AI to help with essays

Teachers also using the generative technology to aid with lesson planning, with hopes it could ease the burden of their workload
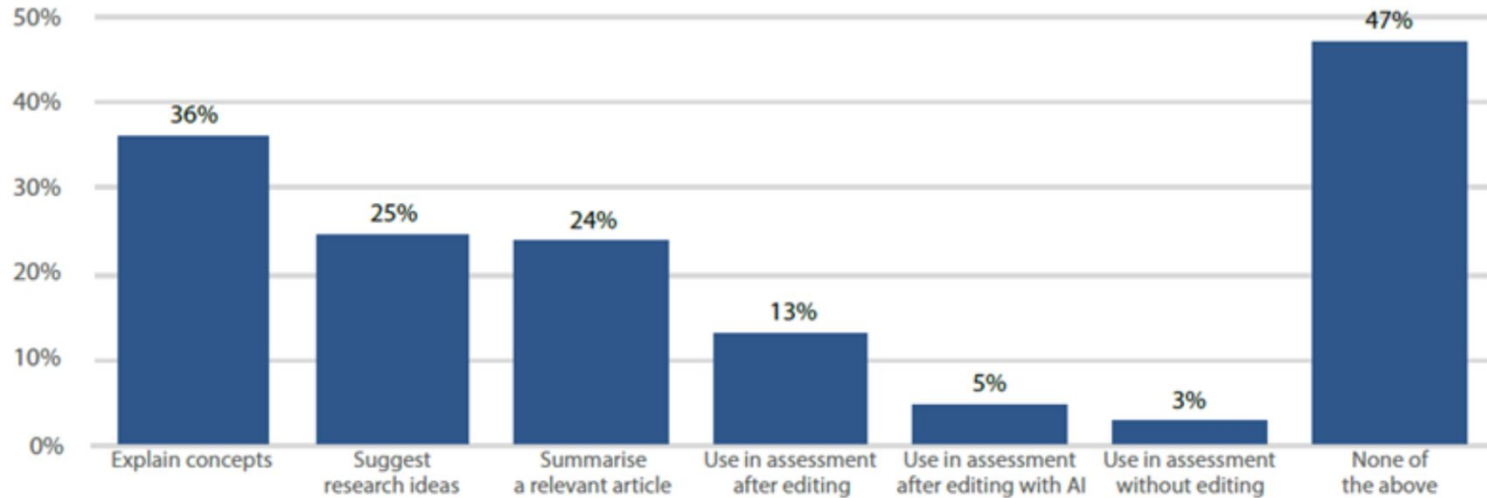

📷 One in 20 students admitted they had copied and pasted unedited AI-generated texts into assessments. Photograph: John Walton/PA

Source: The Guardian, 1 Feb 2024. Link

Usage (by students)

Figure 4: 'When thinking about using generative AI to prepare assessed work, which of the following have you ever done? Tick all that apply.' All responses.

Are we ready?

- Over 924 K-12 educators in **USA**, more than 79 % say their districts still do not have clear policies on the use of artificial intelligence tools ([EdWeek Research Center, Nov-Dec 2023](#)).

- Over 450 schools and universities over the **world**, fewer than 10% have developed institutional policies and/or formal guidance concerning the use of generative AI applications ([UNESCO, May 2023](#))

Policy (example)

In this course **you are allowed to use AI** (e.g. ChatGPT). However, you must be aware that:

- If you provide minimum effort prompts, you will get low quality results. You will need to refine your prompts in order to get good outcomes. This will take work.

- Don't trust anything it says. If it gives you a number or fact, assume it is wrong unless you either know the answer or can check in with another source. You will be responsible for any errors or omissions provided by the tool. It works best for topics you understand.

- AI is a tool, but one that you need to acknowledge using. Please include a paragraph at the end of any assignment that uses AI explaining what you used the AI for and what prompts you used to get the results. Failure to do so constitutes fraud and will be penalised accordingly.

- Be thoughtful about when this tool is useful. Don't use it if it isn't appropriate for the case or circumstance.

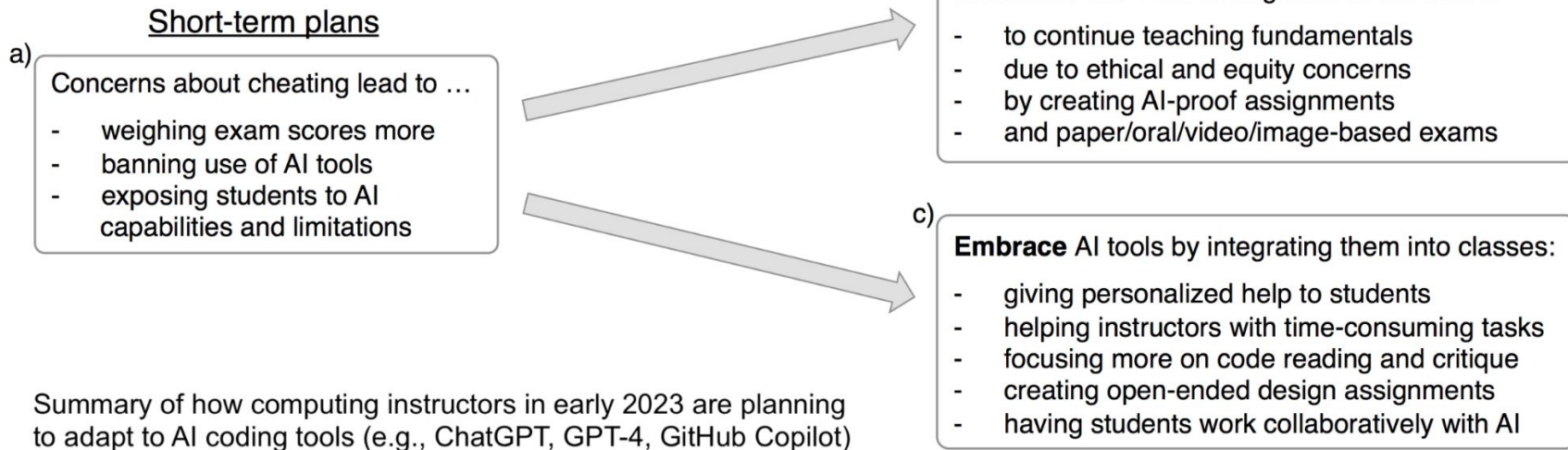This policy is based on Ethan Mollick's.

Source: https://oneusefulthing.substack.com/p/my-class-required-ai-heres-what-ive

Trial

- Topic: Formal methods

- Learning outcome: write formal specifications using the B formal language.

- Research question: does ChatGPT help students to write formal specifications?

- Study design: pre-test post-test single group design

- (Early) results:

    - + boilerplate structure of the specification.

    - - wrong syntax

    - - wrong predicates (invariants, pre/post conditions)

HINT: (for now) it is not a problem to allow students to use ChatGPT during the final exam.

Alternatives

Short-term plans

a)

Concerns about cheating lead to …

- weighing exam scores more
- banning use of AI tools
- exposing students to AI capabilities and limitations

Longer-term plans

b)

Resist the use of AI coding tools in the future:

- to continue teaching fundamentals
- due to ethical and equity concerns
- by creating AI-proof assignments
- and paper/oral/video/image-based exams

c)

Embrace AI tools by integrating them into classes:

- giving personalized help to students
- helping instructors with time-consuming tasks
- focusing more on code reading and critique
- creating open-ended design assignments
- having students work collaboratively with AI

Summary of how computing instructors in early 2023 are planning to adapt to AI coding tools (e.g., ChatGPT, GPT-4, GitHub Copilot)

Source: Sam Lau and Philip J. Guo. From "Ban It Till We Understand It" to "Resistance is Futile": How University Programming Instructors Plan to Adapt as More Students Use AI Code Generation and Explanation Tools such as ChatGPT and GitHub Copilot. ACM Conference on International Computing Education Research (ICER), 2023.

Evidence to back up decisions.

Prediction

…"all programs in the future will ultimately be written by AIs, with humans relegated to, at best, a supervisory role."

…"CS students are not going to need to learn such mundane skills as how to add a node to a binary tree or code in C++. That kind of education will be antiquated, …"

The End of Programming

By Matt Welsh
Communications of the ACM, January 2023, Vol. 66 No. 1, Pages 34-35
10.1145/3570220
Comments (11)

VIEW AS: ☐ ☐ ☐ ☐ ☐ SHARE: ☐ ☐ ☐ ☐ ☐ ☐

I came of age in the 1980s, programming personal computers such as the Commodore VIC-20 and Apple][e at home. Going on to study computer science (CS) in college and ultimately getting a Ph.D. at Berkeley, the bulk of my professional training was rooted in what I will call "classical" CS: programming, algorithms, data structures, systems, programming languages. In Classical Computer Science, the ultimate goal is to reduce an idea to a program written by a human—source code in a language like Java or C++ or Python. Every idea in Classical CS—no matter how complex or sophisticated, from a database join algorithm to the mind-bogglingly obtuse Paxos consensus protocol—can be expressed as a human-readable, human-comprehendible program.
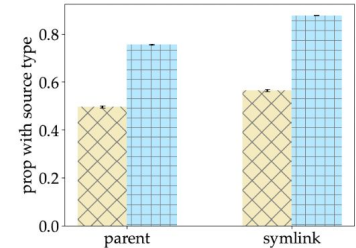
Evidence

- OpenAI's codex-davinci-002

- 47 participants
- experiment group = 33
- control group = 14

- Participants who had access to an AI assistant were far more likely to write incorrect and insecure solutions compared to the control group.
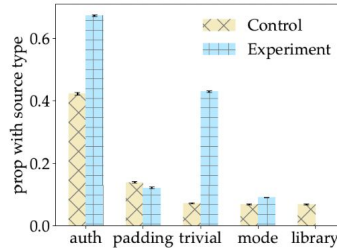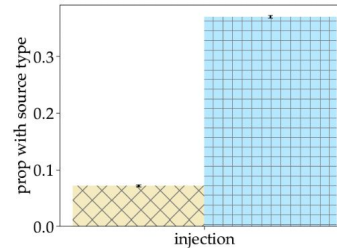

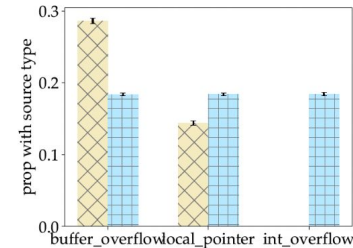(a) Q1 Mistakes: Encryption & Decryption
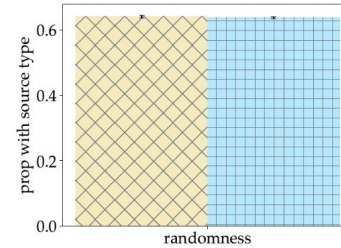
(b) Q2 Mistakes: Signing a Message

(c) Q3 Mistakes: Sandboxed Directory

(d) Q4 Mistakes: SQL

(e) Q5 Mistakes: C Strings

Challenge

- Usage (for learning purposes)

  - Is it worth to use an AI assistant?

  - If yes, then what is the baseline knowledge to use it?

Challenge (2)

- (Learners) Know how

  - What are the subjects to keep/remove from the current curricula?

  - What are the critical skills to develop?

Challenge (3)

- Ethics

    ○ What are the principles that govern the conduct of professionals regarding:

        ■ **Use** of AI assistants?

        ■ **Development** of AI assistants?

Generated with AI · 29 February 2024 at 10:39 am - DALL-E 3

That's all.
Thank you for your attention.