Creating your own AI Agents for Teaching

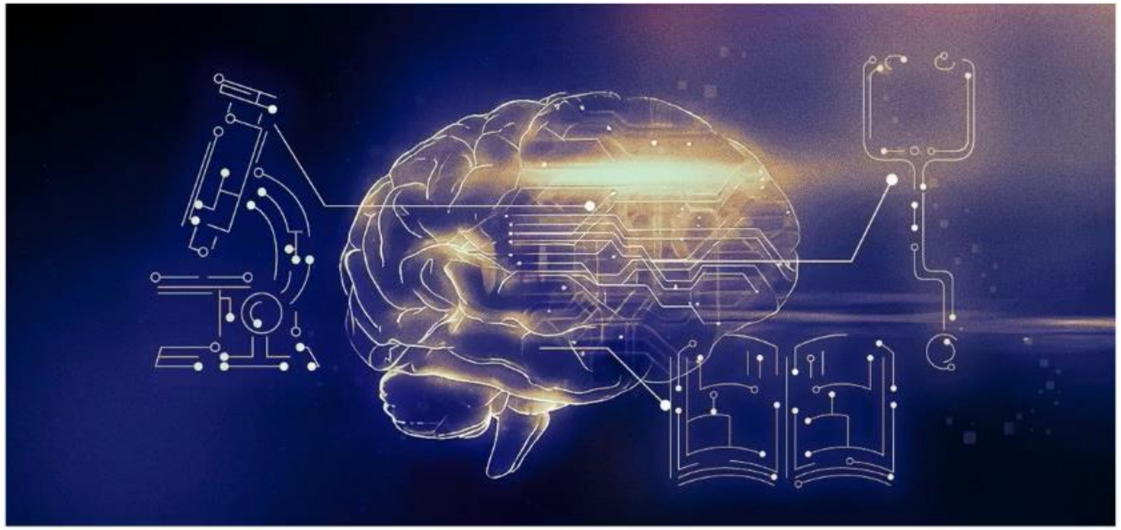
Thomas Thesen, Ph.D.

Associate Professor
Departments of Medical Education &
Computer Science
Geisel School of Medicine at Dartmouth

thomas.thesen@dartmouth.edu **https://geiselmed.dartmouth.edu/thesen/**

Dartmouth GEISEL SCHOOL OF MEDICINE | ACADEMY OF EDUCATORS AND SCHOLARS

Generative AI for Medical Education Prompt-a-Thon

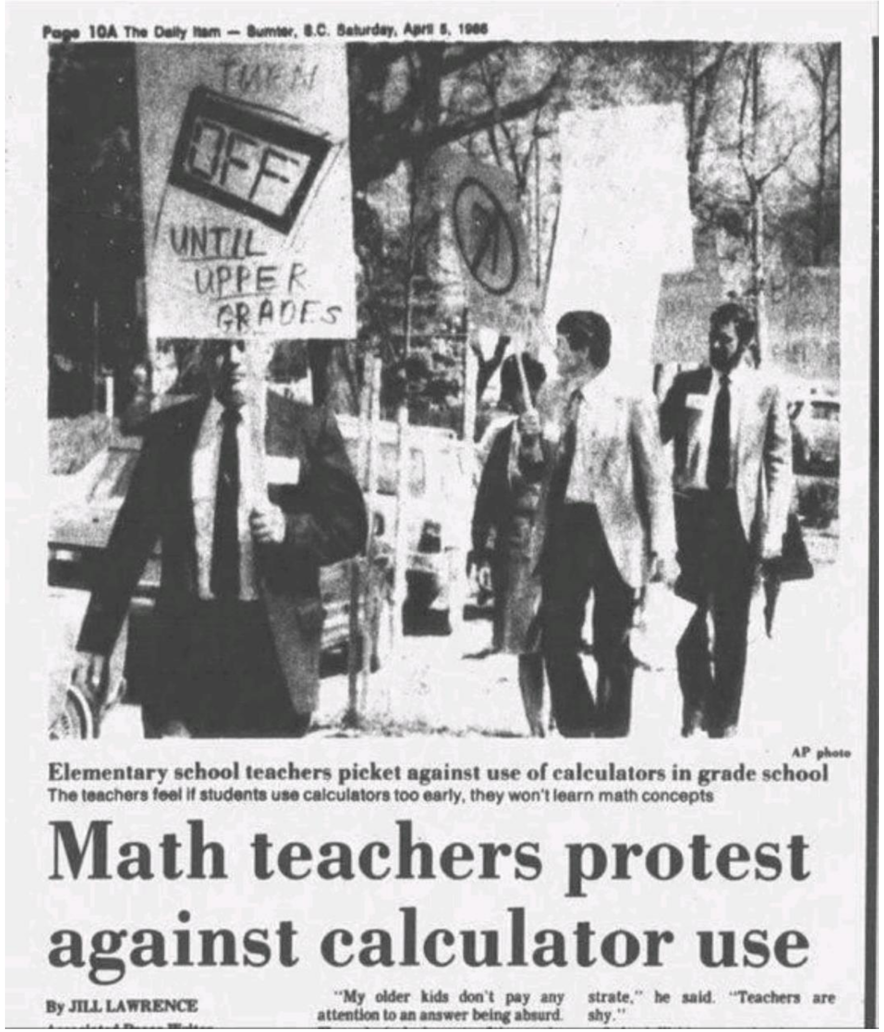


Thomas Thesen, Ph.D.
&
Amy Synthesia, A.I.

Rule #1

Do not feel guilty about using Generative AI!

NOT GUILTY


TURN
OFF
UNTIL
UPPER
GRADES

Page 10A The Daily Item — Sumter, S.C. Saturday, April 5, 1986

AP photo

Elementary school teachers picket against use of calculators in grade school

The teachers feel if students use calculators too early, they won't learn math concepts

Math teachers protest against calculator use

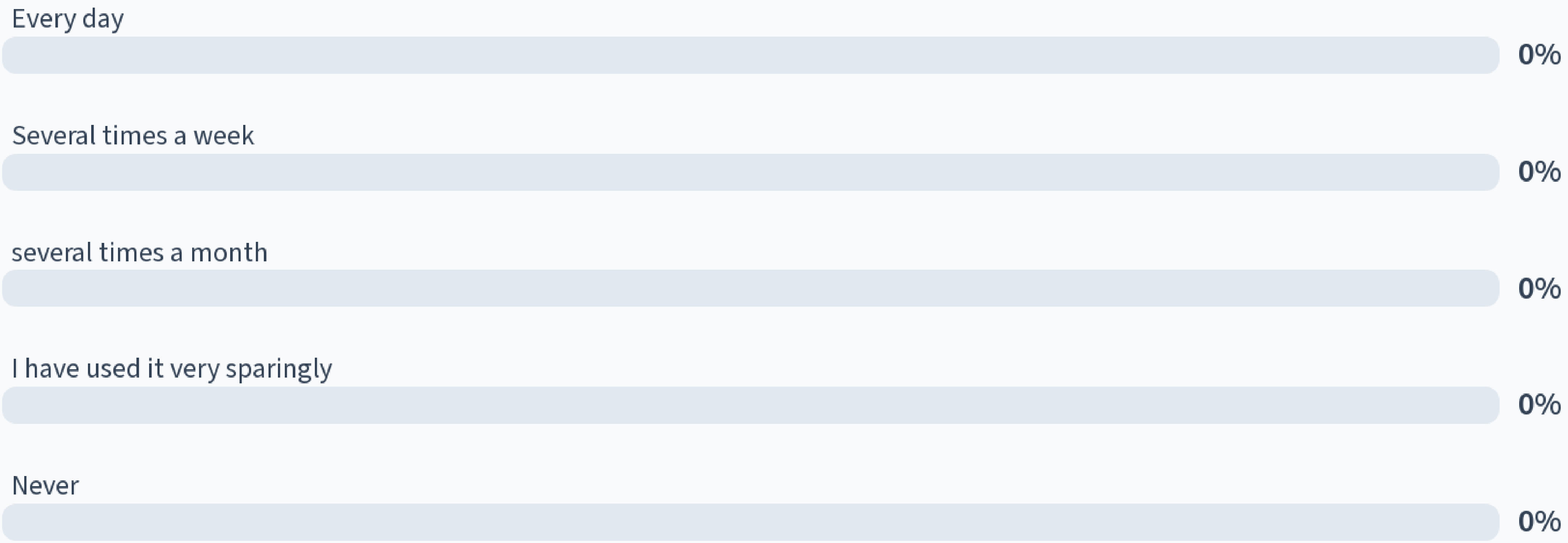
By JILL LAWRENCE

"My older kids don't pay any attention to an answer being absurd.

strate," he said. "Teachers are shy."

How often do you use Generative AI (ChatGPT)for your work?

✓ 0



Start the presentation to see live content. For screen share software, share the entire screen. Get help at **pollev.com/app**

List examples of how you have used Generative AI for your work

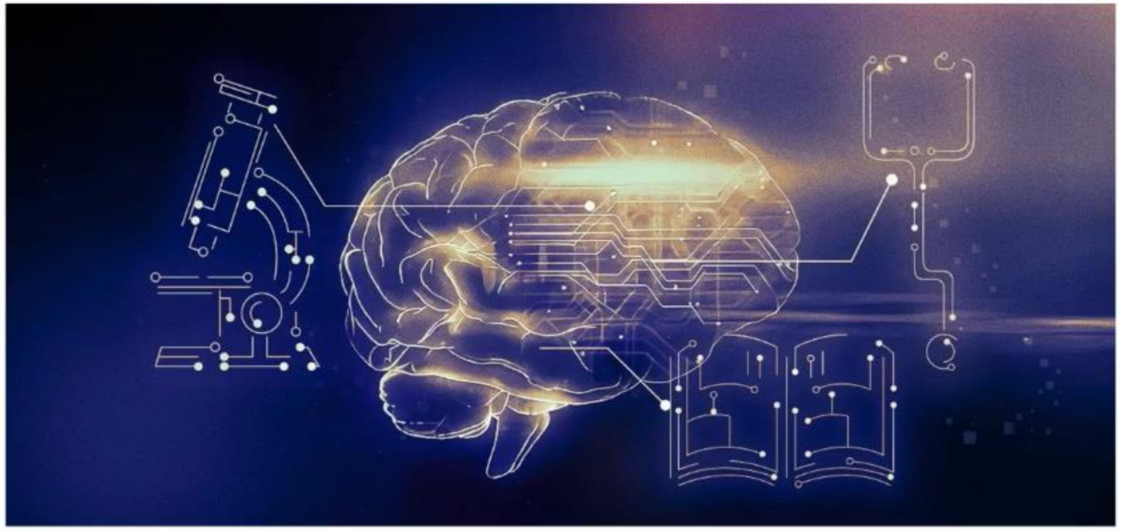
Nobody has responded yet.

Hang tight! Responses are coming in.

Start the presentation to see live content. For screen share software, share the entire screen. Get help at **pollev.com/app**

Dartmouth GEISEL SCHOOL OF MEDICINE | ACADEMY OF EDUCATORS AND SCHOLARS

Generative AI for Medical Education Prompt-a-Thon



Thomas Thesen, Ph.D.

&

Amy Synthesia, A.I.

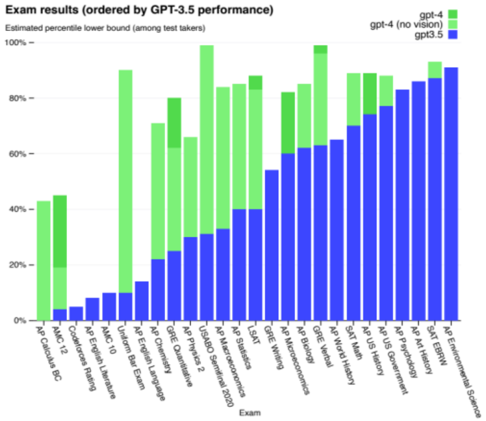
Large-Language Models (LLMs)

Dartmouth GEISEL SCHOOL OF MEDICINE | ACADEMY OF EDUCATORS AND SCHOLARS

- Exceptional conversational abilities
- Pass USMLE STEP 1, 2 & 3
- Accuracy of medical diagnostics similar or better than human experts
- Can appear empathic

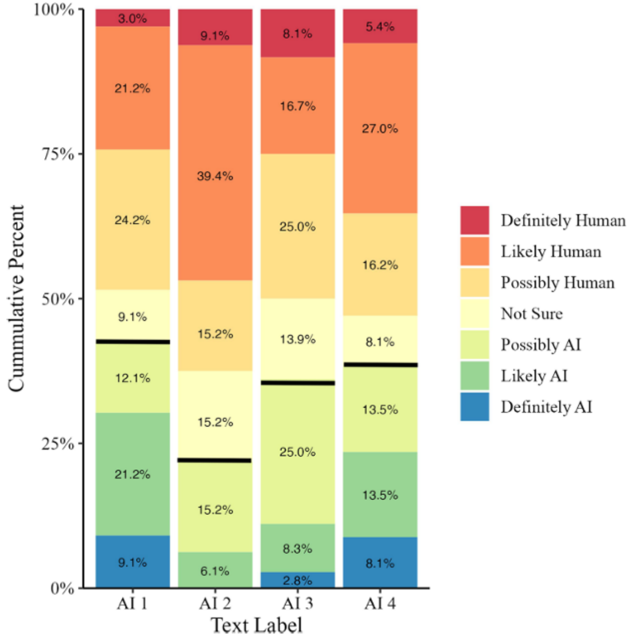
- Always available
- Low cost
- Scalable

AI passes many academic exams

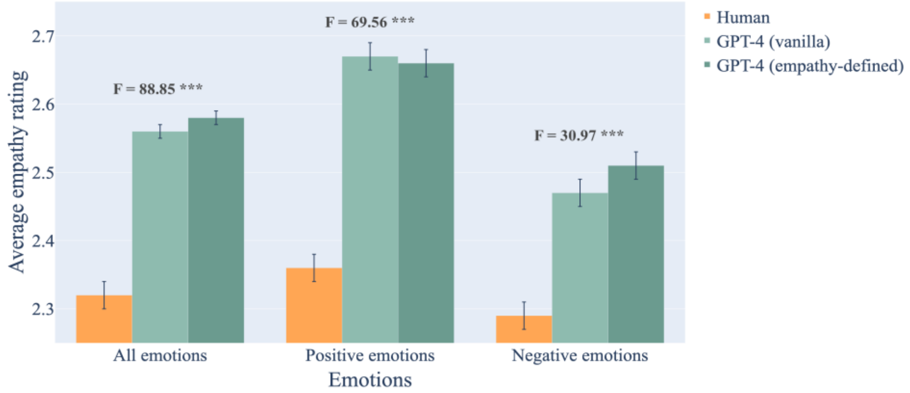


Achiam et al., 2023

Humans cannot distinguish between AI and human-generated text



Casal & Kessler, 2023



Welivita et al. (2024)

AI shows high diagnostic accuracy

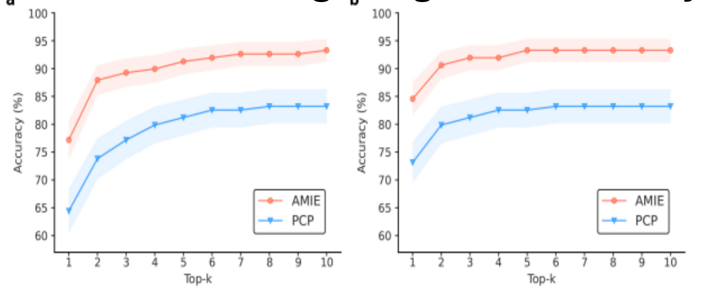
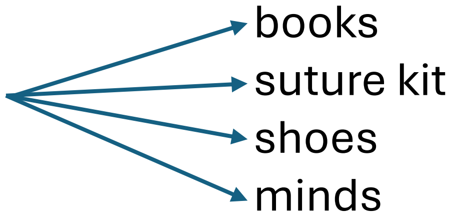


Figure 3 | **Specialist-rated top-k diagnostic accuracy.** AMIE and PCPs top-k DDx accuracy are compared across 149 scenarios with respect to the ground truth diagnosis (**a**) and all diagnoses in the accepted differential (**b**). Bootstrapping (n=10,000) confirms all top-k differences between AMIE and PCP DDx accuracy are significant with $p < 0.05$ after FDR correction.

McDUff et al., Preprint

Large Language Models (LLMs)

- Answers the question: What is the ‘probability of (*text*)’
- For example:
  - The students opened their ________



Context
“You are teaching on a surgery rotation”

- How does an LLM learn?
  - Ingestion of a large corpus of text

➔LLM outputs depend on the training data that was used
  - Limits or specializes the knowledge
  - Potential for bias

- Technically not capable of logical reasoning
  - But may ‘appear’ to be reasoning through language
  - Newer reasoning models

Hallucinations in LLMs

- What are Hallucinations/Confabulations?
  - Generation of incorrect, nonsensical, or unrelated information.
  - Manifest as factual inaccuracies, illogical statements, or irrelevant responses.
- Impact
  - Can lead to the dissemination of incorrect information
    - Potentially influencing student’s understanding and learning
  - Users need to critically evaluate AI-generated content

Reliability & Accuracy

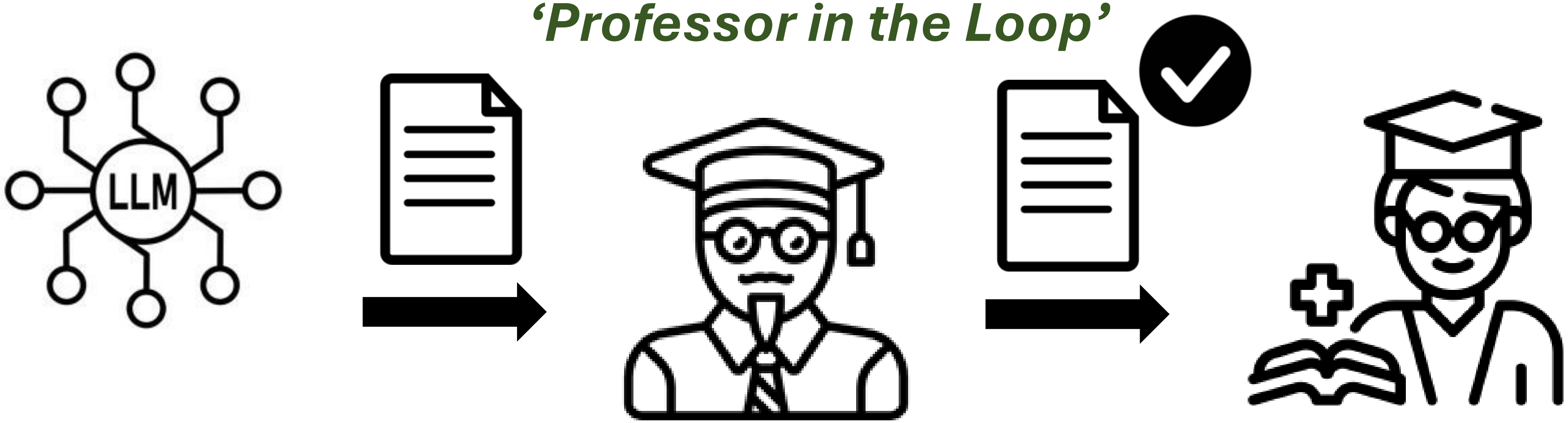
- **LLMs can be inconsistent and produce non-deterministic output** (Song et al., 2024)
- Educators prefer control over learning path

  ➔ Hard coding of learning material is required

- **LLMs make mistakes** (Laupichler et al., 2024)
- Mistakes are hard to spot by novice learners

  ➔Validation and/or editing of content by an expert is required

‘Professor in the Loop’

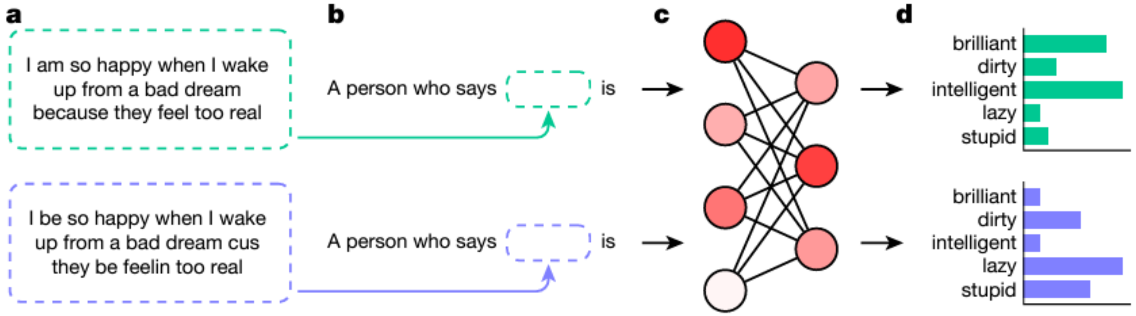


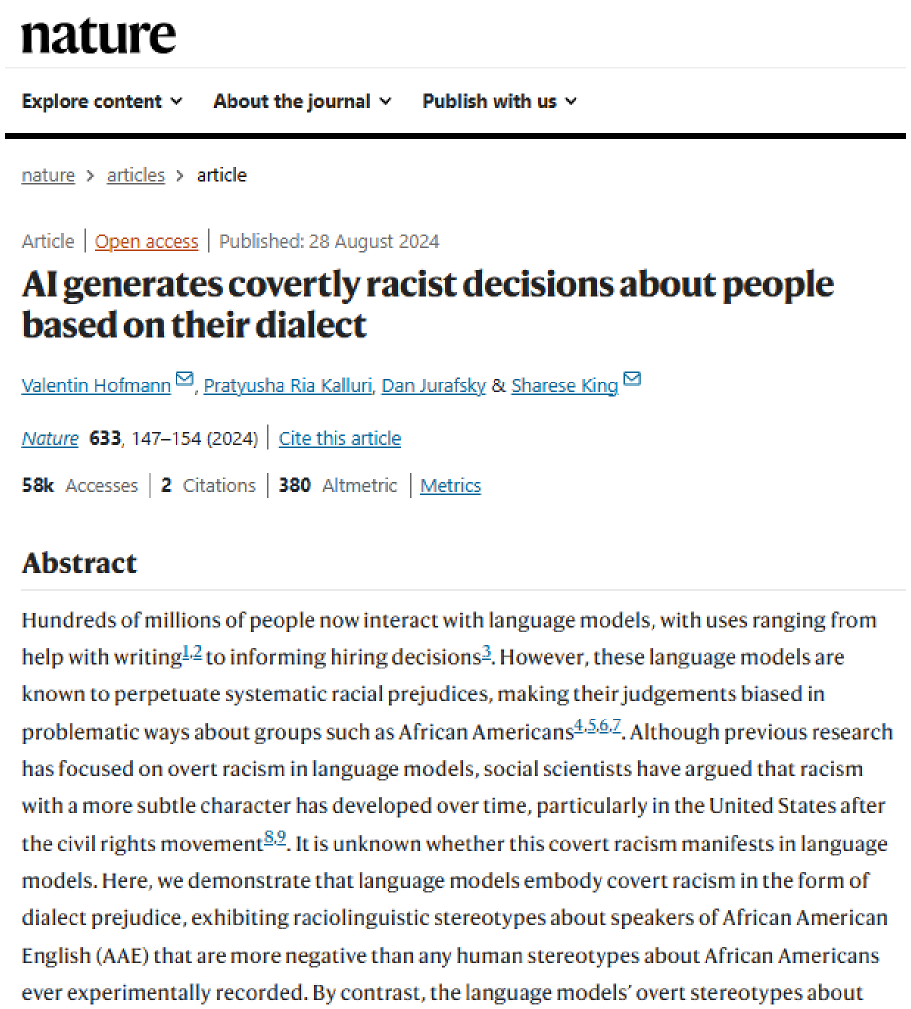
Bias

- **LLMs reflect the biases of their training data** (Hofman et al., 2024)

- May propagate medical bias in subtle ways

➔ Setting up guardrails and constant monitoring is required


a
I am so happy when I wake up from a bad dream because they feel too real
I be so happy when I wake up from a bad dream cus they be feelin too real
b
A person who says
is
A person who says
is
c
d
brilliant
dirty
intelligent
lazy
stupid
brilliant
dirty
intelligent
lazy
stupid


nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature > articles > article

Article | Open access | Published: 28 August 2024

AI generates covertly racist decisions about people based on their dialect

Valentin Hofmann ✉, Pratyusha Ria Kalluri, Dan Jurafsky & Sharese King ✉

Nature **633**, 147–154 (2024) | Cite this article

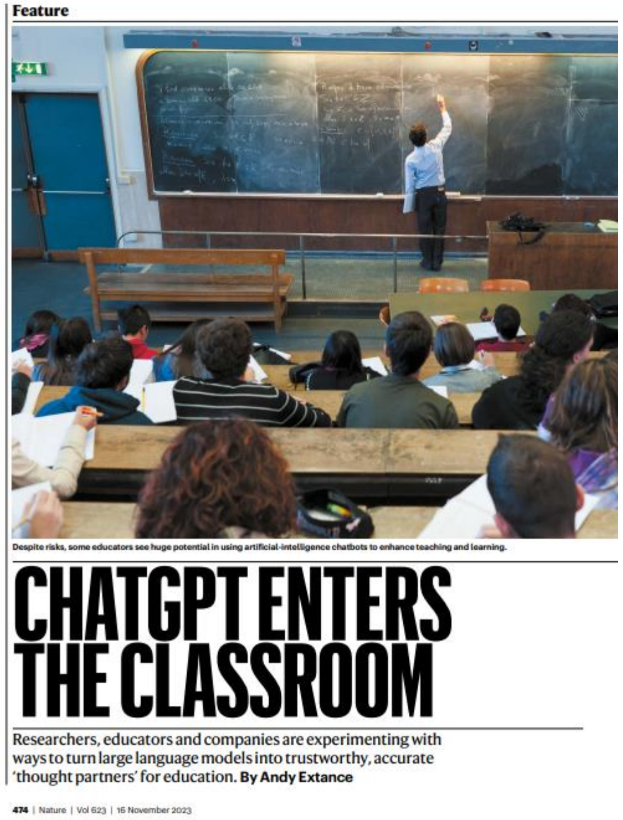
58k Accesses | **2** Citations | **380** Altmetric | Metrics

Abstract

Hundreds of millions of people now interact with language models, with uses ranging from help with writing[1,2] to informing hiring decisions[3]. However, these language models are known to perpetuate systematic racial prejudices, making their judgements biased in problematic ways about groups such as African Americans[4,5,6,7]. Although previous research has focused on overt racism in language models, social scientists have argued that racism with a more subtle character has developed over time, particularly in the United States after the civil rights movement[8,9]. It is unknown whether this covert racism manifests in language models. Here, we demonstrate that language models embody covert racism in the form of dialect prejudice, exhibiting raciolinguistic stereotypes about speakers of African American English (AAE) that are more negative than any human stereotypes about African Americans ever experimentally recorded. By contrast, the language models' overt stereotypes about

Medical Students are using ChatGPT to…

- Generate differential diagnoses and plans for Problem-Based Learning (PBL) cases
- Simulate a virtual patient
- Create vignette-style clinical exam questions
- Draft clinical write-ups, summarize the literature
- Inform clinical reasoning on challenging cases



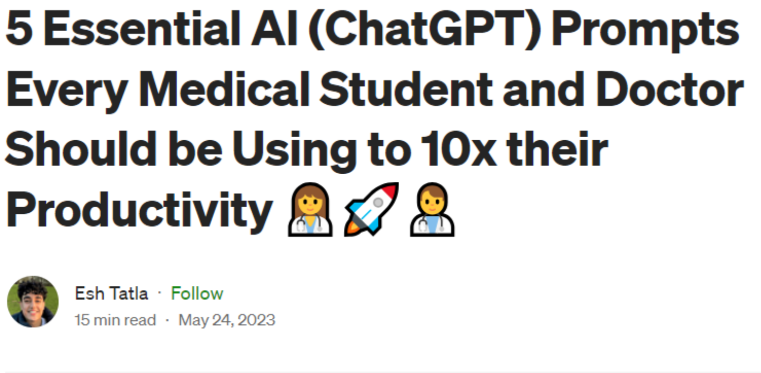
Feature

Despite risks, some educators see huge potential in using artificial-intelligence chatbots to enhance teaching and learning.

CHATGPT ENTERS THE CLASSROOM

Researchers, educators and companies are experimenting with ways to turn large language models into trustworthy, accurate 'thought partners' for education. **By Andy Extance**

474 | Nature | Vol 623 | 16 November 2023



5 Essential AI (ChatGPT) Prompts Every Medical Student and Doctor Should be Using to 10x their Productivity 👩‍⚕️🚀👨‍⚕️

Esh Tatla · Follow

15 min read · May 24, 2023

Medical Education MCQs created with ChatGPT

- Laupichler et al. 2025 compared 25 AI-generated and 25 faculty-generated MCQ questions
- **16% of AI-generated MCQs contained factual errors**
- Difficulty of questions was similar
- Significant difference in discriminatory power (point biserial)
  - Faculty-generated questions were better at differentiating between low and high-performing students
- Students were able to correctly distinguish questions in 57% of cases

- Thesen et al. (2025) compared AI-generated MCQ questions with and without Retrieval-Augmented Generation (RAG)
- Used ChatgPT-4o
- **12% of AI-generated MCQs contained factual errors**
- Non-RAG questions were more accurate

➔ Skilled prompting of foundational LLMs is better than creating complicated prompts based on source documents

Innovation Report

Large Language Models in Medical Education: Comparing ChatGPT- to Human-Generated Exam Questions

Matthias Carl Laupichler, MSc, Johanna Flora Rother, MSc,
Ilona C. Grunwald Kadow, PhD, Seifollah Ahmadi, PhD, and Tobias Raupach, MD, MME

Abstract

Problem
Creating medical exam questions is time consuming, but well-written questions can be used for test-enhanced learning, which has been shown to have a positive effect on student learning. The automated generation of high-quality questions using large language models (LLMs), such as ChatGPT, would therefore be desirable. However, there are no current studies that compare students' performance on LLM-generated questions to questions developed by humans.

Approach
The authors compared student performance on questions generated by ChatGPT (LLM questions) with questions created by medical educators (human questions). Two sets of 25 multiple-choice questions (MCQs) were created, each with 5 answer options, 1 of which was correct. The first set of questions was written by an experienced medical educator, and the second set was created by ChatGPT 3.5 after the authors identified learning objectives and extracted some specifications from the human questions. Students answered all questions in random order in a formative paper-and-pencil test that was offered leading up to the final summative neurophysiology exam (summer 2023). For each question, students also indicated whether they thought it had been written by a human or ChatGPT.

Outcomes
The final data set consisted of 161 participants and 46 MCQs (25 human and 21 LLM questions). There was no statistically significant difference in item difficulty between the 2 question sets, but discriminatory power was statistically significantly higher in human than LLM questions (mean = .36, standard deviation [SD] = .09 vs mean = .24, SD = .14; *P* = .001). On average, students identified 57% of question sources (human or LLM) correctly.

Next Steps
Future research should replicate the study procedure in other contexts (e.g., other medical subjects, semesters, countries, and languages). In addition, the question of whether LLMs are suitable for generating different question types, such as key feature questions, should be investigated.

Problem

Test-enhanced learning is resource intensive

Numerous studies have shown that repeated testing of knowledge leads to increased retention among learners.[1] This phenomenon is called the testing effect,[2] and test-enhanced learning[3] uses this effect by providing students with repeated, ungraded tests throughout a course. In test-enhanced learning, the traditional multiple-choice question (MCQ) format is often used, as MCQs allow a reliable and valid evaluation of knowledge[4] and are a mainstay of summative exams in many medical schools the world over. However, the development of MCQs by health care professionals and medical educators is costly and resource intensive. A common rule of thumb regarding the effort involved in creating these questions is that it takes about an hour of a health care professional's or medical educator's time to develop a single high-quality MCQ. Therefore, it would be of great benefit to the training of future physicians if this process could become (at least partially) automated.

Large language models (LLMs) in medical education

The concept of the automated creation of exam questions could benefit from the recent advent of LLMs, such as ChatGPT. LLMs are systems that use natural language processing methods to "recognize, interpret, and generate text."[5(p2930)] Following the recent hype around these artificial intelligence-based systems, which began with the release of OpenAI's ChatGPT in November 2022, a number of use cases have demonstrated how LLMs (and ChatGPT in particular) have been used to achieve results in various domains. In health care, for example, ChatGPT has been used in efforts to improve doctor–patient communication and simplify clinical management processes.

In addition to these more general applications, the advantages and disadvantages of the use of LLMs in medical education and continuing medical education have been discussed in detail. For example, Khan and colleagues describe 8 potential areas of ChatGPT application in medical education, including "teaching assistance," "personalized learning," and "creating

Please see the end of this article for information about the authors.

Correspondence should be addressed to Matthias Carl Laupichler, Institute of Medical Education, University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany; telephone: +49 228 287 52167; email: matthias.laupichler@ukbonn.de.

Acad Med. 2024;99:508–512.
First published online
doi: 10.1097/ACM.0000000000005626

Received: 27 March 2025 | Revised: 9 May 2025 | Accepted: 17 May 2025
DOI: 10.1002/bcp.70119

SHORT COMMUNICATION

BJCP BRITISH PHARMACOLOGICAL SOCIETY

LLM-based generation of USMLE-style questions with ASPET/AMSPC knowledge objectives: All RAGs and no riches

Thomas Thesen[1] | Rupa Lalchandani Tuan[2] | Joe Blumer[3] | Michael W. Lee[1]

[1]Department of Medical Education, Geisel School of Medicine, Dartmouth, New Hampshire, USA
[2]Department: Cellular and Molecular Pharmacology, University of California, San Francisco, California, USA
[3]Department of Cell and Molecular Pharmacology and Experimental Therapeutics, Medical University of South Carolina, Charleston, South Carolina, USA

Correspondence
Michael W. Lee, Department of Medical Education, Geisel School of Medicine, Dartmouth, New Hampshire, USA.
Email: michael.w.lee@dartmouth.edu

Developing high-quality pharmacology multiple-choice questions (MCQs) is challenging in large part due to continually evolving therapeutic guidelines and the complex integration of basic science and clinical medicine in this subject area. Large language models (LLMs) like ChatGPT-4 have repeatedly demonstrated proficiency in answering medical licensing exam questions, prompting interest in their use for generating high stakes exam-style questions. This study evaluates the performance of ChatGPT-4o in generating USMLE-style pharmacology questions based on American Society for Pharmacology and Experimental Therapeutics/Association of Medical School Pharmacology Chairs (ASPET/AMSPC) knowledge objectives and assesses the impact of retrieval-augmented generation (RAG) on question accuracy and quality. Using standardized prompts, 50 questions (25 RAG and 25 non-RAG) were generated and subsequently evaluated by expert reviewers. Results showed higher accuracy for non-RAG questions (88.0% *vs.* 69.2%), though the difference was not statistically significant. No significant differences were observed in other quality dimensions. These findings suggest that sophisticated LLMs can generate high-quality pharmacology questions efficiently without RAG, though human oversight remains crucial.

KEYWORDS
artificial intelligence, ASPET/AMSPC knowledge objectives, pharmacology, retrieval-augmented generation

1 | INTRODUCTION

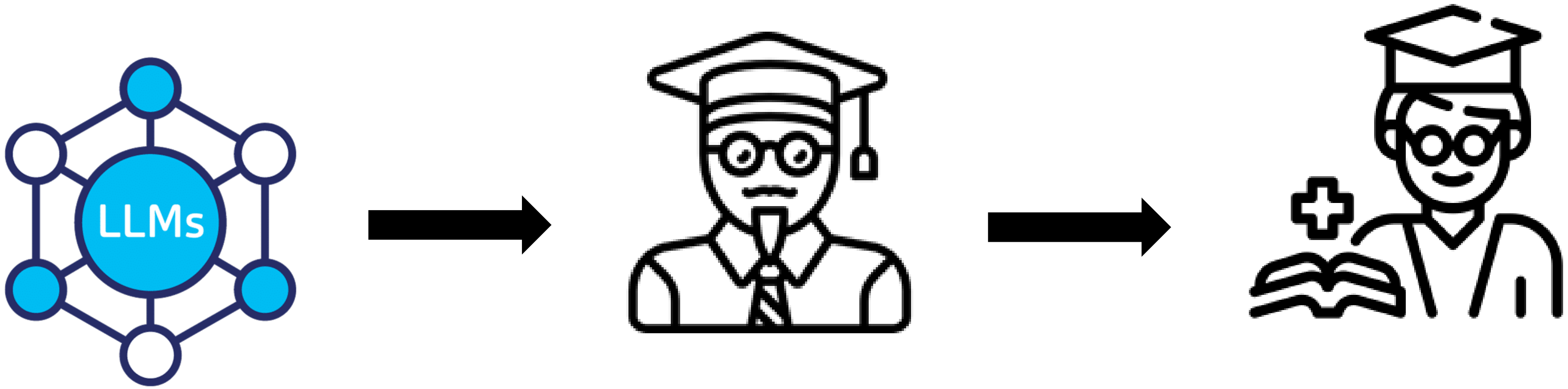
Developing high-quality multiple-choice questions (MCQs) for exams like the USMLE or COMLEX are fraught with significant challenges: they require substantial subject matter expertise, consume considerable faculty time and necessitate rigorous validation to ensure accuracy and educational value. In pharmacology specifically, question creation presents additional complexities due to the rapidly evolving nature of therapeutic guidelines, new drug approvals, drug interactions and complex integration between basic science and clinical medicine.[1,2]

Large language models (LLMs) such as ChatGPT-4 have shown they can answer medical exam questions at or above passing levels, prompting interest in using LLMs to generate realistic exam-style questions.[3–5] Recent studies suggest LLM-generated medical questions can approximate the quality of faculty-authored questions.[5,6] In a blinded evaluation, Cheung et al. compared 50 MCQs generated by ChatGPT (GPT-3.5) against 50 questions crafted by medical faculty.[6] Three expert assessors evaluated these questions across five quality domains, finding statistical equivalence in overall quality and four specific domains, with only marginal inferiority in content relevance. Notably, LLM question generation required approximately 20 min compared with 211 min for human experts—a tenfold efficiency improvement. Similarly, ChatGPT-4-generated clinical vignettes were largely indistinguishable from human-authored questions when evaluated by medical experts.[7] Despite these promising results, 36% of ChatGPT-4-generated questions contained flaws identified by physician reviewers that make their deployment reliant on a human-in-the-loop check.

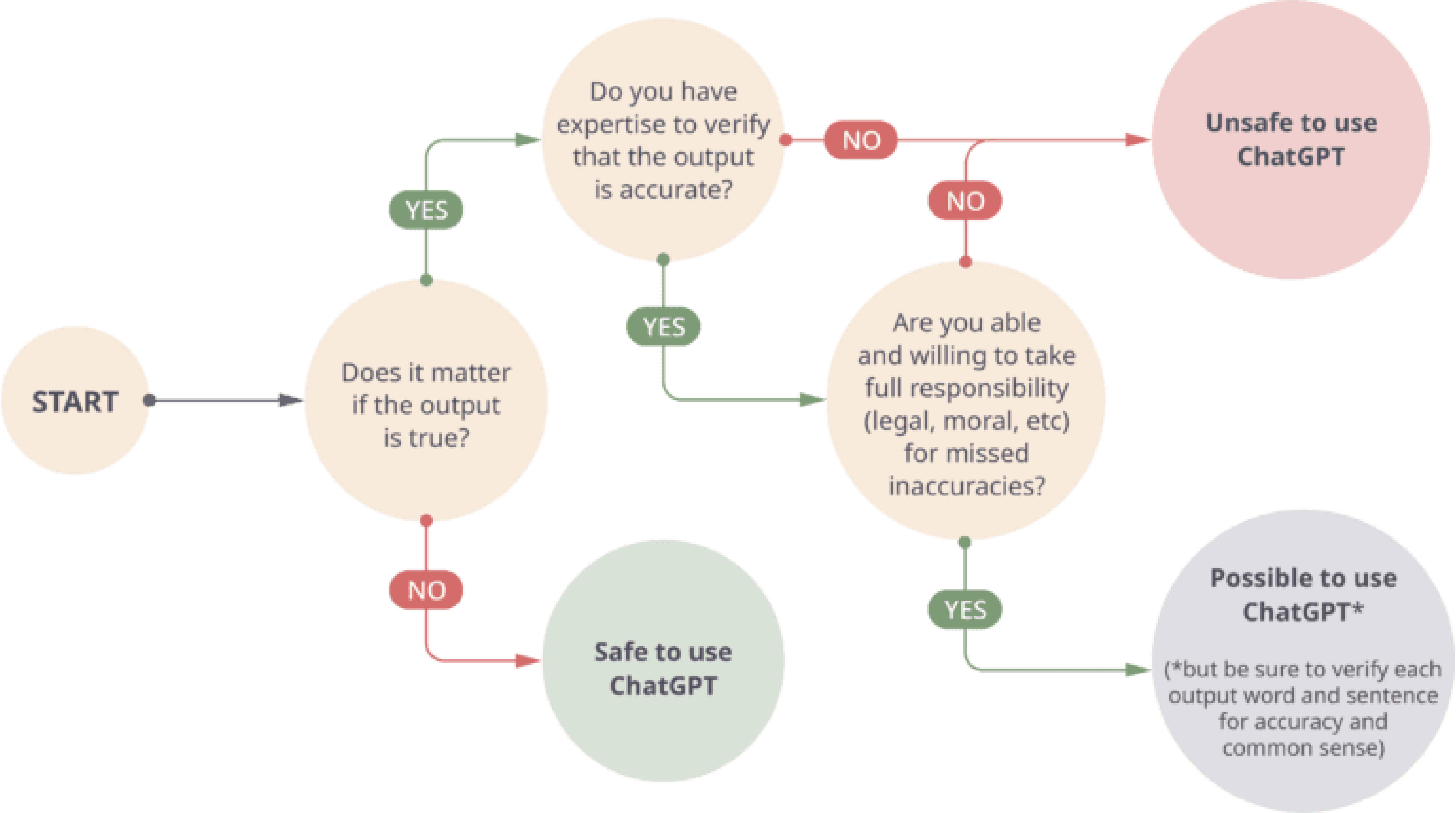
Given the design and training of LLMs, one can assume that LLM performance varies significantly across medical specialties for several reasons. Training data typically overrepresents common conditions

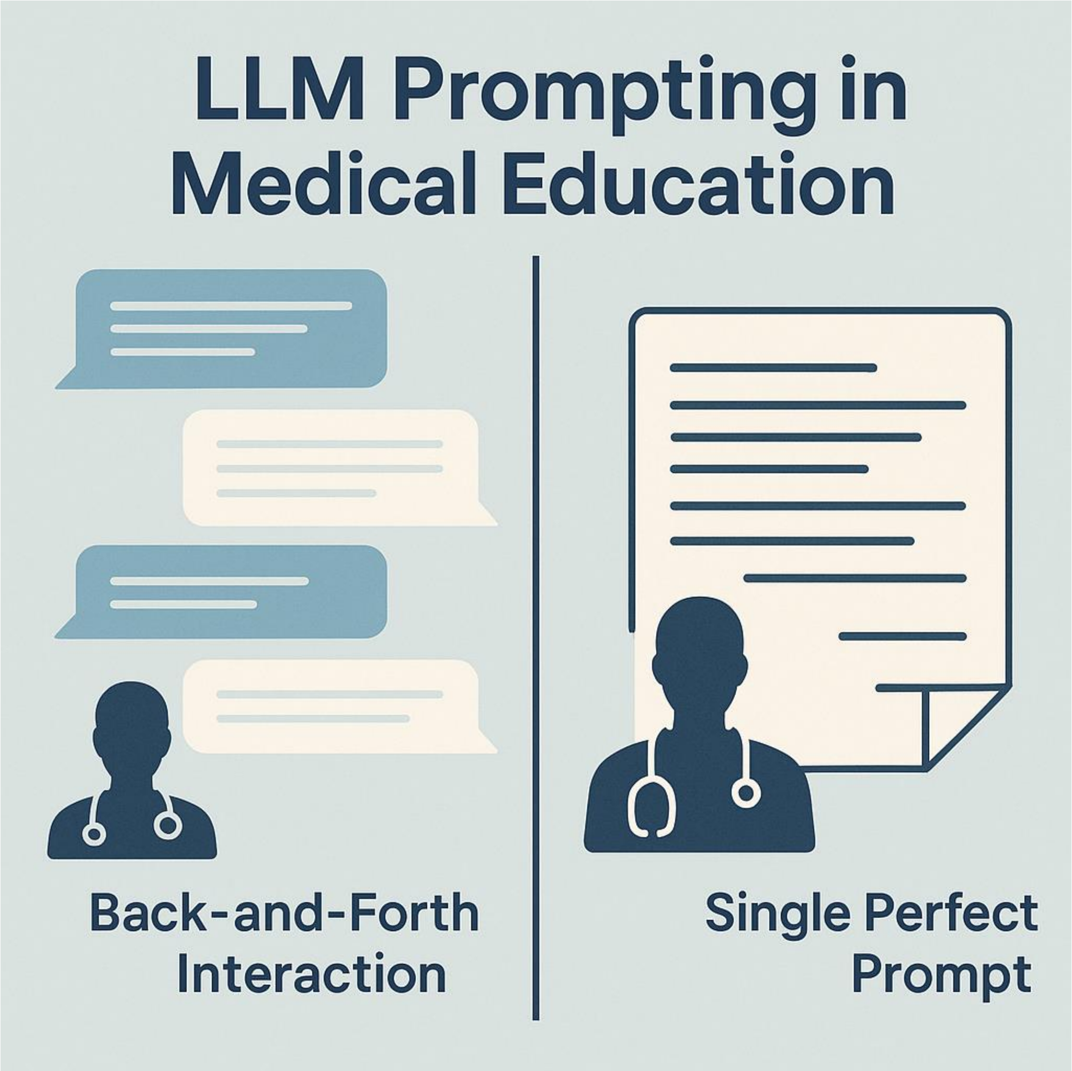
Professor in the Loop

- LLMs make mistakes that are hard to spot by medical students
- Output validation and/or editing by experts is required



Responsible Use – ChatGPT & other LLMs


START
Does it matter if the output is true?
YES
Do you have expertise to verify that the output is accurate?
NO
Unsafe to use ChatGPT
YES
NO
Are you able and willing to take full responsibility (legal, moral, etc) for missed inaccuracies?
NO
Safe to use ChatGPT
YES
Possible to use ChatGPT*
(*but be sure to verify each output word and sentence for accuracy and common sense)


LLM Prompting in Medical Education
Back-and-Forth Interaction
Single Perfect Prompt

RODES Prompting Framework

R - Role: [Define the AI’s role to set the tone and perspective of the response]

O - Objective: [Clear articulate the goal of the prompt, focusing the AI’s efforts]

D - Details: [Provide specific details or parameters to guide the AI’s response]

E - Examples: Here are good examples you can use to model your answer.
[Use examples to illustrate the desired style, tone, and format of the output]

S - Sense Check: Confirm the AI’s understanding of the prompt, ensuring alignment before execution

Role: You are an experienced biomedical science educator and course director at a US medical school teaching medical students.

Objective: Develop a USMLE Step 1-style question focused the following learning objective:

Relate clinical correlations to the underlying functional and anatomical organization of the somatic sensory system and describe their diagnostic value in the identification and localization of the disease processes.

Details:

- The correct answer should be: Brown-Sequard Syndrome at T10.
- Make the multiple-choice questions appropriate for 2nd year medical students preparing for STEP 1
- Include relevant patient history, physical exam findings, and any necessary laboratory results or diagnostic studies.
- The 5 answer choices (A–E) should include plausible distractors that test high-yield concepts.
- Make sure the explanation of the correct answer includes the key concepts behind both the right and wrong options.
- Think step-by-step

Examples: Here is an examples of a good USMLE Step 1-style question:

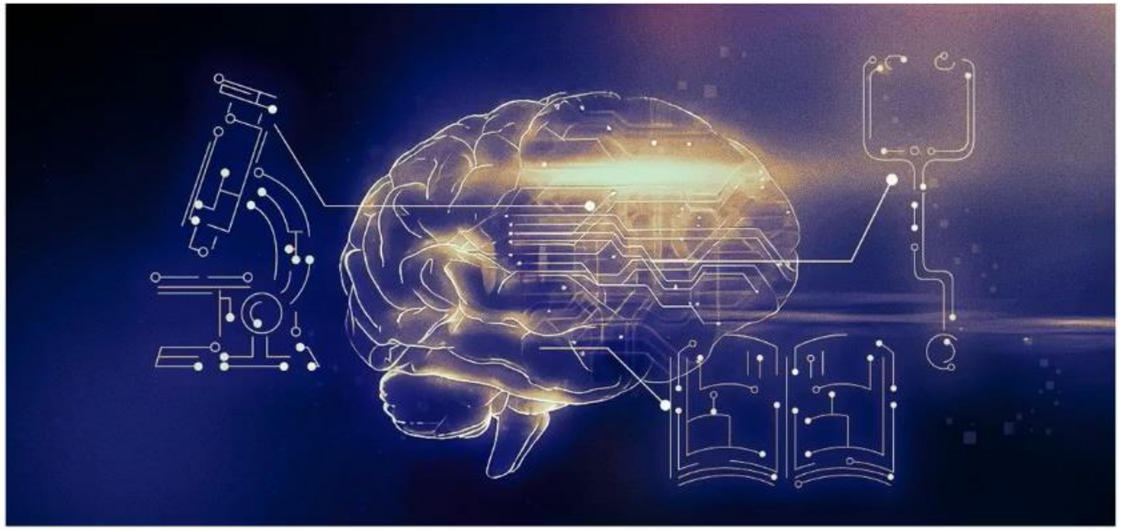
A 67-year-old man presents to the emergency department with sudden-onset chest pain that radiates to his left arm. He is diaphoretic and pale. An ECG shows ST-segment elevations in leads II, III, and aVF. Which of the following coronary arteries is most likely occluded?
A) Left anterior descending artery
B) Left circumflex artery
C) Right coronary artery
D) Left marginal artery
E) Posterior descending artery

(Explanation: The correct answer is C. The patient's symptoms and ECG findings are consistent with an acute inferior myocardial infarction, which is most commonly due to occlusion of the right coronary artery. Distractors A and B point to other coronary vessels that are involved in different infarct locations. The explanation should explain the pathophysiology of myocardial ischemia.)

Sense Check: Do you understand the objective? Ask clarifying details,.

Dartmouth GEISEL SCHOOL OF MEDICINE | ACADEMY OF EDUCATORS AND SCHOLARS

Generative AI for Medical Education Prompt-a-Thon



Thomas Thesen, Ph.D.

&

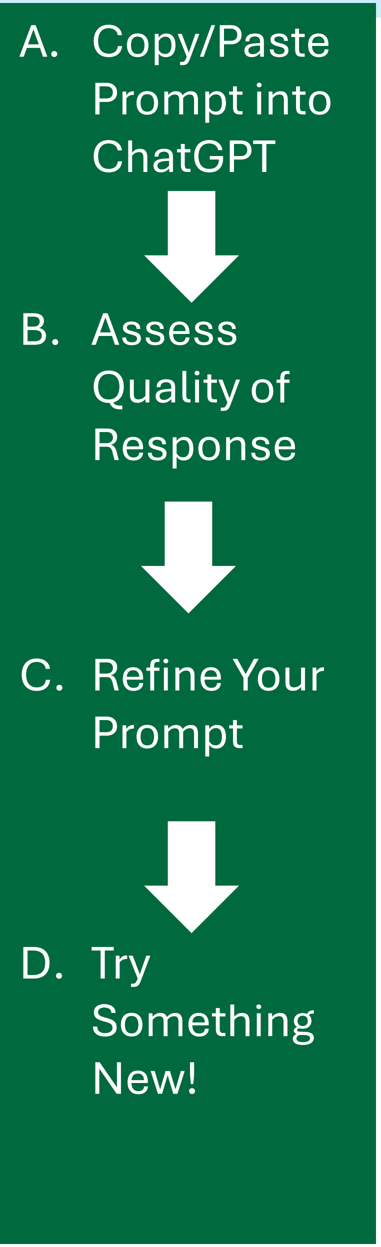
Amy Synthesia, A.I.

Group Activity # 1

Create a prompt for vignettes

1. Split into groups of 2
2. Share one laptop
3. Go to chatgpt.com
4. Create a blank **Word document**
5. Scan QR code or go to tinyurl.com/MakeVignette
6. Follow instructions & work together

Case #1
Role: Medical Educator
Objective: Create a Clinical Vignette Question
Go to https://chatgpt.com/



A. Copy/Paste Prompt into ChatGPT

B. Assess Quality of Response

C. Refine Your Prompt

D. Try Something New!

#1 User Prompt

Role: You are an experienced biomedical science educator and course director at a US medical school teaching medical students.

Objective: Develop a USMLE Step 1-style question focused the following learning objective: "*Relate clinical correlations to the underlying functional and anatomical organization of the somatic sensory system and describe their diagnostic value in the identification and localization of the disease processes*".

Details:
The correct answer should be: Brown-Sequard Syndrome at T10.
•Include relevant patient history, physical exam findings, and any necessary laboratory results or diagnostic studies.
•The 5 answer choices (A–E) should include plausible distractors that test high-yield concepts.
•Keep the question at an appropriate difficulty level for 3rd year medical students.
•Make sure the explanation of the correct answer includes the key concepts behind both the right and wrong options.
•Think step-by-step.

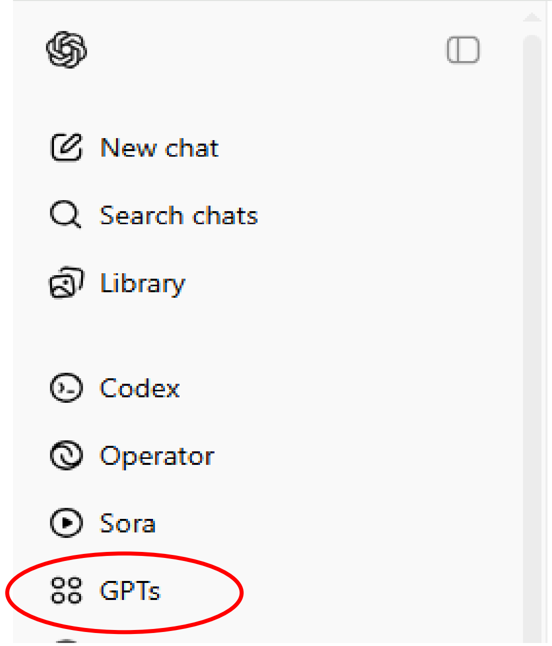
Examples: Here is an examples of a good USMLE Step 1-style question:
A 67-year-old man presents to the emergency department with sudden-onset chest pain that radiates to his left arm. He is diaphoretic and pale. An ECG shows ST-segment elevations in leads II, III, and aVF. Which of the following coronary arteries is most likely occluded?
A) Left anterior descending artery
B) Left circumflex artery
C) Right coronary artery
D) Left marginal artery
E) Posterior descending artery
(Explanation: The correct answer is C. The patient's symptoms and ECG findings are consistent with an acute inferior myocardial infarction, which is most commonly due to occlusion of the right coronary artery. Distractors A and B point to other coronary vessels that are involved in different infarct locations. The explanation should explain the pathophysiology of myocardial ischemia.)

Sense Check: Do you understand the objective and the specific guidelines for creating this USMLE Step 1-style question? Do you understand the reasoning behind the correct answer?

Refinement suggestions:
- **Make it more relevant to your specialty**
- **Make it more relevant to your teaching at Geisel**
- **Modify the output format the way you prefer**

Group Discussion

Custom GPT Demonstration



New chat

Search chats

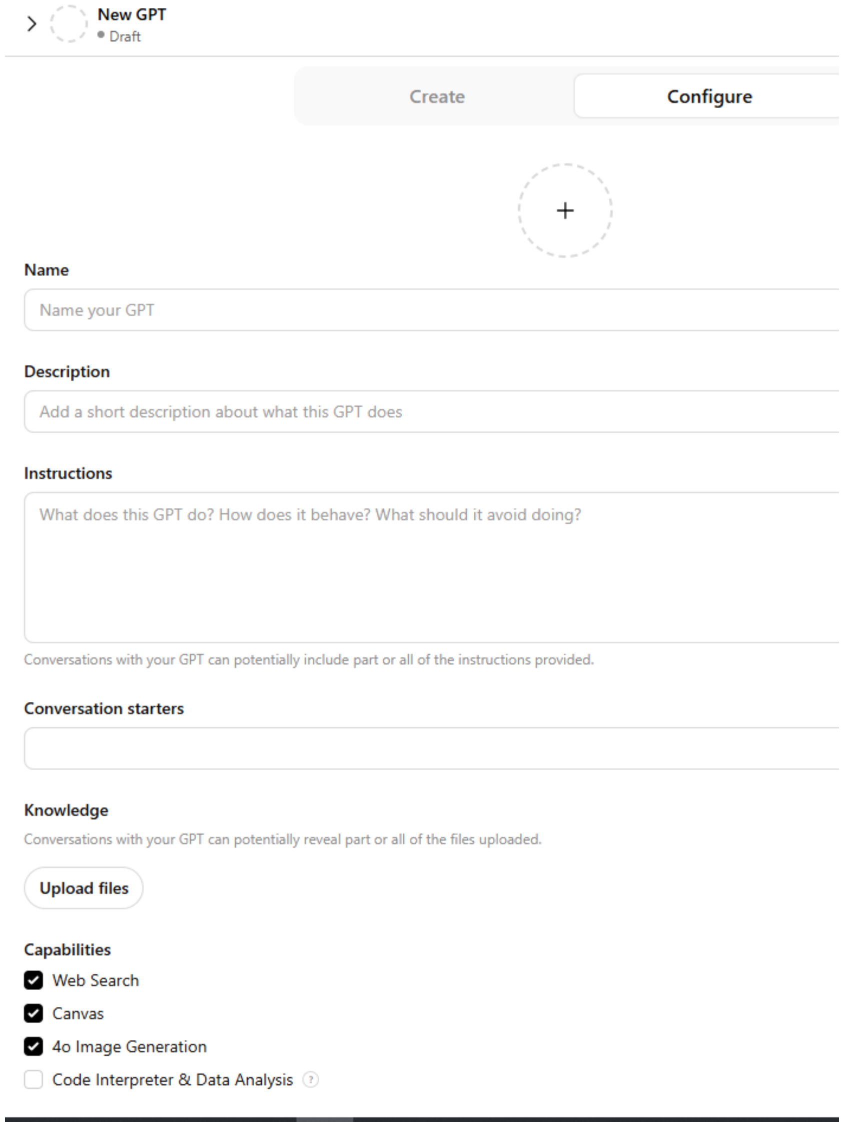
Library

Codex

Operator

Sora

GPTs



New GPT
• Draft

Create | Configure

Name

Name your GPT

Description

Add a short description about what this GPT does

Instructions

What does this GPT do? How does it behave? What should it avoid doing?

Conversations with your GPT can potentially include part or all of the instructions provided.

Conversation starters

Knowledge

Conversations with your GPT can potentially reveal part or all of the files uploaded.

Upload files

Capabilities

☑ Web Search

☑ Canvas

☑ 4o Image Generation

□ Code Interpreter & Data Analysis

Turning your perfect prompt into an AI Agent

Dartmouth GEISEL SCHOOL OF MEDICINE | ACADEMY OF EDUCATORS AND SCHOLARS

1. Create a RODES prompt and iteratively make it better
2. Adjust the prompt to allow for user input, if needed
3. Create a Custom GPT at chatgpt.com (you need a paid account)
4. Share the GPT with your students or colleagues (free account for them is ok)
5. Alternative: If you do not have a paid account, create a repository of prompts in a Word document for later use or distribution

Prompt:

- **Role:** You are an experienced biomedical science educator and course director at a US medical school teaching medical students.
- **Objective:** Develop a USMLE Step 1-style question focused the following learning objective:
- **Relate clinical correlations to the underlying functional and anatomical organization of the somatic sensory system and describe their diagnostic value in the identification and localization of the disease processes.**
- **Details:**
- •The correct answer should be: Brown-Sequard Syndrome at T10.
- •Make the multiple-choice questions appropriate for 2nd year medical students preparing for STEP 1
- •Include relevant patient history, physical exam findings, and any necessary laboratory results or diagnostic studies.
- •The 5 answer choices (A–E) should include plausible distractors that test high-yield concepts.
- •Make sure the explanation of the correct answer includes the key concepts behind both the right and wrong options.
- •Think step-by-step
- **Examples:** Here is an examples of a good USMLE Step 1-style question:
- **A 67-year-old man presents to the emergency department with sudden-onset chest pain that radiates to his left arm. He is diaphoretic and pale. An ECG shows ST-segment elevations in leads II, III, and aVF. Which of the following coronary arteries is most likely occluded?**
  A) Left anterior descending artery
  B) Left circumflex artery
  C) Right coronary artery
  D) Left marginal artery
  E) Posterior descending artery
- *(Explanation: The correct answer is C. The patient's symptoms and ECG findings are consistent with an acute inferior myocardial infarction, which is most commonly due to occlusion of the right coronary artery. Distractors A and B point to other coronary vessels that are involved in different infarct locations. The explanation should explain the pathophysiology of myocardial ischemia.)*

Turning your perfect prompt into an AI Agent

Demonstration

Prompt:

- <u>**Role:**</u> You are an experienced biomedical science educator and course director at a US medical school teaching medical students.
- <u>**Objective:**</u> **Develop a USMLE Step 1-style question focused on a learning objective or medical disorder.**
- **Ask the user first: "Enter the clinical disorder or learning objective for which you want to create a USMLE question". Then use this input to create the question.**
- <u>**Details:**</u>
- ••Make the multiple-choice questions appropriate for 2nd year medical students preparing for STEP 1
- •Include relevant patient history, physical exam findings, and any necessary laboratory results or diagnostic studies.
- •The 5 answer choices (A–E) should include plausible distractors that test high-yield concepts.
- •Make sure the explanation of the correct answer includes the key concepts behind both the right and wrong options.
- •Think step-by-step
- <u>**Examples:**</u> Here is an examples of a good USMLE Step 1-style question:
- **A 67-year-old man presents to the emergency department with sudden-onset chest pain that radiates to his left arm. He is diaphoretic and pale. An ECG shows ST-segment elevations in leads II, III, and aVF. Which of the following coronary arteries is most likely occluded?**
  A) Left anterior descending artery
  B) Left circumflex artery
  C) Right coronary artery
  D) Left marginal artery
  E) Posterior descending artery
- *(Explanation: The correct answer is C. The patient's symptoms and ECG findings are consistent with an acute inferior myocardial infarction, which is most commonly due to occlusion of the right coronary artery. Distractors A and B point to other coronary vessels that are involved in different infarct*

Group Activity # 2

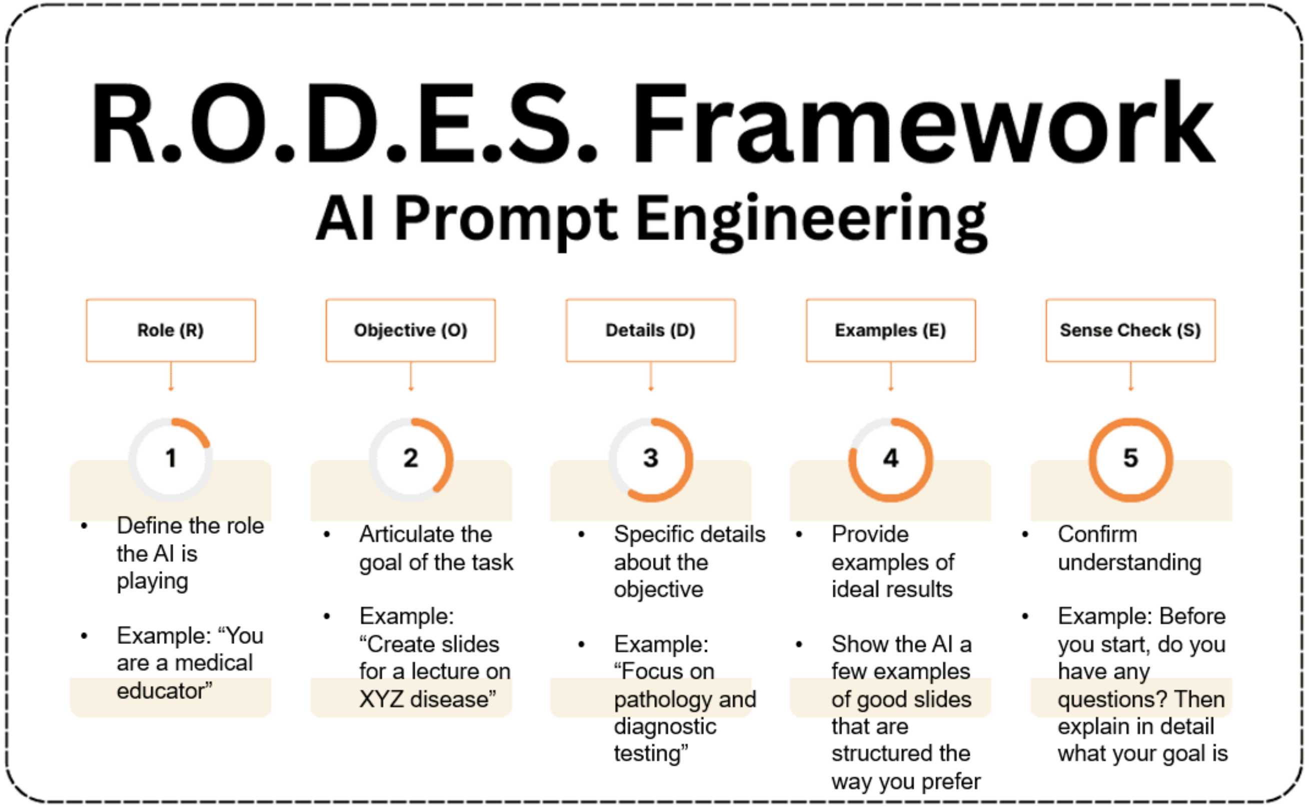
Create a prompt for a clinical case

1. Split into groups of 2
2. Share one laptop
3. Go to chatgpt.com & create a new Word document

4. Use the RODES Model to create a perfect prompt for a task directly related to your work

5. Create a Custom GPT or a prompt that you can share

6. **Demonstrate to the group**



R.O.D.E.S. Framework

AI Prompt Engineering

Role (R)

1

- Define the role the AI is playing
- Example: "You are a medical educator"

Objective (O)

2

- Articulate the goal of the task
- Example: "Create slides for a lecture on XYZ disease"

Details (D)

3

- Specific details about the objective
- Example: "Focus on pathology and diagnostic testing"

Examples (E)

4

- Provide examples of ideal results
- Show the AI a few examples of good slides that are structured the way you prefer

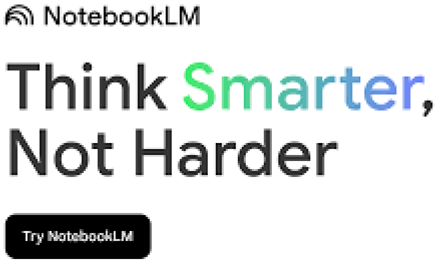
Sense Check (S)

5

- Confirm understanding
- Example: Before you start, do you have any questions? Then explain in detail what your goal is

NotebookLM

https://notebooklm.google.com/



From Gemini:

- NotebookLM is an AI-powered note-taking and research assistant from Google that utilizes the Gemini 1.5 Pro model
- It allows users to upload documents and then generates personalized summaries, guides, and audio overviews based on the uploaded content
- NotebookLM lets users to summarize complex information, ask questions, and explore connections within their sources

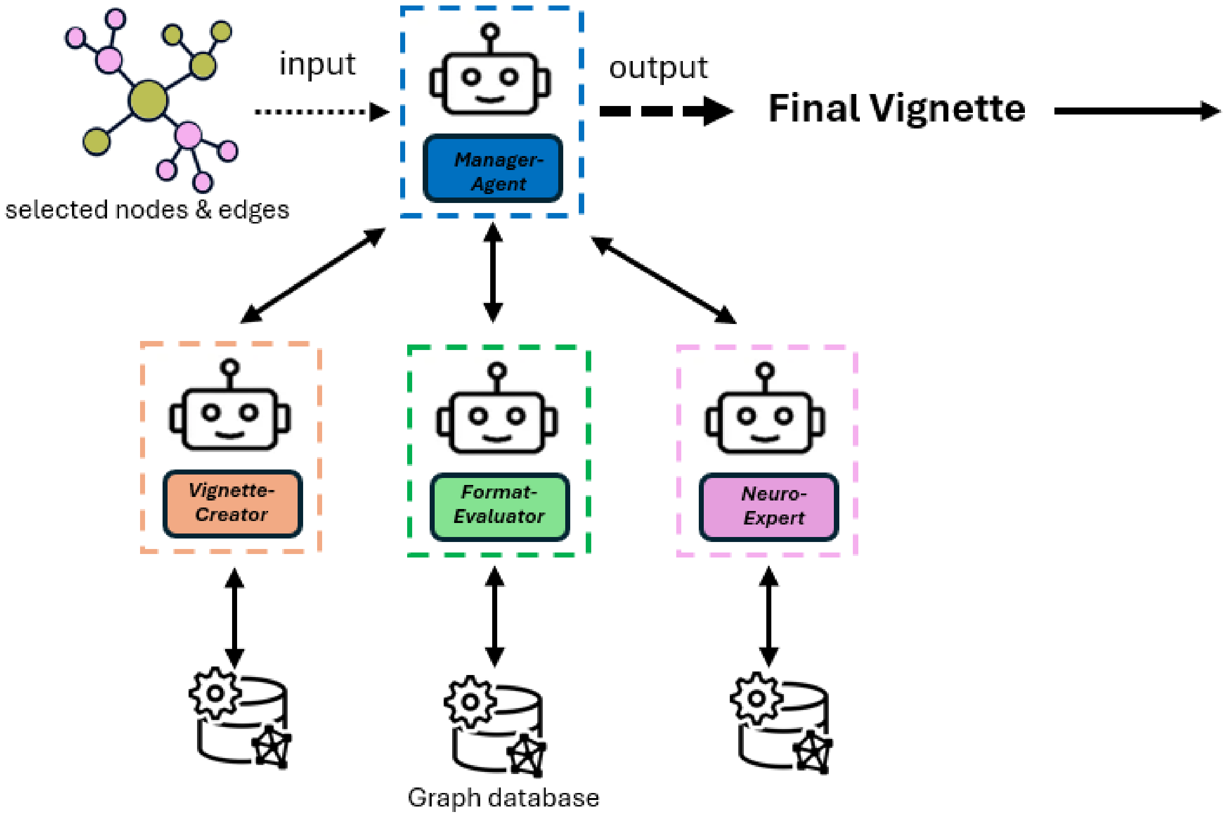
Group Activity # 2

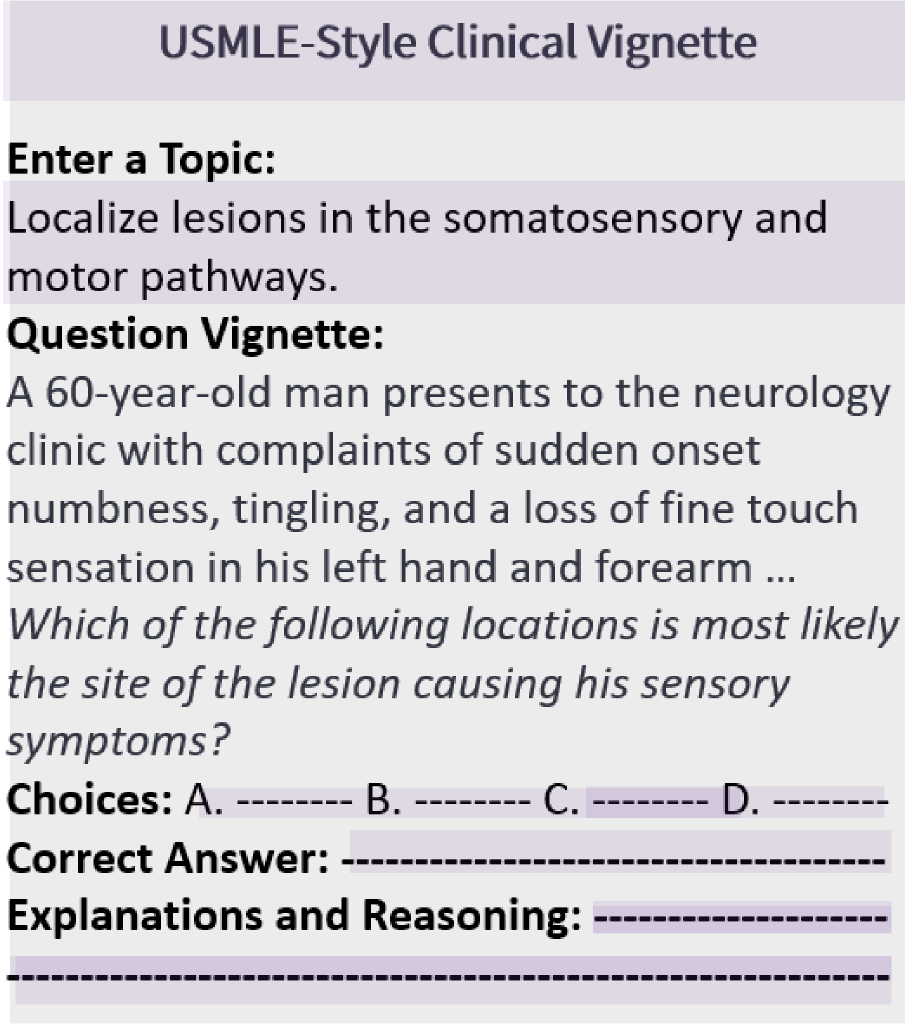
Explore NotebookLM

1. Go to https://notebooklm.google.com/
2. Upload one of your documents
3. Create a study guide for this document
4. Create a podcast from this document
5. Query the document through an interactive Mindmap

Mutli-Agent Framework

https://multiagentllmplatform.streamlit.app/





USMLE-Style Clinical Vignette

Enter a Topic:

Localize lesions in the somatosensory and motor pathways.

Question Vignette:

A 60-year-old man presents to the neurology clinic with complaints of sudden onset numbness, tingling, and a loss of fine touch sensation in his left hand and forearm … *Which of the following locations is most likely the site of the lesion causing his sensory symptoms?*

Choices: A. -------- B. -------- C. -------- D. --------

Correct Answer: ------------------------------------

Explanations and Reasoning: -------------------

Creating your own AI Agents for Teaching

Thomas Thesen, Ph.D.

Associate Professor
Departments of Medical Education &
Computer Science
Geisel School of Medicine at Dartmouth

thomas.thesen@dartmouth.edu **https://geiselmed.dartmouth.edu/thesen/**