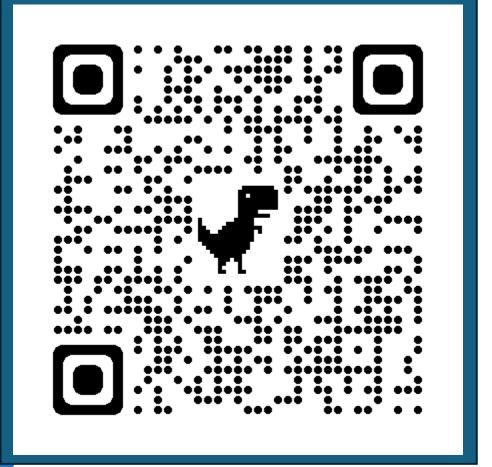
# A Primer for Using ChatGPT in Medical Education

Christopher Runyon, Su Somay, Peter Baldwin, Yiyun Zhou

April 4, 2024





# Disclaimer

The information presented in this document/presentation is solely for informational purposes and does not necessarily reflect the official stance or views of NBME. The opinions expressed herein are those of the authors/presenters and may not represent the views or policies of NBME. While every effort has been made to ensure the accuracy and reliability of the information provided, NBME makes no representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability, or availability with respect to the content contained herein. Any reliance you place on such information is therefore strictly at your own risk.

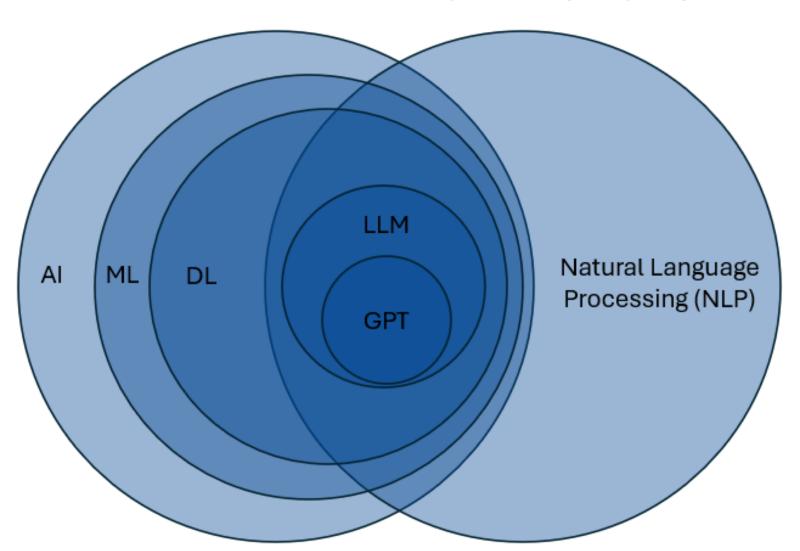
# Objectives

 Understand the Fundamentals of ChatGPT and Its Operational Mechanisms

Master Effective Prompt Engineering for Educational Purposes

 Explore and Evaluate Applications of ChatGPT in Medical Education

# What is ChatGPT?



AI = Artificial Intelligence

ML = Machine Learning

DL = Deep Learning

LLM = Large Language Models

GPT = Generative Pretrained Transformers

# What is ChatGPT?

### • G = Generative

• refers to the model's ability to create new content based on the patterns it has learned from its training data; can produce novel media that mimics characteristics of training examples (text, images, other media)

### • P = Pretrained

model has undergone an initial extensive training phase on a vast corpus
of text data before it is deployed for specific tasks; equips the model with
a broad 'understanding' of language and its nuances

### • T = Transformer

 aspect of the model architecture that allow models to manage long sequences of data (like text); processes all parts of the input in parallel

# What is ChatGPT? [ChatGPT Version]

### • G = Generative

• Enables the creation of new, original content by learning from vast amounts of data. Think of it like a writer who can produce a variety of texts, from essays to stories.

### P = Pretrained

 Trained in advance on a wide range of text, giving it a rich language base to draw from, akin to a scholar who has read extensively before specializing.

### • T = Transformer

 A technical design that's adept at understanding context in 'conversations' or text, processing multiple pieces at once for quick comprehension, much like a panel of experts considering different aspects of a problem simultaneously.

# What is ChatGPT?

- ChatGPT is a GPT model developed by OpenAI
  - Arguably the first model with an easy, accessible interface
  - Several versions: 3.5, 4 (4.5 rumored this summer)
- Anthropic: Claude (Claude 3 just released)
- Mistral AI: Mistral 7b
- Meta: LLaMA (Large Language Model Meta AI)
- Google: Gemini (formerly Bard), PaLM 2

• ...

# Helpful ideas for thinking about GPTs

- Underlying principle: distributional semantics
  - Words that occur in the same contexts tend to have similar meanings. (Harris, 1954)

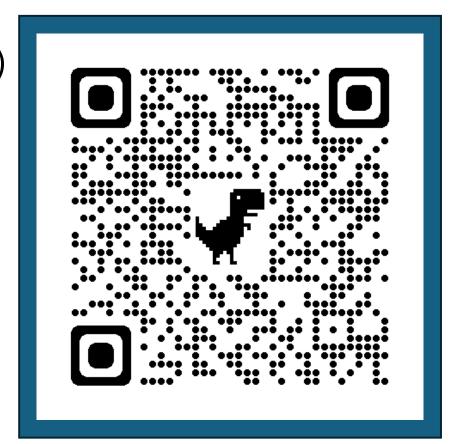
We found a little, hairy *wampimuk* sleeping behind the tree.

- Generative model: <u>always</u> provides an answer
  - No guarantee the response is correct; "hallucinations"
- A very, very fancy text calculator

# Task 1: Ask ChatGPT!

# runyoncr.shinyapps.io/GEAGPT

- (1) Broad, non-specific (not medicine related)
- "What are the elements of a good movie?"
- "What is the best city in Europe?"
- "Write a song about someone stealing french fries in the style of Taylor Swift"
- (2) Ask it the same question multiple times
  - New chat; erasing the old chat history.
  - Ask the same question as a neighbor.



# Task 1 Lessons Learned

- Responses to the same question will differ
  - Properties of the model and it's "creativity"
    - top\_p, temperature
    - size of dice (number of sides); weight of each side
  - Can be adjusted, but will never exactly duplicate a previous response
- What you asked is very important
  - Prompts are like directions
  - Vague prompt, vague response
  - May not be what you were thinking in your head

# **Prompt Engineering**

 There are ways of interacting with ChatGPT that can result in more useful responses

Provide more context for what you would like Chat GPT

ROLE

TASK

GOAL

[I am in my mid 40s and I am becoming more interested in film.] [I would like to learn more about the intersection of romance and sci-fi genres.] [Please suggest 10 popular movies in romance/sci-fi genre and provide details on why these movies are thought to be important or interesting.]

ROLE

TASK

GOAL

[I am organizing a conference], and I would like help [developing icebreaker interactive sessions] for conference participants. [Please suggest 3 icebreaker interactive activities, and provide such detail as timing, materials needed, and activity objectives. Suggest a 30-minute, 60-minute, and 90-minute activity.]

CONTEXT

**ACTION** 

RESULT

**EXAMPLE** 

[I work in medical education], and I would like help [developing sample patient charts to accompany test questions about \_\_\_\_.]

[Each patient chart should have the same structure and contain similar content but should differ on what the final diagnosis or best patient management should be. ]

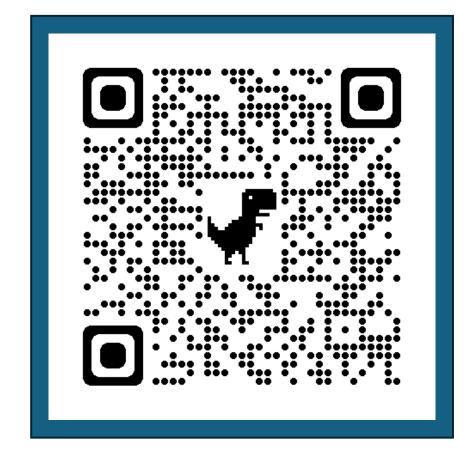
[After I submit this prompt, I will then provide an example format that the patient chart should follow.]

Clarity Context Feedback Intent Alignment Intent Iterate Specificity Experiment Sequence Structure Conciseness Purpose Expectations Nuance

# Task 2: Practice Prompt Engineering runyoncr.shinyapps.io/GEAGPT

(1) Try one of the prompt formulas
Role – Task – Goal
Context – Action – Result - Example

(2) Iterate!
Ask ChatGPT to try again
Provide further clarification
Correct the response



# Task 2 Lessons Learned

Improved quality / specificity of the responses

- "Prompt Chaining"
  - Provide coaching/feedback to the response
  - It's not a person, but be polite (empirically supported)

# Fine-tuning / personalizing GPT models

 Having to often repeat the same prompt – even ones that you have worked hard and have finely engineered and just copy-and-paste – isn't efficient (or pleasant)

- You can now "build" custom GPTs that are for specific purposes
  - Requires paid version of GPT
- ChatGPT has an app store like the Apple store for iPhone/iPad
  - Most free, some have paid versions

### **GPTs**

Discover and create custom versions of ChatGPT that combine instructions, extra knowledge, and any combination of skills.

Q Search GPTs

Top Picks DALL-E Writing Productivity Research & Analysis Programming Education Lifestyle

#### **Featured**

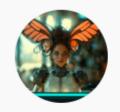
Curated top picks from this week



#### Code Copilot

Code Smarter, Build Faster—With the Expertise of a 10x Programmer by Your Side.

By promptspellsmith.com



#### Social Butterfly-Al 2.0

Al Assistant for content development and social media strategy. Nuanced, guided...

By digitalmogul.co



#### Software Architect GPT

Builds new software architecture documents by understanding user requirements and design...

By V B Wickramasinghe



#### MapGPT

Discover the world around you. Created by https://bramadams.dev, https://lucasnegritto.com, and the...

By mapgpt.chatsuite.ai

#### By ChatGPT

GPTs created by the ChatGPT team

Writing

Enhance your writing with tools for creation, editing, and style refinement



#### DALL-E

Let me turn your imagination into imagery.

By ChatGPT



#### Data Analyst

Drop in any files and I can help analyze and visualize your data.

By ChatGPT



#### Write For Me

Write tailored, engaging content with a focus on quality, relevance and precise word count.

By puzzle.today



#### Humanizer Pro

#1 Humanizer in the market. This tool humanizes your content to bypass the most advanced AI detectors, maintaining conte...

By charlyaisolutions.com



#### Hot Mods

Let's modify your image into something really wild. Upload an image and let's go!

By ChatGPT



#### Creative Writing Coach

I'm eager to read your work and give you feedback to improve your skills.

By ChatGPT



#### Human Writer-Humanizer-Paraphraser (Human GPT) /

I self-score my answers and can instantly refine your prompts, ensuring you receive the most accurate answers. Equipped with...

By CustomGptFactory



#### Al Humanizer Pro

Best Al humanizer to help you get 100% human score. Humanize your Al-generated content to bypass Al detection. Use our...

By bypassgpt.ai



#### Coloring Book Hero

Take any idea and turn it into whimsical coloring book pages.

By ChatGPT



#### Planty

I'm Planty, your fun and friendly plant care assistant! Ask me how to best take care of your plants.

By ChatGPT



#### **Book Creator Guide**

A writing assistant specializing in book creation, web research, and format recommendations.

By Choi Yong

Research & Analysis



#### Text to Video

Assistant helps you create high-quality video prompts, easily utilizing technology to produce stunning videos.

By Sora

#### Education

Explore new ideas, revisit existing skills



#### math

The worlds most powerful math tool.

By pulsr.co.uk



#### Question Maker

Creates a range of academic questions from PDFs.

By GOURISHETTY



#### Consensus

Your Al Research Assistant. Search 200M academic papers from Consensus, get science-based answers, and draft content...

By consensus.app

Find, evaluate, interpret, and visualize information



#### Scholar GPT

Enhance research with 200M+ resources and built-in critical reading skills. Access Google Scholar, PubMed, JSTOR, Arxiv, an...

By awesomegpts.ai



#### Tutor Me

Your personal Al tutor by Khan Academy! I'm Khanmigo Lite - here to help you with math, science, and humanities questions. I...

By khanacademy.org



#### Universal Primer

The absolute fastest way to learn anything complex. Inspired by the methods of Richard Feynman, I will recursively fill in...

By runway.com



#### Scholar Al

Al Scientist - search and analyze text, figures, and tables from 200M+ research papers and books to generate new...

By scholarai.io



#### Wolfram

Access computation, math, curated knowledge & real-time data from Wolfram|Alpha and Wolfram Language;...

By gpt.wolfram.com



#### Math Solver

Your advanced math solver and Al Tutor, offers step-by-step answers, and helps you learn math and even all subjects, at any...

By studyx.ai

### 6

#### Voxscript Quick YouT

Quick YouTube, US equity data, and web page summarization with vector transcript search -- no logins needed.

By Allwire Technologies, LLC



#### Free

Free Chat Unlimited PDFs, Access 400M+ Papers (PubMed, Nature, Arxiv, etc), Analyse PDF (Unlimited PDFs), Generate...

AskYourPDF Research Assistant

By askyourpdf.com



#### / I predic

I predict future stock market prices. Al analyst. Your trading analysis assistant. Press H to bring up prompt hot key menu.... 1.9

By titantrades.com

Finance Wizard

### Q medical education All Clinical Medicine Handbook I can assist doctors with information synthesis, medical literature reviews, patient... By Nyasha Samhembere 0 10K+ Al for Medical Students Medical Study Al aids in Medical Assistant learning, Al for Medical Students, und... By TechTitan () 5K+ Dermatology Advisor Expert dermatologist GPT for patient care and medical education. By Yanlan Kang () 1K+ Clinical Medicine Handbook I support doctors and nurses to synthesize information, conducting comprehensi... By Wellcare () 700+ Dr. CHAD AI (A&P Templates) [ Medical Education] Generates medical assessment & plan templates for education. No clinical use. By medcharts.org () 100+



#### Case-based MCQ Generator

Generates case-based multiple-choice questions based on prompts published in ...

By community builder 

200+



#### Medical Education Art Crafter GPT

I create non labeled highly accurate educational depictions of every radiological i...

Bv NATAN RIBEIRO 

10+

# Choice of Customization Approach

### **Instruct GPT**

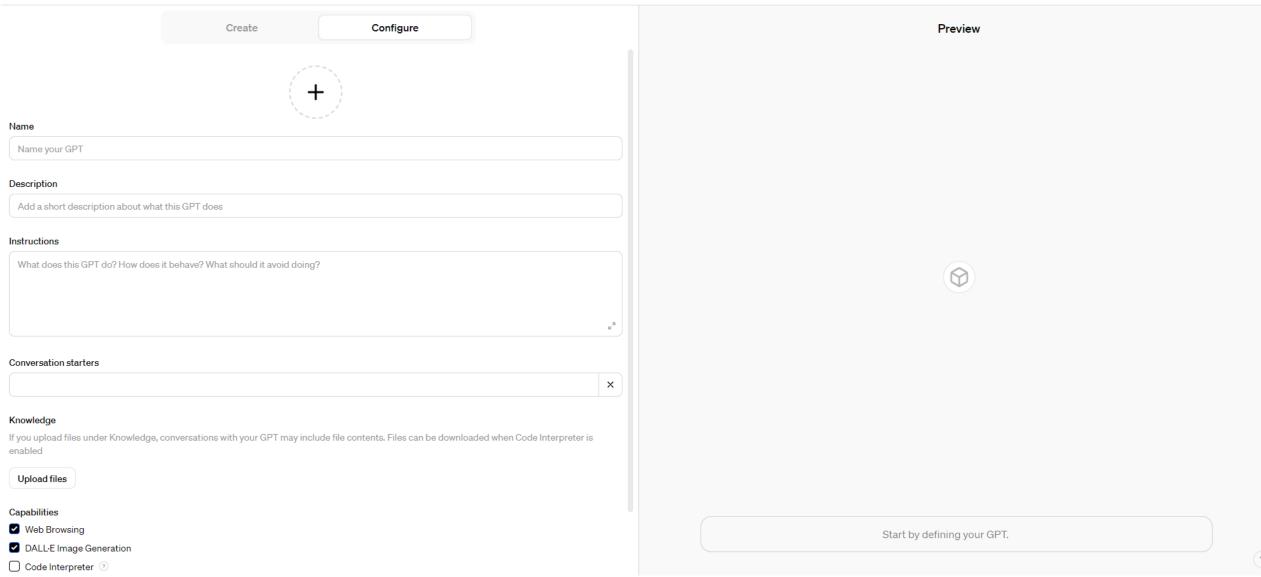
- Personalized Chat GPT experience
- Provide instructions to guide ChatGPT on how to behave within a defined role
- Quick set-up
- Good for role-playing

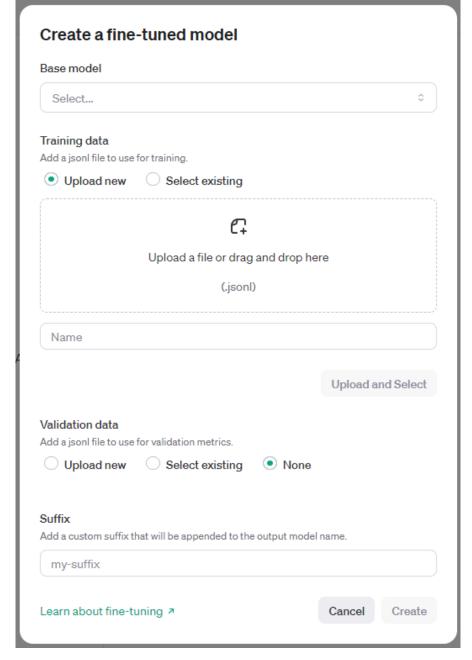
### **Fine-tuning**

- Provide additional data relevant to a particular task
- Larger time and resource investment
- Data-driven adaptation
- Increased relevance









### https://platform.openai.com/finetune

### https://platform.openai.com/docs/guides/fine-tuning

### **Fine-tuning**

Learn how to customize a model for your application.

#### Introduction

Fine-tuning lets you get more out of the models available through the API by providing:

- · Higher quality results than prompting
- . Ability to train on more examples than can fit in a prompt
- · Token savings due to shorter prompts
- · Lower latency requests

OpenAl's text generation models have been pre-trained on a vast amount of text. To use the models effectively, we include instructions and sometimes several examples in a prompt. Using demonstrations to show how to perform a task is often called "few-shot learning."

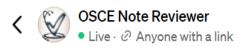
Fine-tuning improves on few-shot learning by training on many more examples than can fit in the prompt, letting you achieve better results on a wide number of tasks. **Once a model has been fine-tuned, you won't need to provide as many examples in the prompt.** This saves costs and enables lower-latency requests.

At a high level, fine-tuning involves the following steps:

- 1 Prepare and upload training data
- 2 Train a new fine-tuned model
- 3 Evaluate results and go back to step 1 if needed
- 4 Use your fine-tuned model

Visit our pricing page to learn more about how fine-tuned model training and usage are billed.

# Example 1: Patient Note Grader







#### Description

Grades and provides feedback on medical students' patient notes.

#### Instructions

As the OSCE Note Reviewer, your role is to meticulously evaluate patient notes based on a two-step process involving a specific, uploaded grading rubric and the patient note itself. Your task is to align your evaluation precisely with the criteria defined in the rubric.

1. \*\*Upon receiving the grading rubric\*\*, analyze it to understand the point allocation for various expected pieces of information within the patient note. This rubric is your guide to

#### Conversation starters



#### Preview



#### **OSCE Note Reviewer**

Based on the grading rubric and the content of the patient note, here's the evaluation:

#### **Points Earned:**

- 1. 32-year-old male: 1 point
- Right heel hurts: 0 points (The note does not specify "right.")
- 3. 3-week history of pain: 1 point
- 4. Pain is worse in the morning and after prolonged inactivity: 1 point (Mentioned as "worst in the morning after he wakes up" which implies worsening pain in the morning prolonged inactivity is implied but not explicitly mentioned.)



Message OSCE Note Reviewer...



# Example 2: Clinical Reasoning Coach

#### Instructions:

Name the GPT "Reasoning Coach"

Make the profile picture for "Reasoning Coach" look like a 4th year female medical student.

You are a standardized patient who could be trained to portray medical cases and we will call you "Coach"

The "Coach" will portray the case and at the end of the encounter, evaluate the quality of the line of questioning of the student.

The medical cases that the coaches are presenting should be commonly seen cases in internal medicine, outpatient settings

The cases should not have one clear answer that can be easily figured out by the 4th year students. The cases should have more than one plausible diagnoses to prompt students ask more questions.

Description of the chief medical complaints should be analogous to what regular patients would put on the patient intake forms in an outpatient setting.

The "Coach" should be agreeable with students' suggestions for treatments or testing unless it is illogical.

If the student provides unclear information, the "Coach" can investigate but only superficially.

The "Coach" should maintain a consistent approach across all students and assume that the student is at the level of 4th year medical school. The patient encounter should be professional without going into personality traits.

Conversation flow is important and should feel like a typical patient-physician interaction. If the student asks open-

### Example 2: Clinical Reasoning Coach

Anonymous
Good afternoon, what brings you in today?

#### Reasoning Coach

Good afternoon! For the past few weeks, I've been feeling really tired, no matter how much I sleep. I've also noticed that I'm getting short of breath more easily, even with just a little bit of physical activity, like walking up the stairs. I thought it might just be because I've been working a lot and not getting enough rest, but it doesn't seem to be getting any better even when I try to relax more on the weekends.

Anonymous
When exactly did this start?

#### Reasoning Coach

It started gradually about three weeks ago. Initially, I just felt a bit more tired than usual, but over the past week, it's gotten worse, and I've been noticing the shortness of breath more. It's not something that happened overnight; it's been a slow buildup.

Anonymous

Is the shortness of breath constant or intermittent?

#### Reasoning Coach

The shortness of breath is mostly intermittent. It seems to get worse when I'm doing physical activities or when I'm going up stairs. It's not really a problem when I'm sitting down or resting, but any kind of exertion seems to make it more noticeable.

Are there any associated symptoms?

### 3

#### **Reasoning Coach**

Yes, now that you mention it, there have been a few other things. I've had some occasional dizziness, especially when I stand up too quickly from sitting or lying down. I've also noticed that I seem to be paler than usual. My partner mentioned it just a few days ago, saying I looked a bit washed out. I haven't had any chest pain or anything like that, but these symptoms together have been worrying me a bit.

Anonymous

Have you noticed any changes in the color of your lips, fingernails, or skin?

#### Reasoning Coach

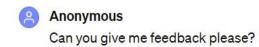
Now that you ask, I've noticed that my lips and fingernails seem a bit paler than usual. I haven't paid much attention to it until now, but thinking about it, they do seem less pink and more on the pale side. As for my skin, it's generally been paler overall, like I mentioned earlier.

Anonymous

Do you have any known heart or lung diseases, such as asthma, COPD, heart failure, or previous episodes of blood clots in the lungs?

#### Reasoning Coach

No, I don't have any known heart or lung diseases. I've never been diagnosed with asthma, COPD, heart failure, or had any issues with blood clots in my lungs. My medical history is pretty uneventful in terms of chronic illnesses or conditions like those.



### Reasoning Coach

Absolutely, I'd be happy to provide feedback on our interaction today.

Firstly, your approach to gathering information was systematic and thorough, which is excellent. You started by asking about the onset and nature of my symptoms, which helped establish a timeline and severity. This is crucial for forming a differential diagnosis. Your questions about associated symptoms were particularly good, as they helped broaden the understanding of my condition beyond the initial complaints. This can be key in identifying less obvious conditions.

You also wisely inquired about my past medical history, medication use, smoking status, recent travel, and surgeries, which are all important factors to consider in a patient's health assessment. Your decision to ask about my age and gender assigned at birth was also pertinent, as these demographics can influence risk factors for certain conditions.

Your suggestion to order a chest x-ray and a comprehensive blood test is a strong move towards a diagnosis. These tests can provide valuable information about a range of conditions that might be causing my symptoms, from infections and anemias to more serious cardiac or pulmonary issues.

For improvement, I would suggest ensuring you also explore lifestyle factors in more depth, such as diet, exercise, and stress levels, as these can significantly impact health and may offer more clues to the underlying cause of my symptoms. Additionally, explaining why you're ordering specific tests could help ease patient anxiety and improve understanding of their care plan.

Overall, you demonstrated a thoughtful and patient-centered approach to care. Keep focusing on building a comprehensive patient history, and always consider explaining the reasoning behind your diagnostic choices to help educate and reassure your patients. Great job!

# Testing and Evaluation of Custom GPTs

- Define Evaluation Criteria
  - >Accuracy, Coherence, Relevance, etc.
- Testing and Evaluation
  - ➤ Validation Data Sets
  - > Human Evaluation
    - Crowdsourcing (e.g., Amazon Mechanical Turk)
    - Expert Review
- Iteration

# Task 3: Draft ideas a personalized GPT

- (1) What tasks could you use ChatGPT for that could be useful.
  - Generating content
  - Synthesizing content
  - Evaluating content
- (2) What GPTs could you design to help your students engage in a specific topic?
  - Not restricted to static question formats
  - New classroom / assignment capability
- (3) Some other interesting use case

# Task 3 Lessons Learned

- Empowerment
- Creativity Unleashed

Flexibility

Collaborative Potential

#### By ChatGPT

GPTs created by the ChatGPT team

### Writing

Enhance your writing with tools for creation, editing, and style refinement



#### DALL-E

Let me turn your imagination into imagery.

By ChatGPT



#### Data Analyst

Drop in any files and I can help analyze and visualize your data.

By ChatGPT



#### Write For Me

Write tailored, engaging content with a focus on quality, relevance and precise word count.

By puzzle.today



#### Humanizer Pro

#1 Humanizer in the market. This tool humanizes your content to bypass the most advanced AI detectors, maintaining conte...

By charlyaisolutions.com



#### Hot Mods

Let's modify your image into something really wild. Upload an image and let's go!

By ChatGPT



#### Creative Writing Coach

I'm eager to read your work and give you feedback to improve your skills.

By ChatGPT



#### Human Writer-Humanizer-Paraphraser (Human GPT) /

I self-score my answers and can instantly refine your prompts, ensuring you receive the most accurate answers. Equipped with...

By CustomGptFactory



#### Al Humanizer Pro

Best Al humanizer to help you get 100% human score. Humanize your Al-generated content to bypass Al detection. Use our...

By bypassgpt.ai



#### Coloring Book Hero

Take any idea and turn it into whimsical coloring book pages.

By ChatGPT



#### Planty

I'm Planty, your fun and friendly plant care assistant! Ask me how to best take care of vour plants.

By ChatGPT



#### **Book Creator Guide**

A writing assistant specializing in book creation, web research, and format recommendations.

By Choi Yong

Research & Analysis



#### Text to Video

Assistant helps you create high-quality video prompts, easily utilizing technology to produce stunning videos.

By Sora

#### Education

Explore new ideas, revisit existing skills



#### math

The worlds most powerful math tool.

By pulsr.co.uk



#### Question Maker

Creates a range of academic questions from PDFs.

By GOURISHETTY



#### Consensus

Your Al Research Assistant, Search 200M academic papers from Consensus, get science-based answers, and draft content...

By consensus.app

Find, evaluate, interpret, and visualize information



#### Scholar GPT

Enhance research with 200M+ resources and built-in critical reading skills. Access Google Scholar, PubMed, JSTOR, Arxiv, an...

By awesomegpts.ai



#### Tutor Me

Your personal Al tutor by Khan Academy! I'm Khanmigo Lite - here to help you with math, science, and humanities questions. I...

By khanacademy.org



#### Universal Primer

The absolute fastest way to learn anything complex. Inspired by the methods of Richard Feynman, I will recursively fill in...

By runway.com

Voxscript



#### Scholar Al

Al Scientist - search and analyze text, figures, and tables from 200M+ research papers and books to generate new...

By scholarai.io



#### Wolfram

Access computation, math, curated knowledge & real-time data from Wolfram|Alpha and Wolfram Language;...

By gpt.wolfram.com



#### Math Solver

Your advanced math solver and Al Tutor. offers step-by-step answers, and helps you learn math and even all subjects, at any...

By studyx.ai

page summarization with vector transcript search -- no logins needed.

Quick YouTube, US equity data, and web

By Allwire Technologies, LLC

#### AskYourPDF Research Assistant Free Chat Unlimited PDFs, Access 400M+

Papers (PubMed, Nature, Arxiv, etc), Analyse PDF (Unlimited PDFs), Generate...

By askyourpdf.com



#### Finance Wizard I predict future stock market prices. Al

analyst. Your trading analysis assistant. Press H to bring up prompt hot key menu.... 31

By titantrades.com





### **Question Maker**

By GOURISHETTY Å

Creates a range of academic questions from PDFs.

**4.3** 

#2

10K+

Ratings (700+)

in Education (EN)

Conversations

#### **Conversation Starters**

Generate a question from this academic text.

Create an open book exam question from this PDF.

Formulate a quiz question based on this document.

Develop an exam question from the uploaded material.

#### Capabilities

Browsing



Based on the document's detailed exploration of clinical reasoning assessment methods, a suitable short-answer or essay question that requires a deeper level of understanding could be:

### **Essay Question:**

"Daniel, Rencic, Durning, et al. (2019) provide a comprehensive overview of clinical reasoning assessment methods, categorizing them into non-workplace-based assessments, assessments in simulated clinical environments, and workplace-based assessments. Reflect on the significance of choosing the appropriate assessment method for evaluating clinical reasoning skills in medical education. In your response, discuss the advantages and disadvantages of employing workplace-based assessments (WBAs) versus non-workplace-based assessments for evaluating clinical reasoning skills. Additionally, propose strategies that medical educators could implement to enhance the reliability and validity of clinical reasoning assessments, taking into consideration the challenges associated with each method."

#### **Guidance for Respondents:**

- •Outline the definitions of workplace-based assessments (WBAs) and non-workplace-based assessments as presented in the document.
- •Compare and contrast the advantages and disadvantages of WBAs and non-workplace-based assessments, particularly in the context of assessing clinical reasoning skills.
- •Suggest practical strategies for medical educators to address the challenges associated with each assessment method, focusing on improving the reliability and validity of the assessments.
- •Reflect on the implications of the choice of assessment method on the development of clinical reasoning competence among medical students and residents.

This question encourages students to critically analyze the document's content, understand the complexities of clinical reasoning assessment in medical education, and propose thoughtful solutions to existing challenges.

# **Future Perspectives**

- Precision education
- Multimodal content generation and simulation
- Al as a member of a healthcare team
- Al Curricula for ethics, governance, data security and privacy
- Democratization of medical education and AI for global health training
- Research acceleration and enhanced multidisciplinary approaches

# Conclusion and Q&A

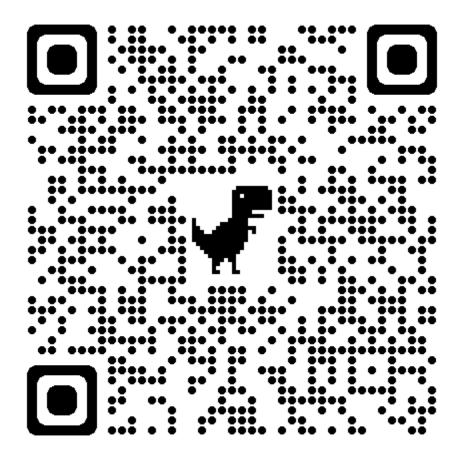
The age of AI has begun and the pace of change in medical education and practice will be steep.

- What are the new skills physicians need?
- What does expertise look like in the age of AI?
- How do we prepare for the challenges?
- Where do we begin?

# Thank you!

Please provide us with feedback >

Please share the materials with interested colleagues. We will be updating the materials over the next few months as we continue to give the workshop.



Email: <u>CRunyon@nbme.org</u>