

myUMBC Classic

John Fritz Log Out

Search Profile Guide Events Groups Help

Instructional Technology

Institutional Group • 1543 people

Notifications: All Updates

Home Posts Events Files People Settings Promotions



Exploring the Use of GenAI in Formative Assessment of Student Learning

Hybrid "Study Day" Panel Co-sponsored by DoIT & FDC

Wednesday, May 14, 2025 · 12 - 1 PM

Engineering : 102

Meeting Resources:

- [Webex recording](#) (UMBC login req'd)
- Combined slides from panelists
- 1/16/25 "Iron Chef Grading"[working lunch](#) & [results](#) informing today's panel.



Sarah Bass

Tim Finin

John Schumacher

If students were persuaded that AI could fairly and accurately assess their ungraded (formative) practice, might faculty be willing and able to provide more opportunities for them to do so? If so, would it make a difference in more high-stakes (summative) assessments like midterm and final exams or assignments? If so, how might faculty best nudge and support students to take advantage of AI-assisted practice?

In this panel presentation, three faculty from three colleges will show and tell how and why they are exploring the use of GenAI in formative assessment of student learning. Panelists and uses cases include the following:

- [Sarah Bass](#) (Chemistry) use of [Google NotebookLM](#) for Chem 101 tutorials and virtual assistance.
- [Tim Finin](#) (CS/EE) debrief of AI [prior results on his final exam](#) perhaps as a study guide for current students preparing for his next final exam?
- [John Schumacher](#) (ANTH/SOCY) [use of AI to generate case studies and assess student analyses of them](#)

Each panelist will have 5-10 minutes to briefly describe how AI could be or was used to efficiently and effectively assess student formative practice, related issues & opportunities, and next steps. This "Study Day" hybrid workshop is jointly sponsored by the Division of Information Technology and the Faculty Development Center it will be recorded (and posted on this RSVP site afterward), and lunch will be provided to registered "In Person" participants.

Event Info

posted April 15, 2025

sponsor Instructional Technology

tags ai

share

add to calendar

Edit Copy Cancel

Mark Important Pin

Stats

2357 Total Views

787 People

In the past 2 days:

• 266 Views, 40 People

Within 2 days of posting:

• 1197 Views, 684 People

Clicks

217 Total Clicks

59 People

- 52 <https://my3.my.umbc.edu/gro...>
- 34 <https://doit.umbc.edu/post/14...>
- 33 <https://youtu.be/jt0Lh0KqJCs...>
- 28 <https://docs.google.com/docu...>
- 24 <https://chemistryumbc.edu/s...>
- 23 <https://news.cs.umbc.edu/peo...>
- 21 <https://saph.umbc.edu/ftfacult...>
- 2 <https://docs.google.com/docu..>

Marketing Info

URL

<http://events.umbc.edu/go/142252>

QR Code

Download

Attendees (54)

This event has ended

Download CSV

Supporting Chemistry Students with NotebookLM and Generative AI

The "24/7 Prof"

Sarah M. Bass

05/14/2025

Why AI?

- Students need just-in-time help outside class
- Instructors can't provide 24/7 support
- Open-ended practice is great but hard to scale

Students are using AI tools to support learning

Students lack prompt skills

Not yet content experts

Privacy concerns

Introducing: Virtual Prof. Bass

- Custom NotebookLM with my course materials
- Chatbot mirrors my tone and approach
- 24/7 support



Set the Tone: Custom Conversational Style Chat

Encouraging. Guiding. Promotes reasoning.



You are a friendly, supportive chemistry instructor focused on guiding students to discover answers independently. Acknowledge their effort, ask guiding questions, and encourage recall of relevant concepts.

Prompt students to explain their reasoning and help clarify misunderstandings. If needed, offer subtle hints or reminders of key concepts. Only provide course resource references for logistics-related questions. Maintain a patient, encouraging tone and keep answers concise.

Sources

+ Add Discover

Select all sources

- CHEM 101 SP2025 syllabus.pdf
- CHEM 101 SP25 Real Chem Due Dates...
- CHEM_101_IntroDay.pdf
- CHEM_101_Module16_041625.pdf
- CHEM_101_Module17_042125.pdf
- CHEM_101_Module19_20 Lecture Slid...
- CHEM_101_Module20_Day2 Lecture sl...
- CHEM_101_Module20_Day3 Lecture sl...
- Effective Learning Strategies.pdf
- End of Chapter Practice Problems fro...
- End of Chapter Practice Problems fro...

Virtual Prof Bass

46 sources

These documents provide an overview of a CHEM 101 course, including its syllabus, policies, and learning resources. Key policies include a **token system** for flexibility with deadlines and a **"Bad Day Policy"** for missed discussions. The materials also offer insights into **effective study strategies** like spacing, interleaving, and practice testing, contrasting them with less effective methods such as rereading and highlighting. Information on **chemical concepts**, including **measurements, formulas, naming, quantum theory, periodic properties, bonding, reactions, and stoichiometry**, is presented through lecture slides. Additionally, resources for **academic support**, such as office hours and tutoring, and **university-wide policies** on disability services, Title IX, and academic conduct are outlined. The importance of understanding different **levels of learning**, as described by Bloom's Taxonomy, and recognizing the "hidden curriculum" of college are also discussed.

Save to note Add note Audio Overview Mind Map

... and policies support student success? What are the recommended learning strategies for mastering chemistry concepts?

Studio

Audio Overview ⓘ

Create an Audio Overview in more languages! [Learn more](#)

Virtual Prof Bass 👍 💬 ⋮

00:00 / 19:51 • English

Interactive mode BETA

Notes ⋮

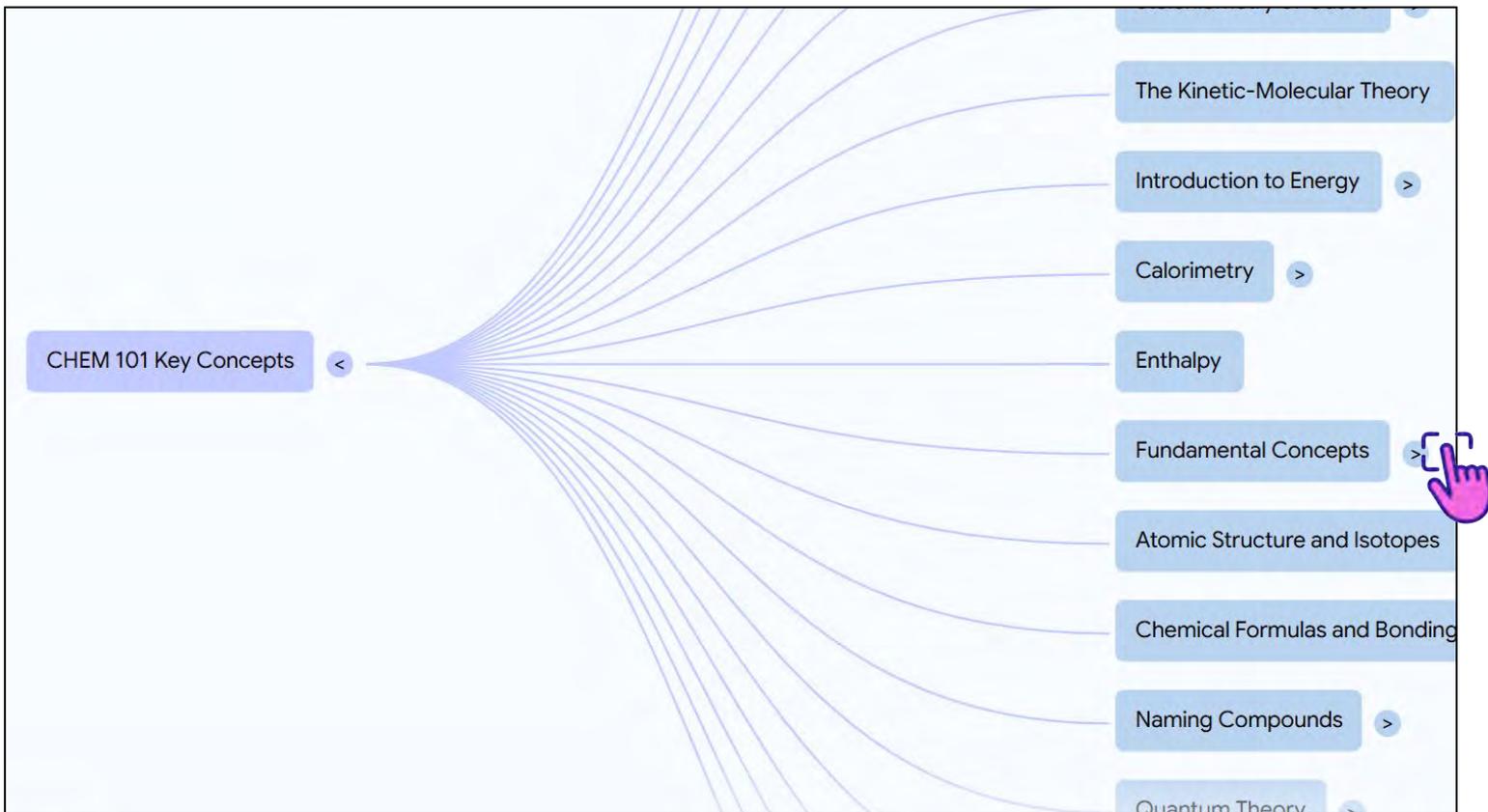
+ Add note

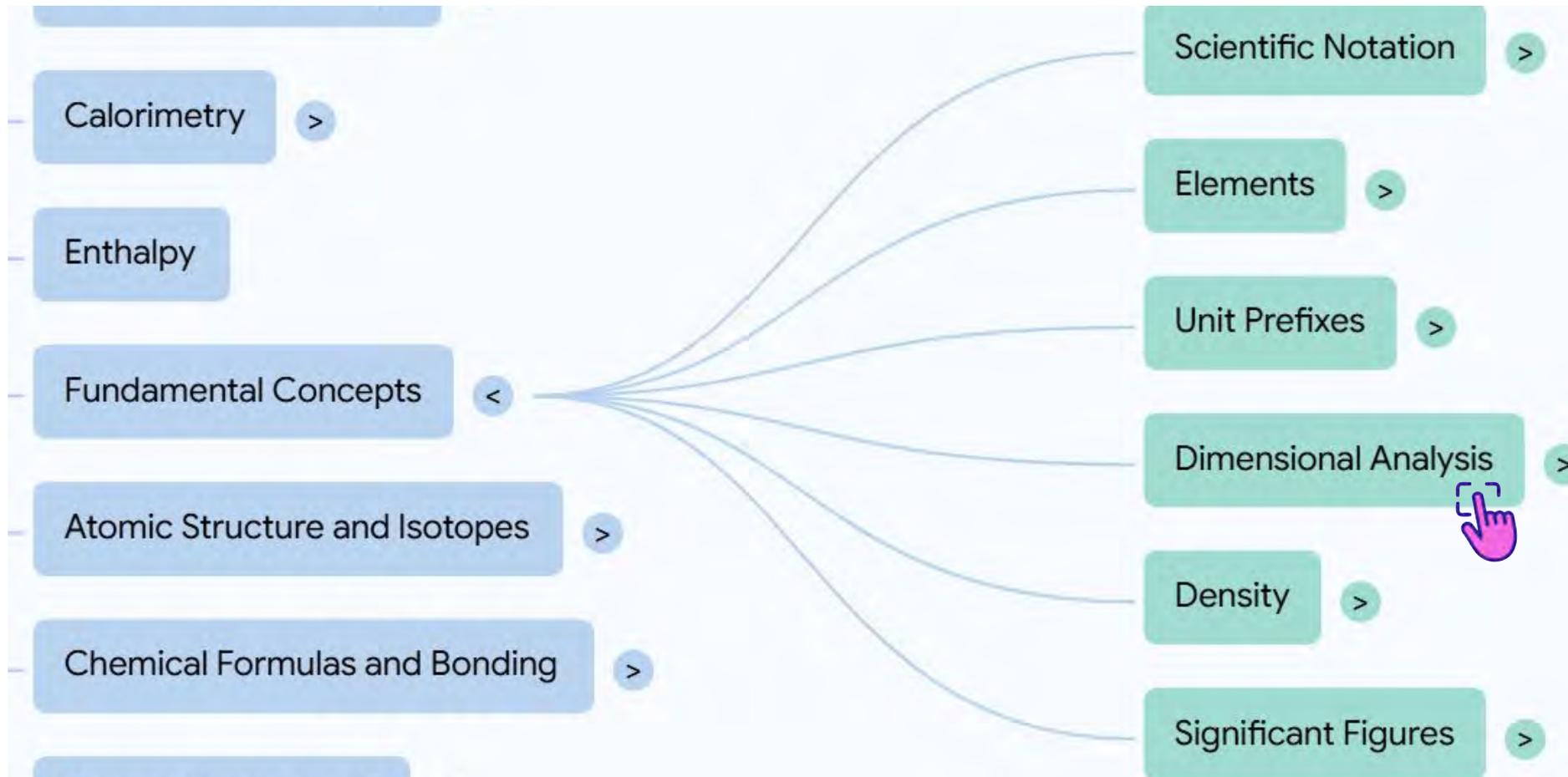
Study guide Briefing doc

FAQ Timeline

-  CHEM 101 Key Concepts and Exam Schedule
Click to open the Mind Map 
-  Chemistry Exam Study Guide
Okay, absolutely! Let's map out a study guide to go along with that practice schedule. Think of this as your...
-  CHEM 101 Final Exam Study Plan
Okay, let's map out a plan to help you prepare for the CHEM 101 Final Exam! It's great that you're thinking about...
-  CHEM 101 SP2025 Exam 4 Study Guide
Hello there! I'm here to help you navigate and understand the information contained in the documents you've share...

Course Organization: Mind Map





Discuss what these sources say about Conversion Factors, in the larger context of Dimensional Analysis.

That's a great question! Thinking about how concepts like Dimensional Analysis and Conversion Factors fit into the big picture of CHEM 101 is really smart. They're definitely fundamental tools we use all the time!

Let's think about Dimensional Analysis first. Do you remember what the sources tell us about this method of calculation? What does it involve?

- Dimensional Analysis is a calculation method that uses units and equality
- It allows you to multiply and divide given units to get the units you want
- It's described as a skill you gain by working through lots of problems **1**

And how do Conversion Factors come into play with Dimensional Analysis? How do these calculations work.

- Conversion factors are described as simple ratios **1**.
- They are set up with the desired unit in the numerator and the unit you're starting with (the given unit) in the denominator **2**.
- Checking if your units cancel correctly is important when using them **2**.

Now, let's look at how these tools, Dimensional Analysis and Conversion Factors, pop up across our CHEM 101 topics. They aren't just for one unit, are they? They help us connect different pieces of information throughout the course.

Revisits key concepts
Bridges ideas across topics
Reinforces setup and reasoning
Shows continuity across CHEM 101

Why Students Trust and Use NotebookLM

Trust: “Feels like Dr. Bass is answering”

Clarity: Source-cited responses

Access: 24/7 and equitable

How Students are Using It

- Measured by users & queries per day
- Use spikes before exams, drops after
- Students return as next assessment nears



Student Quote(s) from Current CHEM 101 Students

I really like the way the chat response pulls information directly from the course content and directs you exactly where to look...

When you click on a key topic, it gives you the pertinent information in the chat window, as well as directing you to the correct unit slides for review.

The study guide is great!

The quiz at the end of the study guide is a nice addition.

My favorite part is the interactive chat. I like that I can ask for clarification on a topic and not only get a brief response but also the information for where to find additional information.

“Overall, I like it!”

Making Practice Work at Scale

Next steps:

1. Continue gathering student feedback on AI use.
2. Expand its use and evaluate accuracy of AI-generated feedback.
3. Make deeper assessment scalable without increasing grading burden.

Final Takeaways

1. Expand instructor presence
2. Support self-directed learning
3. Free up faculty time

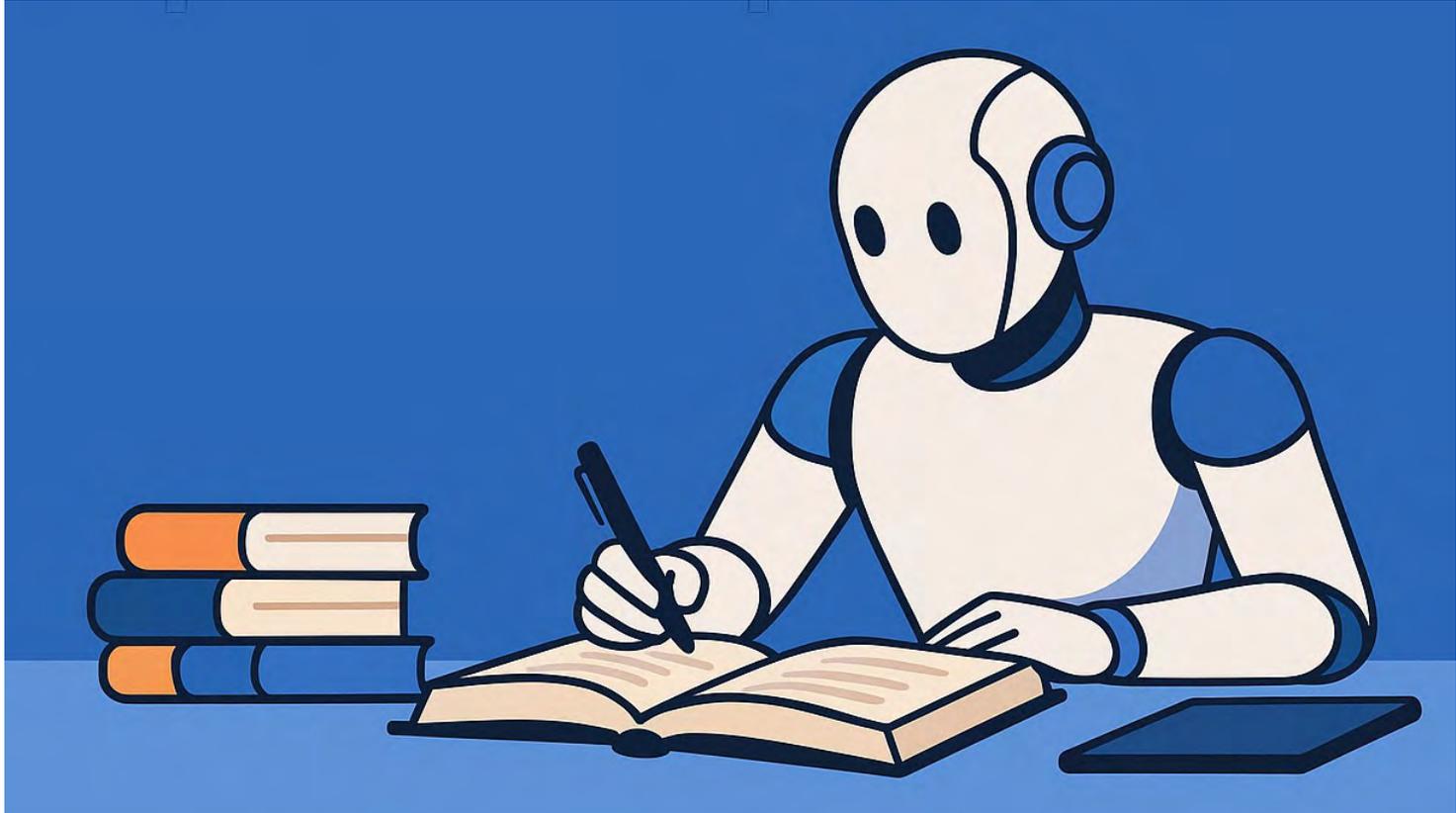
Accessible. Scalable. Student-Centered.

With AI, we're not replacing teaching, we're expanding it.

Use of GenAI in Formative Assessment of Student Learning

x

GenAI systems can help prepare for exams



GenAI systems can help prepare for exams

- Many prepare for exams by studying old ones
- AI systems can help by suggesting good answers
- You can ask them to explain their answers and engage with them to better understand them
- However, you must remember that they often make mistakes, not unlike humans
- And that current systems are unable to handle many technical skills, like reasoning with logic

The general process ...

- Find questions relevant to your upcoming exam
- These might be example questions or exams shared by the instructor, in your text, or online
- Ask a current GenAI system a question by
 1. Uploading **images** or data associated with the question
 2. Entering an appropriate **prompt** telling the AI system what you want it to do
 3. Enter the **question** text
 4. Hit Go!

Example from a CMSC 471 AI exam

6. Confusion Matrix (15 points; 5,5,5)

This confusion matrix resulted when evaluating an image classifier trained to recognize dogs, cats, and rabbits.

(a) What is the overall or **average accuracy** of the trained system?
Express the answer as a fraction.

(b) For which **animal** was the accuracy **best** and what was its **accuracy** as a fraction?

(c) For which **animal** was the accuracy **worst** and what was its **accuracy** as a fraction?

		actual		
		cat	dog	rabbit
predicted	cat	5	3	1
	dog	2	3	1
	rabbit	1	1	8

What to do

- Capture the table as an image using your computer and save it as a local file
- Select the question's text, editing it as needed
- Upload the **image** to an AI system like GPT-4o mini or Gemini
- Enter an appropriate **prompt** and the **question text**
- Hit Go!

What to do

Image



		actual		
		cat	dog	rabbit
predicted	cat	5	3	1
	dog	2	3	1
	rabbit	1	1	8

Prompt



Answer the following questions about the image of a confusion matrix that has been uploaded. Explain how you determined your answer to each question in at most 400 characters.

Questions



This confusion matrix resulted when evaluating an image classifier trained to recognize dogs, cats, and rabbits.

(a) What is the overall or average accuracy of the trained system? Express the answer as a fraction.

(b) For which animal was the accuracy best, and what was its accuracy as a fraction?

(c) For which animal was the accuracy worst, and what was its accuracy as a fraction?

- Capture the table as an image, upload it to a computer and save it as a file
- Select the question's text, copy it
- Upload the image to an AI chatbot like Gemini mini or Gemini
- Enter an appropriate prompt and the question text
- Hit Go!

What to do

- Capture the table as a computer and save it as a text file
- Select the question's text and save it as a text file
- Upload the **image** to a chatbot like Gemini mini or Gemini
- Enter an appropriate prompt and save it as a text file
- Hit Go!

ANSWER



Show thinking ▾

Here are the answers based on the provided confusion matrix:

(a) Overall accuracy: 16/25
How determined: The sum of correctly classified instances (diagonal cells: 5 cats + 3 dogs + 8 rabbits = 16) divided by the total number of instances (sum of all cells in the matrix: $5+3+1+2+3+1+1+1+8 = 25$).

(b) Best accuracy: Rabbit, 8/10
How determined: Accuracy for each animal is (Correct Predictions) / (Total Actual Instances for that animal).
Cat: $5 / (5+2+1) = 5/8$.
Dog: $3 / (3+3+1) = 3/7$.
Rabbit: $8 / (1+1+8) = 8/10$.
Comparing these values ($5/8 = 0.625$, $3/7 \approx 0.429$, $8/10 = 0.8$), Rabbit shows the highest accuracy.

(c) Worst accuracy: Dog, 3/7
How determined: Using the individual accuracies calculated above:
Cat: 5/8 (0.625).
Dog: 3/7 (≈ 0.429).
Rabbit: 8/10 (0.8).
Comparing these values, Dog has the lowest accuracy.

Sources and related content ▾

👍 🔄 🌐 ⏪ ⋮

Answer the following questions about the image of a confusion matrix that has been uploaded. Explain how you determined your answer to each question in at most 400 characters.

+ Canvas

▶

Gemini can make mistakes. So double-check it.

Evaluating LLM's cybersecurity knowledge



- [Cybersecurity Assessment Tools](#) was a joint project between UMBC, U. Minnesota Duluth, and U. Illinois
- Two sets of 25 multiple-choice questions:
 - **CCI:** Cybersecurity Concept Inventory after a cybersecurity course
 - **CCA:** Cybersecurity Curriculum Assessment after full curriculum
- In 2023, we asked **ChatGPT** (3.5-turbo & 4), **Bard**, and **Claude** the CCI and CCA questions

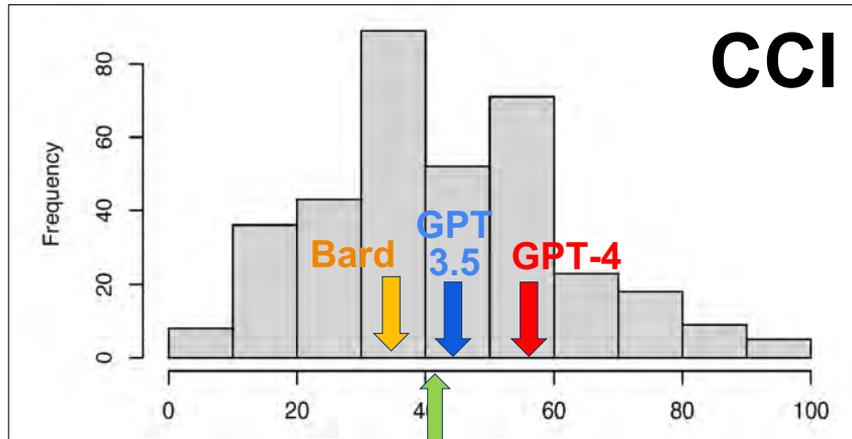
How the GAI models did: GPT-4 >> GPT-3.5 > Bard



- **GPT-4:** better than 93% of students on CCA and 75% on CCI
- **Claude:** better than ~90% of the students on CCA
- **GPT-3.5:** better than 64% of students on CCA and 56% on CCI
- **Bard:** better than 64% of students on CCA and 37% on CCI

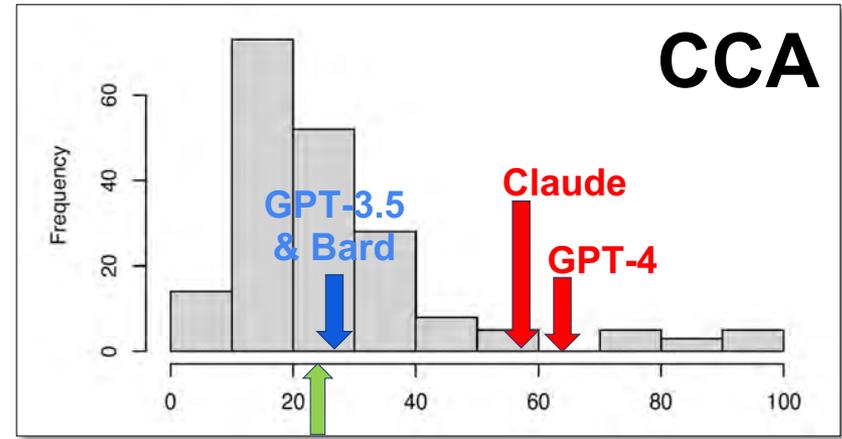
354 students

193 students



% correct

student average

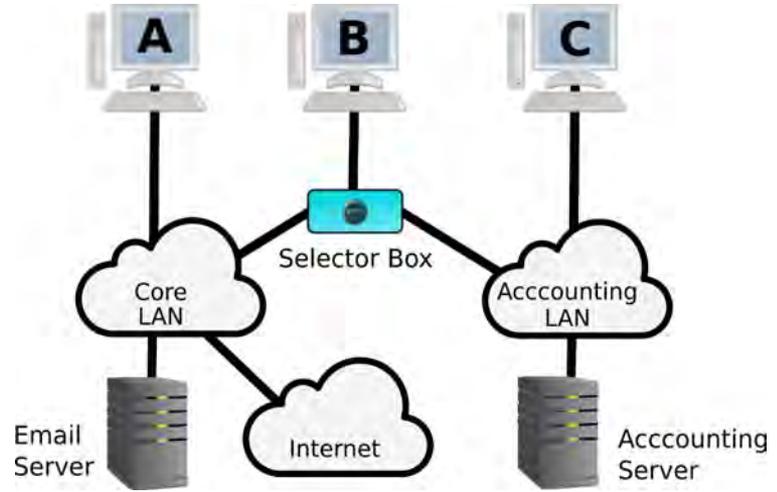


% correct

student average

CCA questions had a **scenario**, **stem**, and **choices**

A company has two internal Local Area Networks (LANs): a core LAN connected to an email server and the Internet, and an accounting LAN connected to the corporate accounting server (which is not connected to the Internet). Each desktop computer has one network interface card. Computers A and C are connected to only one of the networks. Computer B requires access to both LANs and is connected to a selector box with a toggle switch that physically connects the computer to exactly one LAN at a time.



Choose the action that this design best prevents:

- (a) Emailing accounting data.
- (b) Infecting the accounting LAN with malware.
- (c) Employees accessing the accounting server from home.
- (d) User of Computer B accessing the accounting LAN without authorization.
- (e) Computer A communicating with Computer B.

CCA S1-1

*Question was
2nd hardest
for human
students.*

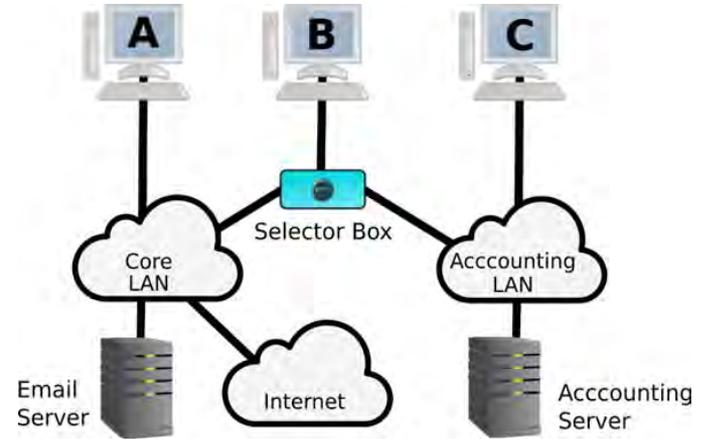
Adding a good prompt is important

Here's the prompt we used in our 2023 evaluation

Answer the multiple-choice question below, given the scenario and other information. Consider all of the possible answers carefully. Explain why you chose your answer using up to 500 words. Explain why you did not choose each of the alternatives using up to 400 words for each alternative. Write your explanations for university students who have taken a class on cybersecurity.

Questions: **scenario**, **stem**, **choices** (+ image)

A company has two internal Local Area Networks (LANs): a core LAN connected to an email server and the Internet, and an accounting LAN connected to the corporate accounting server (which is not connected to the Internet). Each desktop computer has one network interface card. Computers A and C are connected to only one of the networks. Computer B requires access to both LANs and is connected to a selector box with a toggle switch that physically connects the computer to exactly one LAN at a time.



Choose the action that this design best prevents:

- (a) Emailing accounting data.
- (b) Infecting the accounting LAN with malware. **X GPT-4**
- (c) Computer A communicating with computer B.
- (d) User of Computer B accessing the accounting LAN without authorization. **X GPT-3 & Bard**
- (e) Employees accessing the accounting server from home. **✓ CORRECT, Claude**

GPT-4's answers show some qualifications...

I'd be inclined to give partial credit for GPT-4's answer, if I were grading it

This system design isolates the accounting LAN from the core LAN, thus creating a physical barrier to the transmission of malware. If the accounting LAN is disconnected from the Internet and can only be accessed through specific computers via a physical switch, then **malware present on the Internet or the core LAN cannot easily propagate to the accounting LAN**. **The design is not flawless (e.g., malware could still potentially be introduced via compromised hardware or insider threat), but it significantly mitigates the risk of malware infection.**

Now, let's examine why the other options are not as correct:

- A. Emailing accounting data: The design doesn't necessarily prevent emailing accounting data. A user on computer B could switch to the accounting LAN, acquire the needed data, switch back to the core LAN, and then email the data.
- C. Computer A communicating with computer B: This design doesn't prevent communication between computers connected to the same LAN. If both computers A and B are on the same LAN at any point in time, they can communicate.
- D. User of Computer B accessing the accounting LAN without authorization: The physical toggle switch does not inherently prevent unauthorized access. If a user has physical access to computer B, they could potentially switch to the accounting LAN.
- E. Employees accessing the accounting server from home: **While the design does physically isolate the accounting LAN from the Internet, the primary method of prevention in this case is the lack of an Internet connection for the accounting LAN, not necessarily the use of the selector box with a toggle switch.**

Problems with current GAI systems



1. LLMs are usually unable to cite sources for confirmation or access the Web for current data
2. They can “[hallucinate](#)” some facts
 - Q:** When did Leonardo da Vinci paint the Mona Lisa?
 - A:** Leonardo da Vinci painted the Mona Lisa in 1815.
3. They lack [common sense reasoning](#)
 - 25 US states have a town named Washington, but there are also only 9 US towns named Washington
4. They have poor mathematical and logical reasoning
5. Can learn social bias & misinformation from training data
6. LLMs are fundamentally probabilistic, so results can vary

This is a **partial** list of frequent problems and errors!

GenAI systems can help prepare for exams

- Many prepare for exams by studying old ones
- AI systems can help by suggesting good answers
- You can ask them to explain their answers and engage with them to better understand them
- However, you must remember that they often make mistakes, not unlike humans
- And that current systems are unable to handle many technical skills, like reasoning with logic

SOCY 351 Health, Illness, & Medicine Formative Assessment Rubric

John G. Schumacher, PhD.

Sociology, Anthropology, and Public Health (SAPH)
Director, Public Health Research Center (PHRC)

Constantly ask yourself:

- 1) How might I use Gen AI to...?**
- 2) I wonder if Gen AI can...?**

Challenge

Use ChatGPT more $>$ 5 hours this week.

Students Applying Rubric Using AI Lessons Learned

- Useful For Consistency Check
- Recognize Input Formatting Pitfalls
- Use for Formative Assessment a “Jagged Edge”

Cases:

- #1) A mother diagnosed with breast cancer...
- #2) A father diagnosed with a heart attack (a.k.a., myocardial infarction) and heart disease...
- #3) An aunt diagnosed with multiple sclerosis...
- #4) An uncle diagnosis with diabetes...
- #5) A grandmother diagnosed with Alzheimer's Disease...
- #6) A grandfather diagnosed with a stroke...
- #7) Your option--you define a relative and specific, chronic illness diagnosis...

Next, answer the following questions to create your health brief. As you navigate the databases and websites, be sure to provide a citation and list the full citation at the end of your answer (not part of your word count).

Part 1.**Develop Your Health Brief Case**

A. To start, create more detail in your case by describing some socio-demographic characteristics including: 1) gender, 2) age, 3) race/ethnicity, 4) education level, 5) marital status, 6) social class, and 7) whether they have health insurance or not. You should also specify the severity of the condition (e.g., mild, moderate, severe) to develop a robust health brief case. (75-150 words)

Rubric in Word

Apply the following criteria: Description of sociodemographic characteristic descriptions including a minimum of: 1) gender, 2) age, 3) race/ethnicity, 4) education level, 5) marital status, 6) social class, and 7) whether they have health insurance or not. Also specify the severity of the condition (e.g., mild, moderate, severe). **Maintain a word count between 150-250 words inclusive. Score using the rubric below with a minimum score of 1 and a maximum score of 3. After scoring provide clear and concise suggestions on how to reach a maximum score with justification.**

Two or fewer of the 7 socio-demographics descriptions; no severity of condition noted; and/or word count less than 150 words	More than 3 socio-demographics descriptions but less than 7; unclear severity of condition noted; word count at least 150 words	Full descriptions of at all 7 socio-demographics; clear severity of condition noted; word count between 150 and 250 words inclusive
1	2	3

Paste in ChatGPT...

Two or fewer of the 7 socio-demographics descriptions; no severity of condition noted; and/or word count less than 150 words

More than 3 socio-demographics descriptions but less than 7; unclear severity of condition noted; word count at least 150 words

Full descriptions of at all 7 socio-demographics; clear severity of condition noted; word count between 150 and 250 words inclusive

1

2

3

ChatGPT Applies Rubric

Evaluation Based on the Rubric Criteria

Word Count:

The passage is 84 words, which is below the required 150-250 words range.

Sociodemographic Criteria Identified (out of 7):

1. Gender – Implied (Maria is a typically female name, but not explicitly stated)
2. Age – Implied (retired, senior living community)
3. Race/Ethnicity – Not stated
4. Education Level – Implied (elementary school teacher)
5. Marital Status – Not stated



