

Comparing Answers from Various AI Bots: Bing, Gemini, ChatGPT, oh my!

Why?	
In this lesson, students will compare and analyze responses from various generative AI models based on criteria such as accuracy, relevance, completeness, clarity, and creativity. Through hands-on evaluation and critical reflection, they will gain insights into the strengths and limitations of AI technologies, fostering their analytical and digital literacy skills. This exercise is designed to prepare students for a future where AI plays an integral role, encouraging them to think critically about the technology's implications in society and their personal lives.	
Materials Needed	Time needed
Computers with internet access Projector hooked up to a computer with internet access	Approximately 30 - 45 minutes
Objectives	
Students will be able to critically evaluate the performance of various generative AI models based on predefined criteria . Students will explore the strengths and limitations of different generative AI models. Students will develop critical thinking skills by evaluating and comparing AI-generated text. Students will be able to articulate the strengths and limitations of AI-generated content through structured comparison and analysis.	
Key Concepts & Vocabulary	
AI-Generated Text: Information created by artificial intelligence, often imitating human writing. Prompt: A starting point or question given to the AI to guide its text generation. LLM (Large Language Model): A powerful AI model trained on massive amounts of text data, capable of generating human-quality text, but with limitations in understanding and accuracy.	
Lesson	
Introduction Introduce the concept of AI-generated text and its growing role in various fields. Discuss responsible AI use and ethical considerations. Students should never enter personal information in a chat bot. Demonstration On a projector, show students each of the chatbots you want them to use for this	

experience. Demonstrate how to enter text, and look at the answer(s) given by the model.

A list of links is provided in the Sources section at the bottom of this lesson plan. Provide a list of links to students so they can access each page. In some cases, AI chatbots may be blocked by filtering software, so check ahead of time to make sure students will be able to access models.

Student Testing

Either in pairs or individually, students type in questions relevant to a particular topic into the chat for different AI chatbots. They should type the same question to each chatbot.

Students would be well advised to keep all tabs open at the same time to better compare answers side-by-side.

Choose 3-4 different models for students to access, or let them choose from among the options.

Comparison Criteria

Display this list of criteria on the screen for students to look at as they compare the answers from different AI models. Perhaps suggest having students give each model a rating of 1-5 or 1-10 for each criterion:

Accuracy: Evaluate the correctness of the information provided by the AI models. This involves checking facts, figures, and the overall truthfulness of the answers.

Relevance: Assess how relevant the answers are to the questions asked. The responses should directly address the query without providing unnecessary or unrelated information.

Completeness: Determine if the answers provide a comprehensive response to the questions. The responses should cover all aspects of the query without leaving out critical information.

Clarity and Coherence: Evaluate how clear and understandable the answers are. The responses should be well-structured, logically organized, and free of confusing jargon or complex language that isn't explained.

Creativity and Novelty: For more open-ended questions, evaluate the creativity and novelty of the answers. This involves looking at how the AI models can generate unique, insightful, or innovative responses.

Ask students to write down comments explaining why they felt the way they did about their ratings.

Discussion Questions

Which AI model did you find to be the most accurate in its responses? Can you share an example where an AI model provided an exceptionally accurate or inaccurate answer?

Were there instances where AI models provided information that was not directly relevant to the question asked? How did this affect your perception of the model's

usefulness?

Did any of the AI models consistently provide more comprehensive answers than others? How important do you think completeness is in the context of AI-generated responses?

Which AI model produced the clearest and most coherent responses? Can you provide examples of responses that were particularly well-structured or, conversely, confusing? Can you discuss a response that stood out to you as being particularly creative or novel? How do you think AI models generate such unique responses?

Based on the criteria of accuracy, relevance, completeness, clarity, and creativity, which AI model do you think performed the best overall? Why?

Which models were able to search the internet? Do you find this to be an important factor in which one to use?

Did you encounter any responses that raised ethical concerns, such as bias or misinformation? How should developers address these issues in AI models?

Based on your experience, what improvements or advancements would you like to see in future generations of AI models?

Has this exercise changed your perspective on the reliability and utility of AI-generated content? Would you trust an AI model for critical tasks, and why or why not?

Supplemental Activity Ideas

Comparing Google Gemini Drafts: Access the Pedagog.ai lesson plan on [comparing Google Gemini drafts](#)

AI Impact Journal: Have students maintain a journal over a week documenting instances where AI impacts their daily life, whether through social media algorithms, search engines, online shopping recommendations, or virtual assistants. This can heighten awareness of AI's pervasive role in modern life.

AI Bias Investigation: Assign students to research and present on instances where AI has shown bias in real-world applications, such as facial recognition or sentencing algorithms. This can lead to discussions on data quality, representation, and the importance of ethical AI development.

AI and Future Careers: Facilitate a workshop or panel discussion on the future of work in an AI-driven world. Discuss which careers might be impacted by AI and how students can prepare themselves for a future where AI tools are ubiquitous.

Sources - AI Models

OpenAI's ChatGPT - <https://chat.openai.com/>

Google's Gemini - <https://gemini.google.com/app>

Microsoft's Bing Chat / Copilot - <https://www.bing.com/chat>

Anthropic's Claude - <https://claude.ai/chats>

Perplexity AI - <https://www.perplexity.ai/>

Pi - <https://pi.ai/discover>