

Limitations Of

Artificial Intelligence

Evaluation Deck

Created by  AppHaus

 SAP

INTRODUCTION

Artificial Intelligence (AI), while revolutionary, is not without its constraints. Like any tool, it has its strengths and limitations, shaped by the data it's trained on and the algorithms that drive it.

Understanding these boundaries is crucial to harnessing its full potential and ensuring its responsible application in the business landscape.

Within this card set, you'll find the key limitations of AI.

As you explore them, reflect upon the use case ideas you've brainstormed earlier. **How might these limitations impact your envisioned scenarios? Are there creative workarounds or adjustments that can be made to navigate these challenges?**

Engage critically and constructively. By recognizing and addressing these limitations, you'll be better equipped to integrate AI into your business strategies effectively and ethically.

LIMITATION

**Ensure that the
model always
behaves ethically**

LIMITATION

Ensure that the model always behaves ethically

AI models learn from the data they are trained with. In the case of Generative AI, models like GPT learn from vast online data, inheriting both its knowledge and biases. These models **lack a moral compass to discern right from wrong, and have no awareness of cultural nuances**, so they might produce biased or inappropriate content.

Ensuring consistent ethical behavior requires continuous monitoring and fine-tuning, as AI can't inherently understand or uphold human ethics.

DISCUSSION QUESTIONS

Discuss **if this limitation could affect your use case and how**. The following questions might help.

Is the scenario one where biased or prejudiced outputs could lead to discrimination or harm to certain groups?

Will the AI be making recommendations that could have long-term consequences for individuals, such as in career, education, or health?

Is there a potential for the solution to produce outputs that might be considered controversial or offensive in certain contexts or cultures?

LIMITATION

**Ensure a
consistent and
predictable output**

LIMITATION

Ensure a consistent and predictable output

Generative AI models like GPT, due to their complex neural networks and vast amounts training data, **may produce varying outputs at different times even with similar inputs.** This variability poses a challenge in situations requiring consistent, repeatable results.

Although strict prompt engineering strategies can reduce these inconsistencies, it's not possible to eliminate them entirely.

DISCUSSION QUESTIONS

Discuss **if this limitation could affect your use case and how.** The following questions might help.

In your scenario, can unpredictable outputs pose a risk to safety, security, or critical decision-making?

Is consistency in AI-generated outputs critical for legal or compliance reasons?

Do subsequent processes or decisions rely heavily on consistent outputs?

Can inconsistent outputs significantly impact user experience or customer satisfaction?

LIMITATION

Produce convincing yet false answers

LIMITATION

Produce convincing yet false answers

Generative AI models prioritize coherent and contextually relevant responses over absolute truth. Therefore, they can generate answers that sound plausible but are entirely inaccurate (“hallucinations”). **They don’t verify facts but rather rely on patterns from their training.**

This means they can inadvertently spread misinformation or misconceptions, making it crucial for users to cross-check important information from trusted sources.

DISCUSSION QUESTIONS

Discuss **if this limitation could affect your use case and how**. The following questions might help.

Are the stakes high if the AI outputs incorrect information, such as in medical, research, financial or legal sectors?

Are there verification or validation processes (moderation) in place to catch any “hallucinated” outputs?

Is the AI’s generated content used as primary data or evidence for decision-making within the company?

LIMITATION

Provide explanations or reasoning

LIMITATION

Provide explanations or reasoning

While generative AI models can produce answers that fit a given context, it's hard to pinpoint the reasoning behind their outputs. They operate based on patterns learned from extensive and varied datasets, not through a clear, step-by-step reasoning process. During training, they adjust an immense number of internal parameters, defining complex relationships between inputs and outputs.

This complexity means it's often not possible to trace the causal connections between input and generated output.

DISCUSSION QUESTIONS

Discuss **if this limitation could affect your use case and how**. The following questions might help.

Will the AI's output influence critical decision-making processes where understanding the 'why' is as important as the 'what'?

Is there a risk of significant financial, legal, or operational consequences if the reasoning behind an AI's output can't be justified?

Would your target audience or users be more trusting or accepting of AI outputs if they came with explanations?

LIMITATION

Learn from new data on-the-fly

LIMITATION

Learn from new data on-the-fly

Generative AI models need extensive data and computational power for training. While they can produce real-time responses, **updating their foundational knowledge to adapt to new information on-the-fly is currently impractical due to time and resource constraints.**

This limitation restricts their ability to immediately learn from new or updated data.

DISCUSSION QUESTIONS

Discuss **if this limitation could affect your use case and how.** The following questions might help.

Does your use case require the AI model to adapt quickly to new data or situations as they emerge?

Will the AI be interacting with users or systems that constantly provide new data that should be integrated immediately for enhanced performance?

LIMITATION

Address specific, well-defined problems

LIMITATION

Address specific, well-defined problems

Unlike Large Language Models (LLMs), **simple Machine Learning (ML) models can only address specific, predefined problems based on their training data** and struggle with new or unexpected situations.

For example, a model trained to recognize cats in images won't identify dogs without dog-specific data. Similarly, an AI trained to play chess won't excel at Go without retraining on Go data.

DISCUSSION QUESTIONS

Discuss **if this limitation could affect your use case and how**. The following questions might help.

Do you have plenty and high quality data to train your model?

Is the task of the model changing or staying consistent?

Is your problem domain clearly defined with a specific set of criteria?