AI: What to Know

Part 1: The Basics

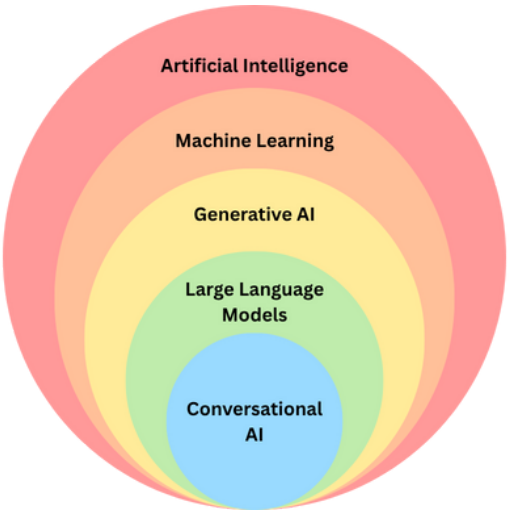
What is AI?

Artificial Intelligence is the process of using computers to perform tasks that typically would require human intelligence.

Examples:

- Image recognition
- Streaming service recommendations
- Digital assistants
- Writing

AI is a broad space and there are many different subtypes of AI that exist. When people talk about AI now they often mean **Generative AI** or **Large Language Models (LLMs**) which are subsets of AI. Some common LLMs include Chat-GPT, Grok, Claude and Gemini.



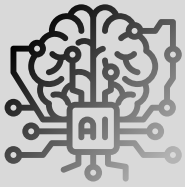
A hierarchy of AI types. Definitions can be found in the glossary on the reverse side.

When to use AI?

AI is a great tool when you are creating something new or want to reason about data. Generative AI models can help on tasks like outlining a presentation or adding images to a presentation.

Models can be custom trained on your data or data can be shared with more general LLMs to reason about the data. For example, if you had student scores on every question for a test you might consider using a LLM to help you identify which topics students struggled with the most and use that to structure a review session.

It's often not practical to train a custom model given the time required and amount of data needed. However, it's important to understand what data it's okay to share with an existing model. Personally identifiable information like users' names should never be shared. If you're unsure about a use case, it's best to consult with your company's IT team.



AI: What to Know

Part 1: The Basics

Glossary

Model: A program that has been trained on data to solve a particular task. The task can be very specific (e.g. Identify students that might be at risk of missing too many school days) or more general (e.g. process a user question and provide a written answer).

Generative AI: An AI model that generates new content (e.g. text, images, audio, or video).

Large Language Models (LLMs): AI systems that have been trained to understand and generate human language. These are typically trained on very large and varied text datasets like webpages available on the internet.

Artificial General Intelligence: A system that can replicate human performance on any task not just those it's been trained on. These are still theoretical and this is what people mean when they talk about AI sentience.

Conversational AI: A system that allows for back and forth conversation with a user (e.g. chat bot). The main challenge for these systems is maintaining the context from earlier parts of the conversation or between conversation sessions.

Machine Learning: A subset of AI in which machines learn from data rather than being specifically programmed for a task.

Bias: Skew or prejudice in a model caused by an unbalanced or prejudiced training data set. This is particularly common in large language models as they are often trained on the internet and therefore can reflect the bias of the general population.

Hallucination: When a generative AI model produces results that are incorrect or misleading. Some causes of hallucination include bias in the training data, incomplete training data, or incorrect reasoning. Since the reasoning is not exposed to the user it is hard to determine the source of hallucinations.

Prompt Engineering: Tweaking the input to a generative AI model to elicit a preferred response.

Adversarial AI: Intentionally feeding bad data into a model to get them to produce misinformation or harmful results. Since many generally available LLMs continue to train on user data this is a risk users need to be aware of.

Rebecca Thomas
www.beccathomas.me/ai-guides/

Created 10/2025