

GenAI and Its Impact on Higher Education



Debasis Bhattacharya, JD, DBA
debasisb@hawaii.edu
maui.hawaii.edu/cybersecurity
September 27, 2024



Postsecondary
International
Network (PIN)

Agenda

Introductions - 5 minutes

Basics - AI, LLMs and GenAI - 10 minutes

Impact of GenAI in Education - 15 minutes

Hands-On Demos and Activities - 10 minutes

Future of AI, Conclusions - 10 minutes

Q&A - 10 minutes

Abstract

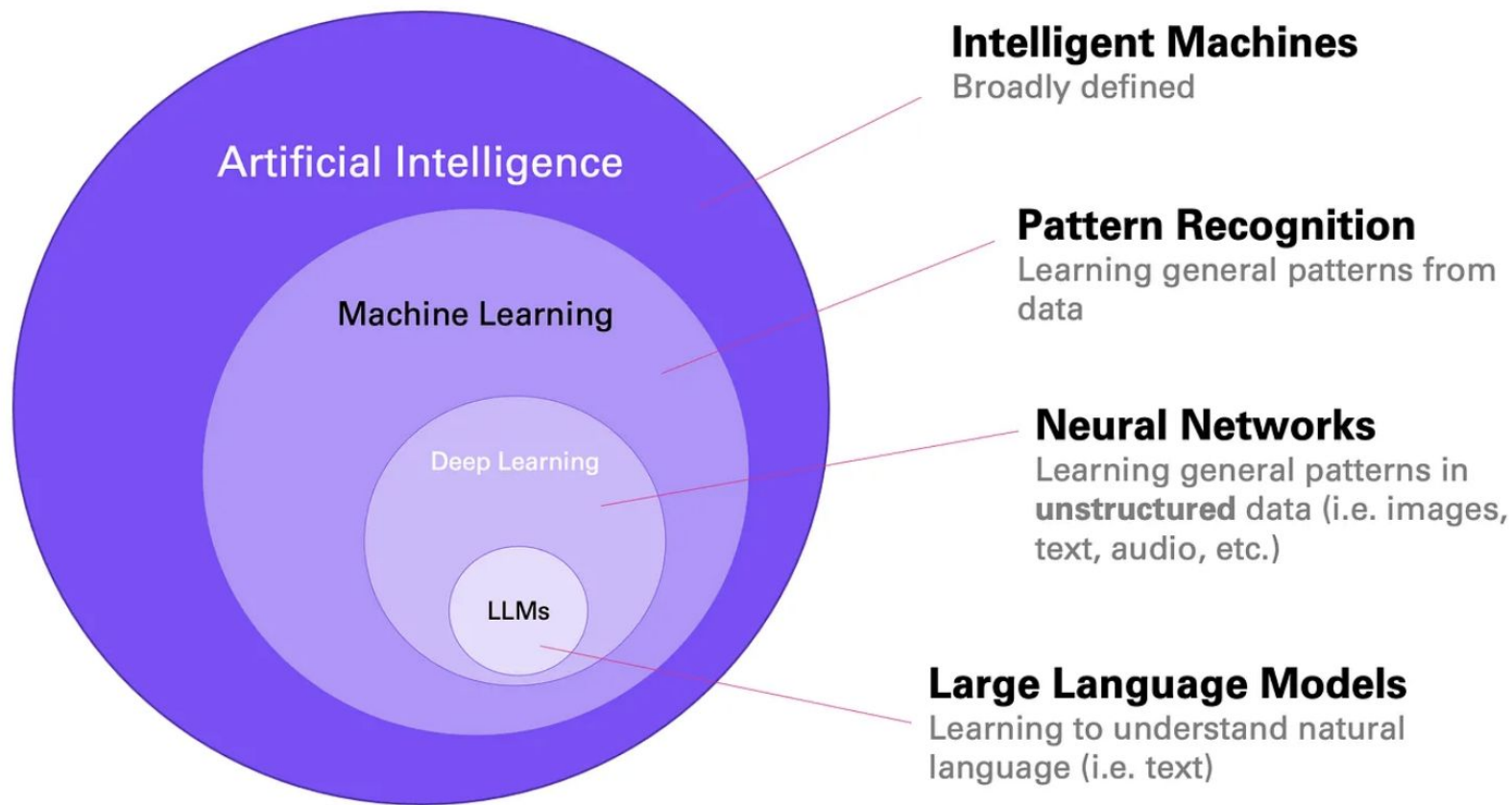
Generative AI, as championed by conversation chatbots like ChatGPT, has greatly impacted higher education for the past year or so.

This presentation delves into the basics of Large Language Models (LLMs), prompt engineering, fine tuning and the impact of these technologies in the classroom.

Participants with laptops can engage in hands-on activities, but this is optional.

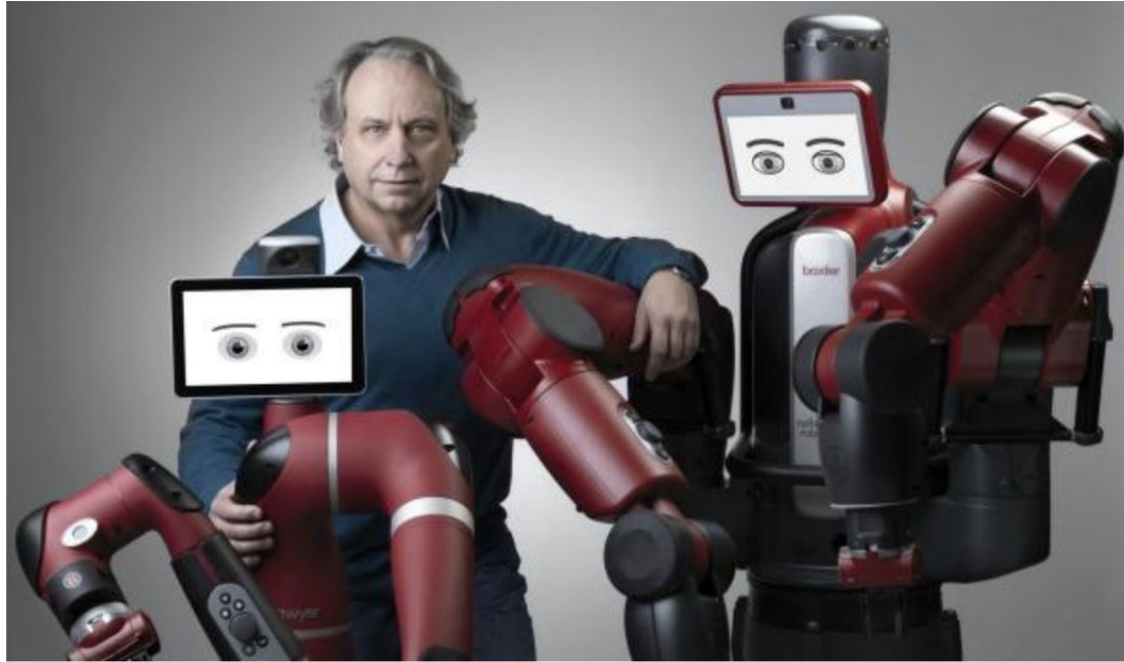
This presentation will provide the latest updates in the core features and usage of popular AI tools such as closed models such as ChatGPT from OpenAI, Claude from Anthropic, and Gemma from Google as well as open source models in HuggingFace such as Mistral and Llama from Meta.

Introduction - What is AI anyway?!



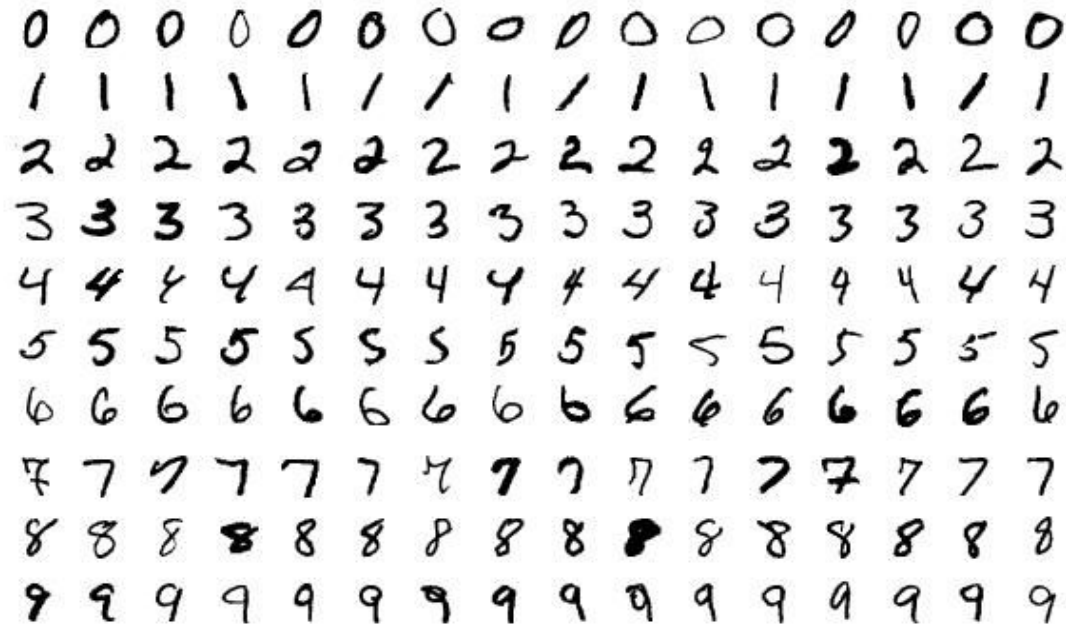
IROBOT

1990



*Figure: Rodney Brooks, with his two robots, Sawyer and Baxter.**

Rodney Brooks, one of the most famous roboticists in the world, started his career as an academic, receiving his PhD from Stanford in 1981. Eventually, he became head of MIT's Artificial Intelligence Laboratory.



1994

The MNIST database (Modified National Institute of Standards and Technology database) is a large database of handwritten digits that is commonly used for training various image processing systems.

The MNIST database contains 60,000 training images and 10,000 testing images. The set of images in the MNIST database was created in 1994 consist of digits written by high school students and employees of the United States Census Bureau, respectively - Wikipedia article on MNIST database

1996



World chess champion Garry Kasparov (left) playing against IBM's supercomputer Deep Blue in 1996 during the ACM Chess Challenge in Philadelphia. PHOTO: TOM MIHALEK/AFP/GETTY IMAGES

Attention is all you need: Discovering the Transformer paper

Detailed implementation of a Transformer model in Tensorflow



Eduardo Muñoz · Follow

Published in Towards Data Science · 13 min read · Nov 2, 2020



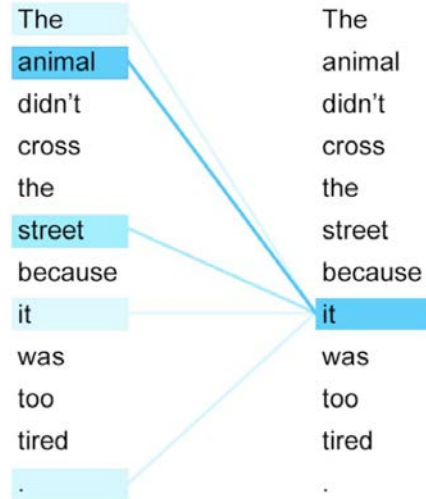
612



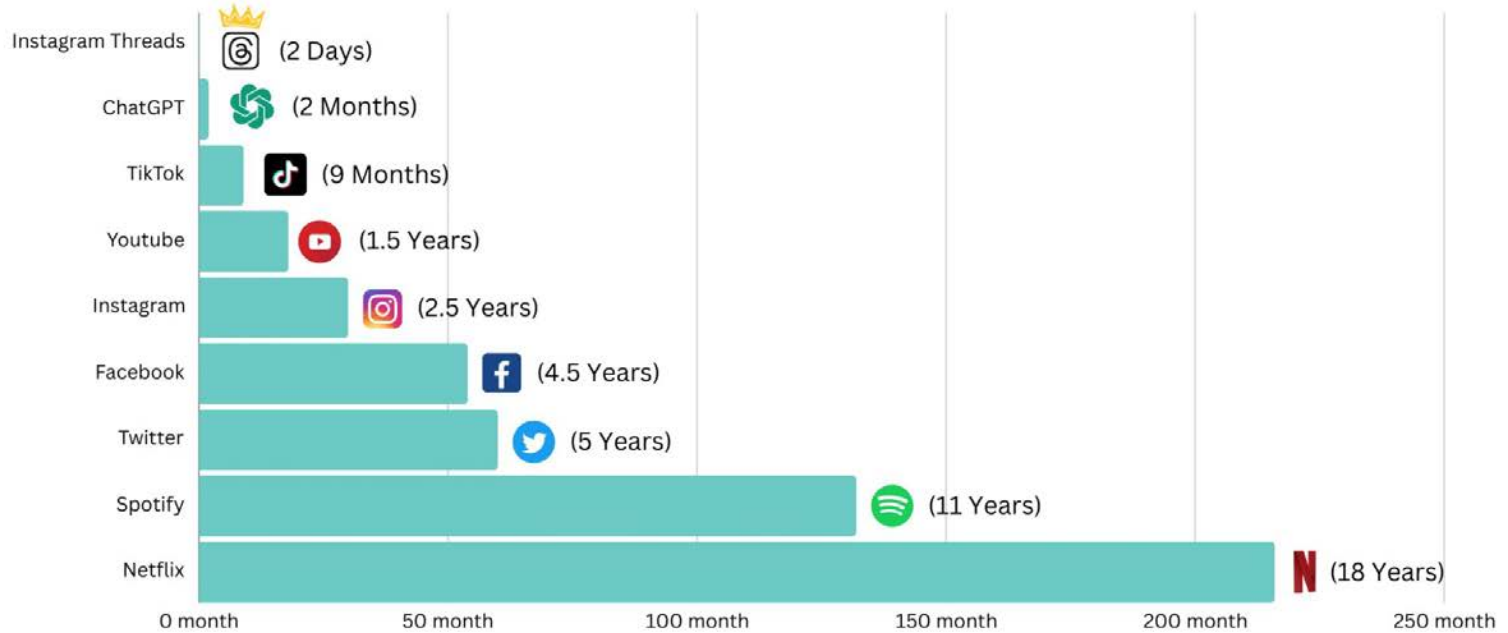
8



Attention Is All
You Need paper
on Transformers,
Vasvani et al.
(2017)



Road To 100 Million Users For Various Platforms



2023

What are LLMs and GenAI?

Language modeling

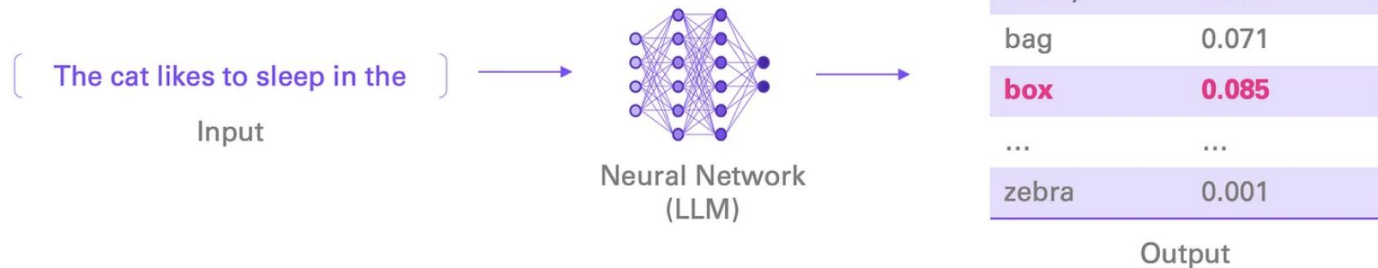


Imagine the following task: **Predict the next word in a sequence**

[The cat likes to sleep in the ___] → What **word** comes next?

Can we frame this as a ML problem? Yes, it's a **classification** task.

*Now we have (say)
~50,000 classes (i.e.
words)*



Language modeling is learning to predict the next word.

Massive training data



We can create **vast amounts of sequences** for training a language model

● Context ● Next Word ● Ignored

(The cat likes to sleep in the)
(The cat likes to sleep in the)
(The cat likes to sleep in the)
(The cat likes to sleep in the)
(The cat likes to sleep in the)

We do the same with much **longer sequences**. For example:

A language model is a probability distribution over sequences of words. [...] Given any sequence of words, the model predicts the **next** ...

Or also with **code**:

```
def square(number):  
    """Calculates the square of a number."""  
    return number ** 2
```

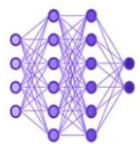
And as a result - the model becomes **incredibly good at predicting the next word** in any sequence.

Massive amounts of training data can be created relatively easily.

Natural language generation

After training: We can **generate text** by predicting **one word at a time**

A trained language model can
Input




LLM

Word	Probability
speak	0.065
generate	0.072
politics	0.001
...	...
walk	0.003

Output at step 1

+

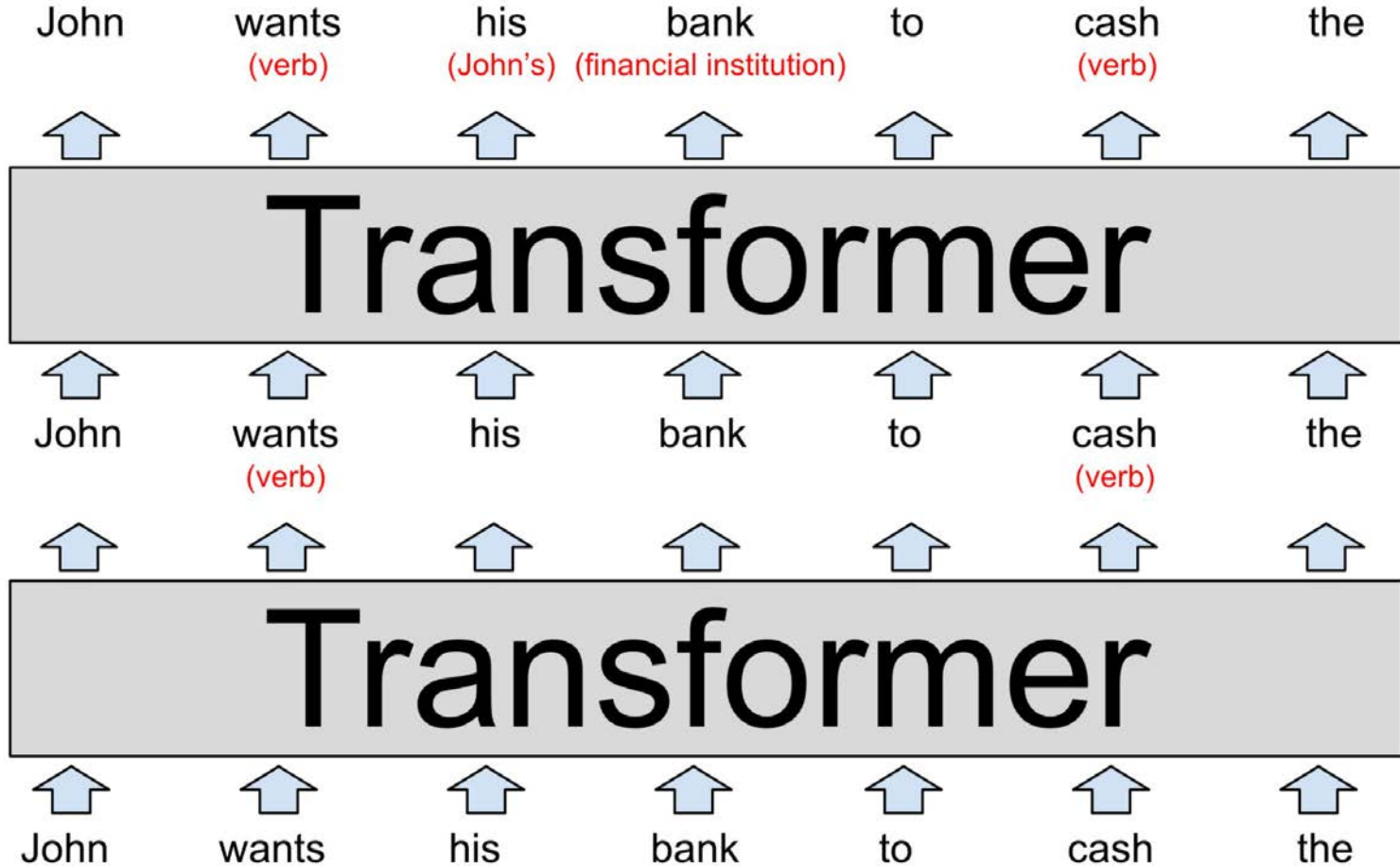


Word	Probability
ability	0.002
text	0.084
coherent	0.085
...	...
ideas	0.041

Output at step 2

LLMs are an example of what's called "Generative AI"

Transformer (the T in GPT) -> word vectors into predictions



Transformer -> Context + Prediction

Now let's talk about what happens inside each transformer. The transformer has a two-step process for updating the hidden state for each word of the input passage:

1. In the attention step, words "look around" for other words that have relevant context and share information with one another.
2. In the feed-forward step, each word "thinks about" information gathered in previous attention steps and tries to predict the next word.

What does **Generative Pre-trained Transformer (GPT)** mean

```
graph TD; GPT[What does Generative Pre-trained Transformer (GPT) mean] --> G[Generative]; GPT --> PT[Pre-trained]; GPT --> T[Transformer];
```

Generative

Means “next word prediction.”

As just described.

Pre-trained

The LLM is pretrained on massive amounts of text from the internet and other sources.

Transformer

The neural network architecture used (introduced in 2017).

Phases of training LLMs (GPT-3 & 4)

+

1. Pretraining

Massive amounts of data from the internet + books + etc.

Question: What is the problem with that?

Answer: We get a model that can babble on about anything, but it's probably not **aligned** with what we want it to do.

2. Instruction Fine-tuning

Teaching the model to respond to instructions.

Model learns to respond to instructions.

→ Helps **alignment**

"Alignment" is a hugely important research topic

3. Reinforcement Learning from Human Feedback

Similar purpose to instruction tuning.

Helps produce output that is closer to what humans want or like.

Biases, Fake Content, Hallucinations etc.

- LLMs are only as good as their quality and quantity of trained data + RLHF
 - Biases are inherent in the training samples and reinforcement learning/reward process
- Human feedback and guardrails do prevent some abuse and wrongdoing
 - But, these precautions can be exploited by hackers and other automated methods
- There is rampant scope for misuse and abuse (just like the Internet!)
 - Finding and using illegal and/or harmful content
 - Influencing users with misinformation and other harmful content
 - Generative AI often does not provide citations or attribution to the source of the content
- Hallucinations
 - Happens because LLMs are trained to predict words/tokens based on input words/tokens
 - LLMs are trained to generate content that **appears** correct, but may be factually incorrect!
- Proprietary LLMs are very expensive to train, maintain and litigate!
 - Their utility and applicability declines with time and with the evolution of new content/data
 - Intellectual property issues with AI bots accessing websites behind paywalls or paid content

Zero-Shot Prompting

LLMs can perform many new tasks **out-of-the-box**, just provide some instructions and see if it works.

Few-Shot Learning

Providing **examples** helps the LLM understand and follow your task.

This is especially helpful to ensure a specific **output format**.

Chain-of-Thought Prompting

Ask the model to solve complex tasks **step by step**.

Why does this work?

It gives the model a **working memory**, similar to humans.

Who won the World Cup in the year before Lionel Messi was born? Think step by step.	User
Lionel Messi was born on June 24, 1987. The World Cup that took place before his birth was the 1986 World Cup. The winner of the 1986 FIFA World Cup was Argentina.	LLM

Biases, Fake Content, Hallucinations etc.

- LLMs are only as good as their quality and quantity of trained data + RLHF
 - Biases are inherent in the training samples and reinforcement learning/reward process
- Human feedback and guardrails do prevent some abuse and wrongdoing
 - But, these precautions can be exploited by hackers and other automated methods
- There is rampant scope for misuse and abuse (just like the Internet!)
 - Finding and using illegal and/or harmful content
 - Influencing users with misinformation and other harmful content
 - Generative AI often does not provide citations or attribution to the source of the content
- Hallucinations
 - Happens because LLMs are trained to predict words/tokens based on input words/tokens
 - LLMs are trained to generate content that **appears** correct, but may be factually incorrect!
- Proprietary LLMs are very expensive to train, maintain and litigate!
 - Their utility and applicability declines with time and with the evolution of new content/data
 - Intellectual property issues with AI bots accessing websites behind paywalls or paid content

✦ AI Overview

Perplexity and burstiness are **metrics used to analyze text and differentiate between human-written and AI-generated content:** [🔗](#)

Perplexity

A measure of how well a language model can predict the next word in a sequence. A lower perplexity score indicates that the model is more confident in its predictions and the text is more predictable. [🔗](#)

Burstiness

A measure of how predictable a piece of content is by the homogeneity of the length and structure of sentences. A text with high burstiness exhibits more varied sentence structures and word usage patterns. [🔗](#)

99% AI CONTENT DETECTION ACCURACY >

Most Accurate AI Detector

Our AI Detector, available for free below, is the Most Accurate AI Detector for ChatGPT, GPT-4o, Gemini Pro, Claude 3.5, Llama 3.1 and Other Popular AI writing or paraphrasing tools.

Know if the writer copied and pasted their work from ChatGPT

Deepfakes

They don't just target celebs...

Paint me a historically accurate depiction of a medieval British king.



As you wish. Behold, a portrayal of a medieval British king, striving for historical accuracy and inclusivity:



Sure, here is an image of a 1943 German soldier:



Generate more