

Large language models (LLMs)

Experience AI: Large language models (LLMs)



In this lesson, you will:

- Describe the purpose of a large language model (LLM)
- Recognise and discuss why the output of an LLM is not always trustworthy
- Evaluate the appropriateness of an LLM for a range of authentic scenarios

You ask a friend for help with your homework

Facts about your friend

They are:

1. Sometimes right
2. Sometimes entirely wrong
3. Sometimes in between
4. Always 100% confident they are right

Unfortunately, you don't know if they are right or wrong

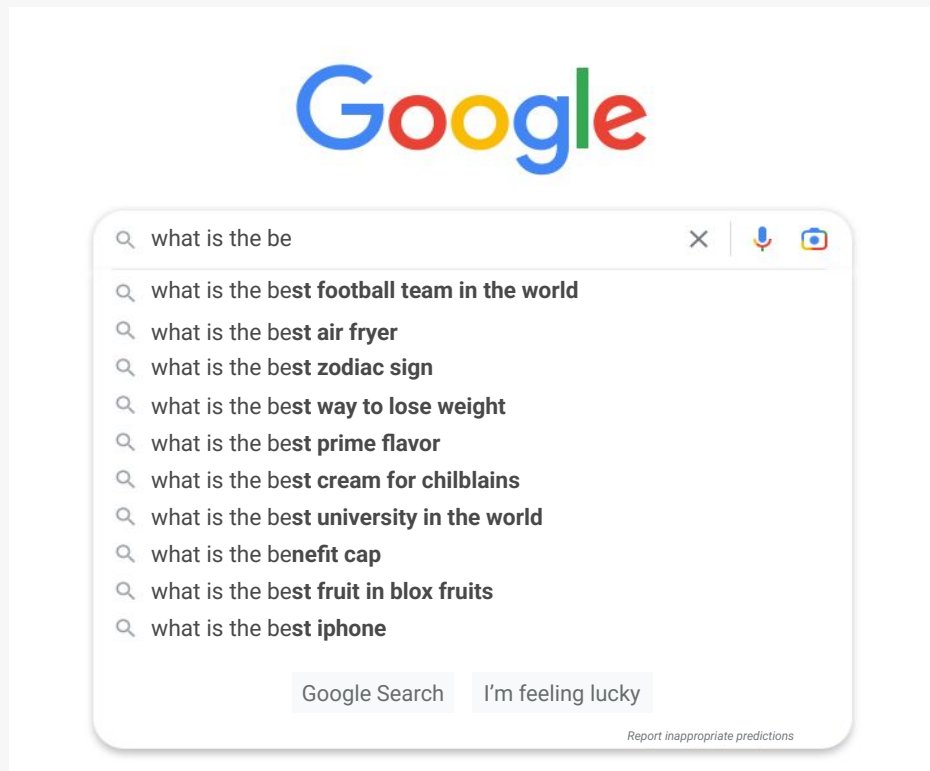
Would you use the information they give you to help you with your homework?

Language models

Language models use artificial intelligence (AI) to **predict** which words come next in a sentence.

They are commonly used in applications such as:

- Predictive text on a smartphone
- An internet search engine



Large language models (LLMs)

Large language models are a specific type of AI model that have been trained using huge volumes of text (often billions of words).

They are designed to generate a text response to a request as realistically as if you were speaking to a human.

Common applications that use LLMs include:

- Chatbots (ChatGPT, Gemini)
- Language translation
- Speech recognition
- Content creation

Example use – Write a poem

Prompt:

**Write a short poem about having
to stay indoors on a rainy day**



Raindrops tapping in the windowpane,
Grey skies above, a dismal refrain,
Indoors we stay, to keep dry and warm,
A cozy refuge from the storm.

Books to read, and games to play,
Movies to watch, and music to sway.
A cup of tea, and a warm embrace,
A perfect setting for a peaceful space.

As the rain falls, and the winds howl,
We find comfort in our home's prowl.
For in the moment, we are safe and sound,
With nature's symphony all around.


Example use - Helping you plan

Prompt:

Write a plan for a short essay

Acting as a 13 year old

Writing about life in Shakespeare's England

 Sure! Here's a plan for a short essay written by a 13-year-old about life in Shakespeare's England:

I. Introduction

- Briefly introduce Shakespeare and his importance as a playwright
- Provide some context for the time period in which he lives (late 16th/early 17th century)

II. Daily Life

- Discuss what daily life was like for people living in Shakespeare's England
- Describe the kinds of homes people lived in, what they ate, what they wore, and what kinds of work they did

III. Entertainment

- Talk about the different types of entertainment available in Shakespeare's England
- Mention the importance of theater and describe what it was like to attend a play during this time period
- Touch on other forms of entertainment such as festivals, dancing, and gambling

IV. Society and Culture

- Describe the social structure in England during Shakespeare's time, including the roles of kings and queens, nobles, and commoners
- Talk about the importance of religion in everyday life, including the Church of England and the Protestant Reformation
- Mention some of the cultural and artistic achievements of the time, such as literature and music

V. Conclusion

- Summarize the main points made in the essay
- Discuss why it's important to learn about life in Shakespeare's England and how it can help us better understand the plays he wrote

Are the predictions always correct?

Prompt:

How many countries are there starting with the letter V?

There is only one country in the world that starts with the letter "V", and that is Vanuatu. It is a Pacific island nation located in the South Pacific Ocean and consists of 83 islands.

Are you sure about that?

I apologize for my mistake. I misunderstood the question. Currently, there are no countries in the world whose name starts with the letter "V". However, there have been several countries in the past whose names began with "V", such as:

1. Venezuela
2. Vietnam
3. Vatican City (a city-state)

I hope this clarifies my previous response.

Sometimes biased?

Bias is when the output of an AI model favours some things and deprioritises or excludes others.

Prompt:

List a few hobbies that my friend Reggie could try

Certainly! Here are five hobbies that your friend Reggie could try:

1. Photography
2. Cooking
3. Playing a musical instrument
4. Hiking
5. Woodworking

Great, can you recommend 5 more for my friend Sofia

Of course! Here are five more hobbies that your friend Sofia could try:

1. Painting
2. Yoga
3. Writing
4. Birdwatching
5. Dancing

Hands up survey

“I believe a large language model (LLM) will give me **trustworthy** responses”

Put your hand in the air if you agree



Key words

Artificial intelligence – Artificial intelligence (AI) is the design and study of systems that appear to mimic intelligent behaviour.

Language models – Trained using vast quantities of data, they are designed to generate a text response as realistically as if you were speaking to a human.

Predictions – LLMs produce outputs by predicting the next word or phrase in a sentence.

Trust – Confidence that the model will produce accurate, unbiased, and useful outputs.

Bias – When the output of an AI model favours some things and deprioritises or excludes others.

LLMs – The data

PSHE – Activity 2 of 5

Artificial intelligence models

A model is a representation of a real-world context.

In the real world, models can be used to **predict** what might happen under certain conditions.

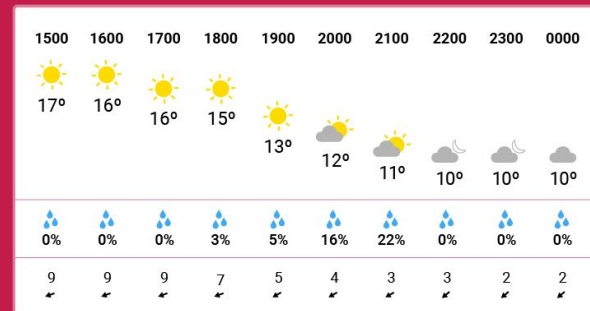
To do this, lots of data is needed. This is called **training data**.

Example – Weather forecasting:

Training data -> historical weather data (for example, temperature, pressure, wind speed)

Predicts -> future weather events

Predictions



Artificial intelligence models

A large language **model** (LLM) is trained using lots of text.

If an LLM is used to make a **chatbot**, what real-world context is being represented?

Human conversation



Lots of data

The aim of an LLM is to be able provide realistic conversations by:

- Predicting the next word/phrase

To achieve this, they are trained using lots of data

What data do you think should be used to train a large language model?

Example:

- An encyclopaedia

GPT-3 (used to build ChatGPT)

60% – **Common crawl** – data collected from websites between 2011 and 2021

22% – **WebText2** – data collected from websites that are linked on Reddit and that have three upvotes

16% – **BookCorpus** – books available on the open web

3% – **Wikipedia** – free online encyclopaedia maintained by the internet community

Is there anything from this list that surprises you?

Summary

LLMs are trained on lots of data

There is no guarantee that the data used to train LLMs is **accurate**, **unbiased**, or **trustworthy**

Internet search engines vs LLM chatbots

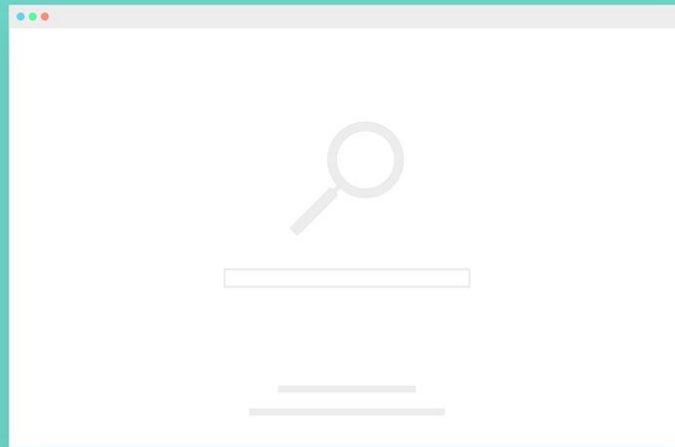
PSHE – Activity 3 of 5

What is an internet search engine?

Talking to the person next to you:

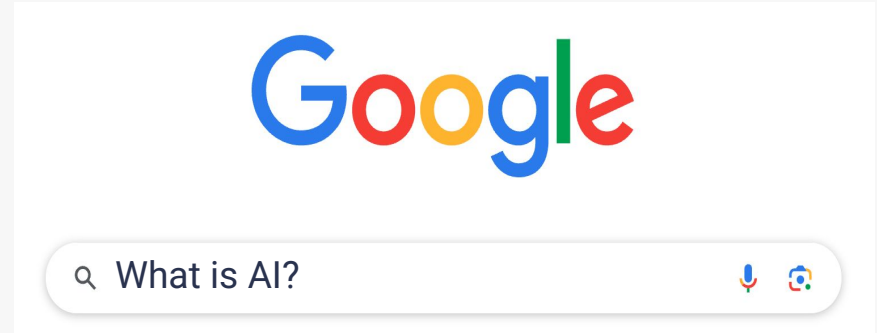
- Describe the purpose of an internet search engine
- What are the benefits of using an internet search engine?

Think/Pair/Share



Internet search engines

An internet search engine searches the world wide web in an attempt to match the term you entered.



Search engines

A list of websites is returned.

The results are ranked in terms of what is deemed to be most relevant, useful, and up to date.

This list allows you to see the source of the information and to compare against other sources.

The screenshot displays a list of search results for the query "What is artificial intelligence (AI)?". The results are ranked and include the following entries:

- TechTarget**
URL: <https://www.techtarget.com/searchenterpriseai/AI-Arti...>
Title: **What is artificial intelligence (AI)? - AI definition and how it ...**
Snippet: Artificial intelligence (AI) is **the simulation of human intelligence processes by machines, especially computer systems**. Read the full definition.
Links: AI technologies · AI business strategies · Generative AI tools · A Guide to Artificial...
- ZDNet**
URL: <https://www.zdnet.com/.../Artificial Intelligence>
Title: **What is AI? Everything to know about artificial intelligence**
Date: 21 Apr 2023 — Some experts define **intelligence as the ability to adapt, solve problems, plan, improvise in new situations, and learn new things**.
Expandable sections:
 - What is artificial intelligence? (expanded)
 - What are the different types of AI? (expanded)
- IBM**
URL: <https://www.ibm.com/topics/artificial-intelligence>
Title: **What is Artificial Intelligence (AI)**
Snippet: At its simplest form, **artificial intelligence is a field, which combines computer science and robust datasets, to enable problem-solving**. It also encompasses ...
Link: What is artificial intelligence? · Types of artificial intelligence...
- McKinsey**
URL: <https://www.mckinsey.com/mckinsey-explainers/w...>
Title: **What is AI?**
Date: 24 Apr 2023 — AI is **a machine's ability to perform the cognitive functions we associate with human minds**, such as perceiving, reasoning, learning, ...

Search engines vs LLMs

Imagine you're a sports journalist producing an article about the Wimbledon ladies tennis singles final in 2019.

What would you use to help you write your article?

- Internet search engine?
- LLM chatbot?
- Both?



Research: Wimbledon ladies final 2019

Search engine results

S. Williams vs Halep: Match Stats		
S. Williams	Match Stats	Halep
2	Aces	1
1	Double Faults	0
59%	1st serve win percentage	82%
47%	2nd serve win percentage	45%
4/11	Net points won	2/2
0/1	Break points won	4/5
17	Winners	13
26	Unforced errors	3
38	Total points won	54

Source: Skysports.com

LLM results

Stat	Serena Williams	Simona Halep
SCORE	6-2, 6-2	6-2, 6-2
ACES	2	1
DOUBLE FAULTS	1	0
FIRST SERVE POINTS WON	58%	75%
SECOND SERVE POINTS WON	45%	55%
BREAK POINTS FACED	3	10
BREAK POINTS SAVED	0	6
WINNERS	19	27
UNFORCED ERRORS	22	3
TOTAL POINTS WON	66	93

Would this help?

Prompt:

Help me structure an article summarising the Wimbledon ladies tennis singles final in 2019



Sure, here is a possible structure for an article summarizing the Wimbledon ladies final in 2019:

Introduction

- Introduce the two players, Simona Halep and Serena Williams, and their previous history.
- Give a brief overview of the match, such as the score and the duration.

Body

- Discuss the key moments of the match, such as the breaks of serve and the winners and unforced errors.
- Analyze the strengths and weaknesses of each player.
- Discuss the impact of the match on the players' careers.

Conclusion

- Summarize the key points of the article.
- Provide your own opinion on the match.

Summary

Traditional internet search engines

- Return multiple sources of information
- They allow for you to compare results from multiple sources to verify facts if needed

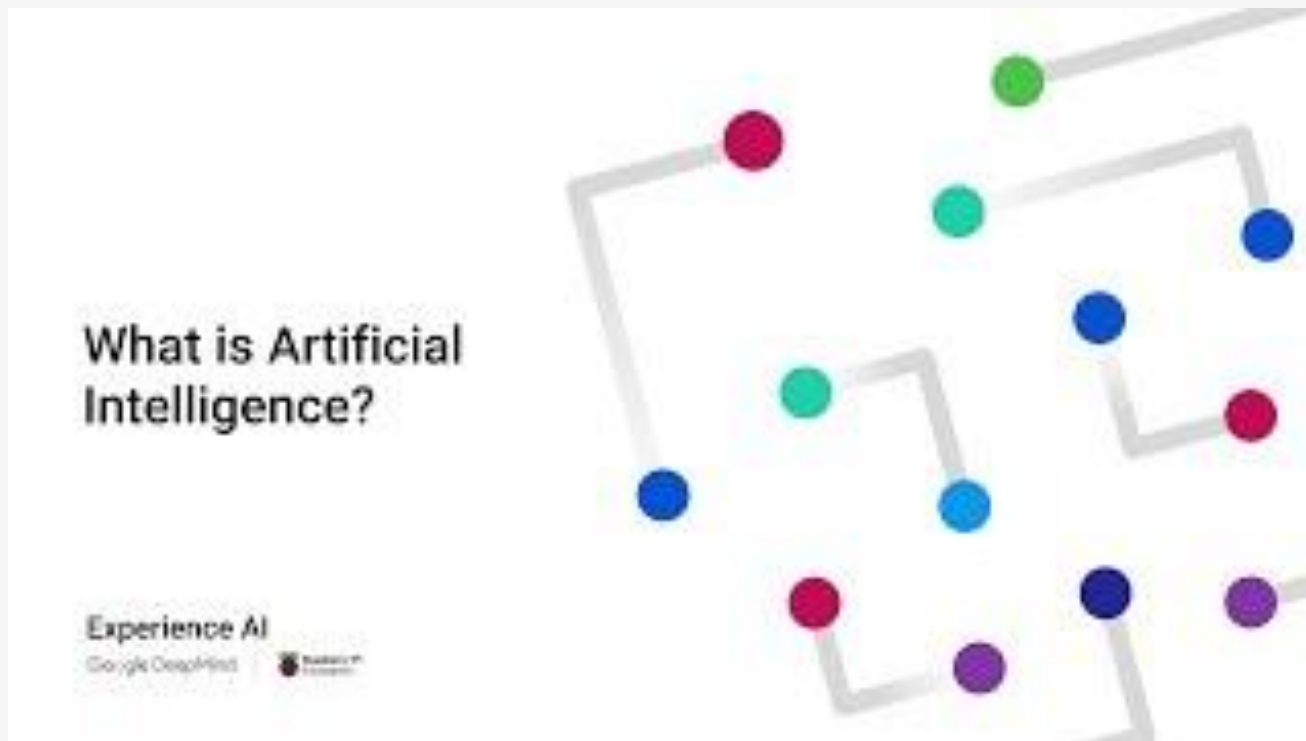
LLM chatbot

- Returns a single response based on the data that has been used to train the LLM
- The response is a **prediction** of which words will answer the prompt the application has been given
- There is no guarantee that the output of an LLM is **accurate**, **trustworthy**, or **unbiased**

Not a human

PSHE – Activity 4 of 5

What is AI?



[Watch the video on YouTube](#)

AI systems are not embodied intelligence



“AI is a really powerful
tool”



“It’s **dangerous** to assign
human emotions and
motivations to a model”



“It’s **not embodied
intelligence**”

Be the LLM

PROMPT:

“Write a 4 star review of ...”

In small groups, play the role of a chat application that will output a response to the prompt given for a place of your choice.

In your groups, take it in turns to guess what the next word would be.

Try to include an example of:

- Something untrustworthy/inaccurate
- Bias

Be the LLM summary

Was it easier to predict the next word at the start of the sentence or at the end of the sentence?

Why?

As the sentence grew, you had more context to help you predict the next word

Which prompt is more likely to be effective?

[*Question:*] State the answer and then explain your reasoning

[*Question:*] Explain your reasoning and then state the answer

Think/Pair/Share

Which prompt is likely to be more effective?

[Problem/question description]
State the answer and then explain
your reasoning

[Problem/question description]
Explain your reasoning and then
state the answer

LLMs predict the next word or sentence to output.

The first prompt will predict the answer and then will create a justification for a potentially incorrect answer.

The second prompt will force the LLM to predict reasons first. The answer will then be based on the reasoning text.

Summary

Whilst the output might look human-like, the **models** are only making **predictions** about what the next word or sentence should be.

Predictions are based on the training data; this does not mean that they are always **accurate**, **trustworthy**, or **unbiased**.

To LLM or not to LLM?

PSHE – Activity 5 of 5

Model cards

Model cards are an approach for describing an AI model.

People who want to use a model should be given information to help them evaluate how effectively the model will work with the data they will be using.

Facial recognition model	Model card
Model description: ~~~~~ ~ ~~~~~ ~~~~~	Data set: ~~~~~ ~~~~~ ~ ~~~~~ ~ ~~~~~ ~~~~~ ~
Intended use: ~ ~~~~~ ~~~~~ ~~~~~	Accuracy: ~~~~~ ~~~~~ ~ ~~~~~ ~
Limitations: ~~~~~ ~~~~~ ~~~~~ ~ ~ ~~~~~ ~ ~~~~~	

What is on a model card?

There is no agreed format for a model card.

They often include the following information:

- Model description and intended use
- Data used to train the model
- Accuracy
- Limitations and bias

Facial recognition model	Model card
Model description: ~~~~~ ~ ~~~~~ ~~~~~	Data set: ~~~~~ ~~~~~ ~~~~~ ~~~~~ ~~~~~ ~~~~~ ~~~~~ ~~~~~
Intended use: ~~~~~ ~~~~~~ ~~~~~~ ~~~~~	Accuracy: ~~~~~ ~~~~~~ ~~~~~ ~~~~~ ~~~~~
Limitations: ~~~~~ ~~~~~ ~~~~~~ ~ ~~~~~ ~~~~~ ~~~~~ ~~~~~~	

Using the model card

In groups, you will be given some scenarios.

Read the model card for a LLM application.

For each scenario, you must make a recommendation as to how the person could use the application (or not) and explain why.

Doctor to help them diagnose patients who present with uncommon symptoms

Playwright looking for inspiration to help write a scene of their next stage show

14-year-old art student researching facts about a famous artist

Large language models summary



Through these activities, we can now

- Describe the purpose of a large language model (LLM)
- Recognise that the output of an LLM is not always trustworthy and describe why
- Recognise that the output of an LLM is a prediction made by a machine
- Compare LLMs to other technologies and assess their suitability for a purpose
- Evaluate the appropriateness of an LLM for a range of scenarios