1

AI

INTRODUCTION

1

AI

Artificial intelligence (AI) is a field of **computer science** that encompasses a range of technologies. Their combined use enables computers to perform specialised, complex tasks that are often **associated with human capabilities.**

AI differs from traditional software in its **autonomy, ability to learn** and improve through interactions with its environment.

Museum of Prague

INTRODUCTION

2

CHATBOT

INTRODUCTION

2

CHATBOT

A chatbot is **software** programmed to **conduct conversations** using either text or voice.

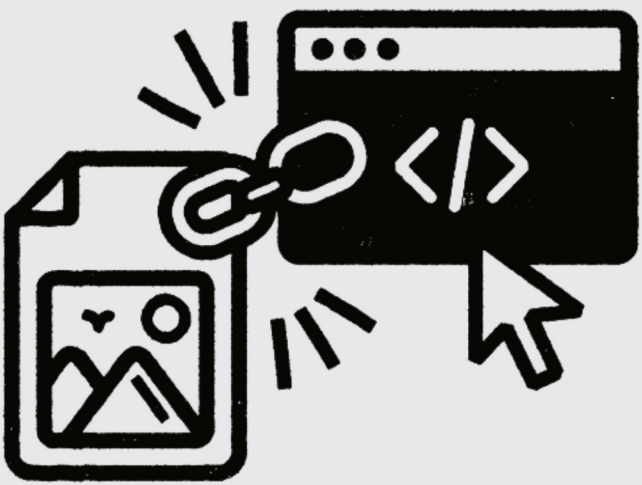
A chatbot **answers** questions, **solves** problems, or **generates** responses across a wide range of topics.

Thanks to extensive training chatbots can mimic human communication.

Museum of Prague

INTRODUCTION

3 EMBEDDING



THEORY

3 EMBEDDING

An embedding is a **numerical, vector-based representation of text,** images, or metadata that enables a chatbot to **understand** meaning, **search for** information, **compare** texts, and maintain the context of a conversation.

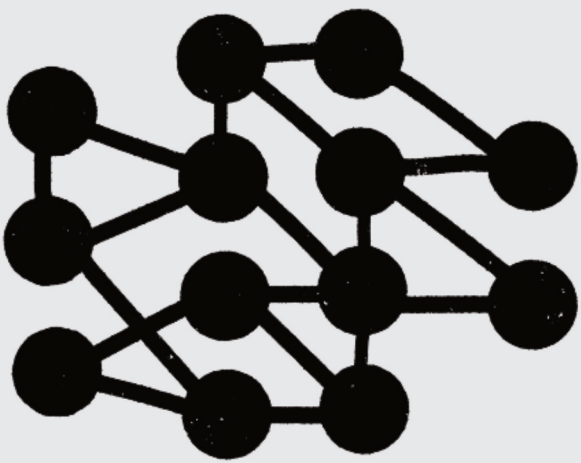
The relationship between words, and thus between numbers, can be visualised as follows:

The word "museum" is close to the word "gallery" in the latent space, but far from "motor oil."

Museum of Prague

THEORY

4 NEURAL NETWORKS



THEORY

4 NEURAL NETWORKS

A neural network is a **fundamental computational structure and a type of artificial intelligence** inspired by how **neurons in the human brain** function. It consists of many interconnected **"nodes"** that work together to process information and learn to recognise patterns in data.

The more data and layers it has, the better it can **understand** more complex problems, such as image recognition or text comprehension.

Museum of Prague

THEORY

5

LLM

THEORY

5

LLM

An LLM (Large Language Model) is a **type of artificial intelligence** capable of **processing human language.**

During training, **it learns** from a massive amount of text data, **enabling it to answer** questions or generate text in a manner similar to a human.

Using patterns in the data, an LLM **predicts** which words should follow to produce a meaningful response. The LLMs that power chatbots are neural networks.

Museum of Prague

THEORY

6 TRAINING

THEORY

6

TRAINING

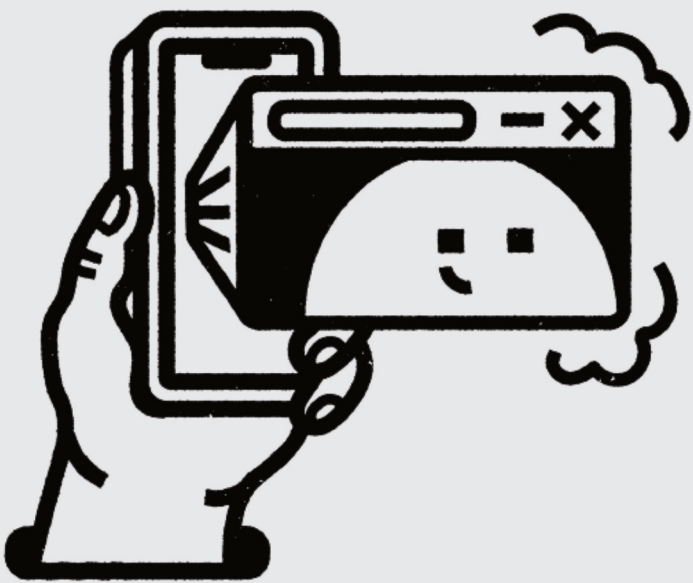
Training is the **process** by which an AI model **learns from large amounts of data** to make predictions or generate outputs.

During training, it **looks for patterns and relationships** in the data and gradually adjusts its "internal settings" to ensure the results are as accurate as possible.

During training, the model **receives a "reward" or "penalty,"** which motivates it to produce more accurate results.

THEORY

7 USER INTERFACE



ELEMENTS

7 USER INTERFACE

The user interface (UI) is the **space** where the **user interacts with the chatbot**, the system.

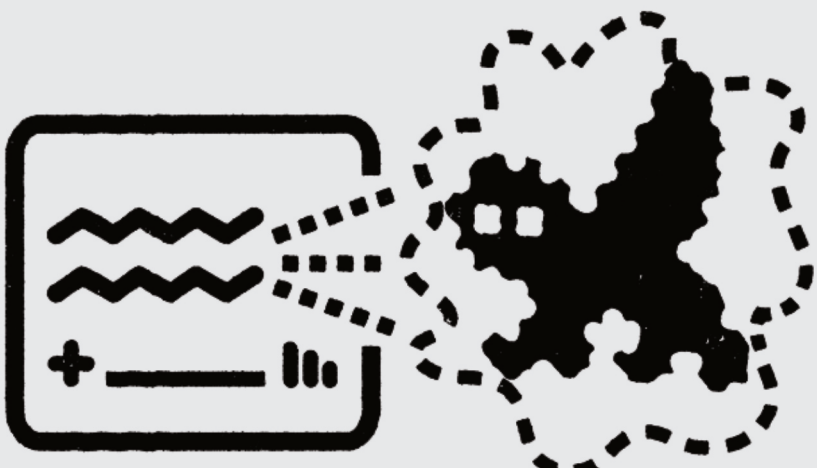
It includes visual and interactive elements:

- text fields,
- speech bubbles
- buttons...

ELEMENTS

8

PROMPT



ELEMENTS

8

PROMPT

A prompt is **the input** that the AI responds to. It can be a question, a task, an instruction, an image, or a description.

A prompt can be entered in **natural language** or in a **structured format** and can specify the type of response we want: text, an image, a graph, etc.

The **clearer** and more specific the prompt is, the **better** the result the AI will typically produce.

Museum of Prague

ELEMENTS

9

TOKEN

ELEMENTS

9

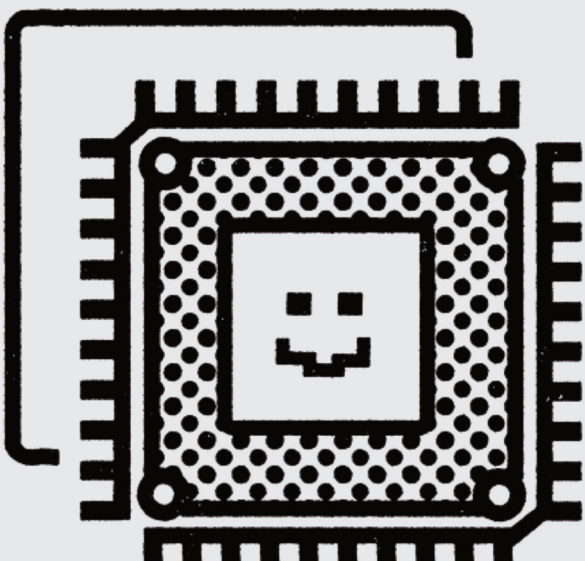
TOKEN

A token is **the basic unit of text for AI**. A single token can be an entire word or just a part of one, it depends on the language and the type of model.

AI breaks text down into individual tokens and process them one by one. Each model has a limit on the number of tokens it can process at once. If a conversation exceeds this limit, the chatbot will start to forget the beginning of the conversation.

ELEMENTS

10 PROCESSORS



ELEMENTS

10 PROCESSORS

AI models require a server with powerful hardware, particularly **graphics processing units (GPUs)**, which can handle many computations simultaneously and are essential for both training and running modern AI.

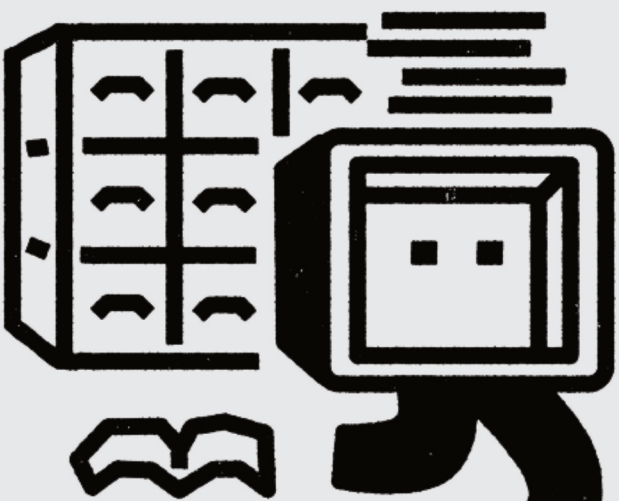
GPUs are similar in performance to the architecture of neural networks.

This significantly speeds up the computation process compared to using a **central processing unit (CPU)**.

Museum of Prague

ELEMENTS

11 DATA CENTRES



ELEMENTS

11 DATA CENTRES

Data centres are **specialised buildings or facilities** that house hundreds to thousands of servers used to store and process data and services. They are designed to operate **24/7**.

Thanks to them, businesses and individual users can access their **online services from anywhere** without maintaining their own high-performance servers.

Museum of Prague

ELEMENTS

12

ENVIRONMENTAL IMPACTS

RISKS

12

ENVIRONMENTAL IMPACTS

Training and running models require powerful servers and **enormous amounts of energy,** as well as a continuous supply of **electricity** and, in some cases, **water** for cooling.

A single prompt may have a small carbon footprint, but with billions of queries per day, the impact adds up, and **consumption rises.**

The growing demand for GPUs and other specialised components also has indirect impacts, such **as emissions from mining and manufacturing.**

Museum of Prague

RISKS

13

MISUSE OF PERSONAL DATA

RISKS

13

MISUSE OF PERSONAL DATA

The chatbot stores and analyses both text and **metadata** that may reveal **sensitive information.** Third parties can link this data together to create a **detailed user profile.**

This profile can be used for **targeted marketing**, risk scoring, or for manipulation and blackmail.

Even content entered "**just for training**" can become part of the model and later be **indirectly reproduced** in responses.

Museum of Prague

RISKS

14 DATA POISONING

RISKS

14 DATA POISONING

Data poisoning is an attack in which **malicious or fraudulent data is intentionally injected into an AI's training data**.

The goal is to **influence the model's behaviour**, such as skewing its outputs or creating a hidden vulnerability. The model then learns **incorrect** patterns and may **generate unreliable or manipulated results.**

Museum of Prague

RISKS

15 HALLUCINATIONS

RISKS

15 HALLUCINATIONS

AI hallucinations are **false information** that is presented **as true.**

The model **merely predicts** the most probable continuation of the text and does not fact-check. When details on a topic are absent, the AI "fills in" based on its training.

LLMs can also create a **misleading context** by combining two different facts.

Museum of Prague

RISKS

16

UNSUPERVISED AGENT

RISKS

16

UNSUPERVISED AGENT

An unsupervised agent in a chatbot setting **poses a risk** because it can independently perform actions that a human may not fully understand or control.

The agent **lacks "common sense,"** so it may act very efficiently but inappropriately.

Museum of Prague

RISKS

17 DATA VALIDATION

MITIGATION

17 DATA VALIDATION

Data validation involves verifying both input and output data to ensure they are **accurate, consistent, and secure.**

This check can be carried out by a human or an automated tool.

Validation is **vital for AI safety. Low-quality or faulty data** can lead to **hallucinations**, misinterpretations of queries, or the execution of unintended instructions.

Museum of Prague

MITIGATION

18

SANDBOXING

MITIGATION

18

SANDBOXING

Sandboxing **stops** the agent from running unverified code or operating outside allowed areas, thus **protecting** both the server and users from harm.

In the context of a chatbot, sandboxing involves **creating a separate, controlled environment.**

Museum of Prague

MITIGATION

19

RENEWABLE RESOURCES

MITIGATION

19

RENEWABLE RESOURCES

Training large models can use as much energy as a small town.

The shift to renewable energy sources is **vital for advancing AI,** making it more sustainable and reducing its climate impact.

Museum of Prague

MITIGATION

20 ANONYMISATION

MITIGATION

20 ANONYMISATION

Data anonymisation involves removing or modifying personal data so that AI cannot identify a specific individual and does not compromise the user's privacy.

Museum of Prague

MITIGATION

21

RAG

MITIGATION

21

RAG

RAG (Retrieval-Augmented Generation) is an AI architecture that combines **document-based information retrieval** with text generation using large language models (LLMs).

The model does not simply respond on the basis of what it has "learned from its parameters", it first locates relevant passages in a knowledge base and then adds them to the prompt as context.

The aim is to **reduce hallucinations**, improve accuracy, and enable the model to work with current or specialised information not used in its training.

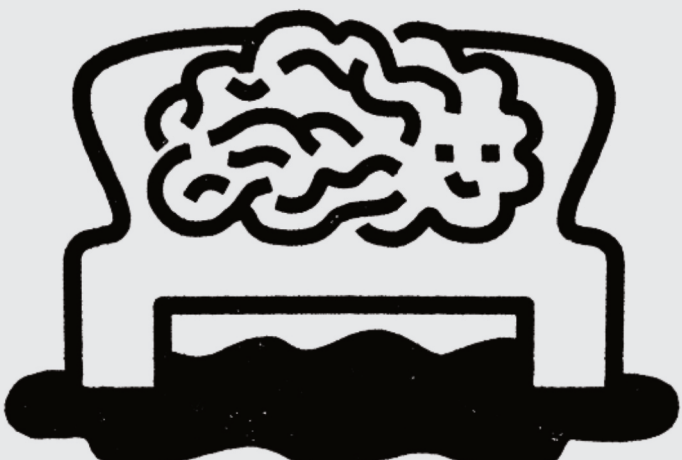
RAG does not require retraining the model; instead, updating the external knowledge base (source material) is **quick and cost-effective.**

Museum of Prague

MITIGATION

22

LOCAL LLM



MITIGATION

22

LOCAL LLM

A local LLM (large language model) is a **dataset stored on** a personal or corporate device, so there is **no need to send** queries to remote data centres.

A smaller model also means **lower power consumption.** By bypassing the cloud centre, the carbon footprint is reduced.

Another advantage of local LLMs is greater **security**. The data remains on devices locally.

Museum of Prague

MITIGATION

23

MODEL OPTIMALISATION

MITIGATION

23

MODEL OPTIMALISATION

Model optimisation is achieved, for example, by pruning models, converting them to more efficient formats (quantisation), or fine-tuning only a subset of parameters (with lower energy consumption).

The result is models that run **faster, more cost-effectively, and more sustainably,** which is particularly crucial for AI, which otherwise requires enormous amounts of energy and water to cool data centres.

Specialised hardware designed solely for AI computations (TPUs) can also be used instead of general-purpose GPUs.

MITIGATION

24 FUTURE OUTLOOK

24 FUTURE OUTLOOK

Where do you think AI is headed?

Museum of Prague

0 CREDITS

0 CREDITS

Team Lead: MgA. Vojtěch Leischner, PhD.
Project Manager: Mgr. Monika Švajková
Editors: Mgr. Monika Švajková, Mgr. Martina Mikolas
Technical Proofreading: Ing. Ondřej Kuželka, PhD.
Language Consultant: Jack Schroeder, PhD.
Graphic Design: BcA. Sarah Belejová

Museum of Prague

0 CREDITS

0 CREDITS

Published by Methodological Centre for the Implementation of AI in Museology.

2026

Museum of Prague