LECTURE 2: MACHINE LEARNING AND LLMS

AARHUS BSS
DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

10 SEPTEMBER 2025

MAGNUS BENDER
ASSISTANT PROFESSOR

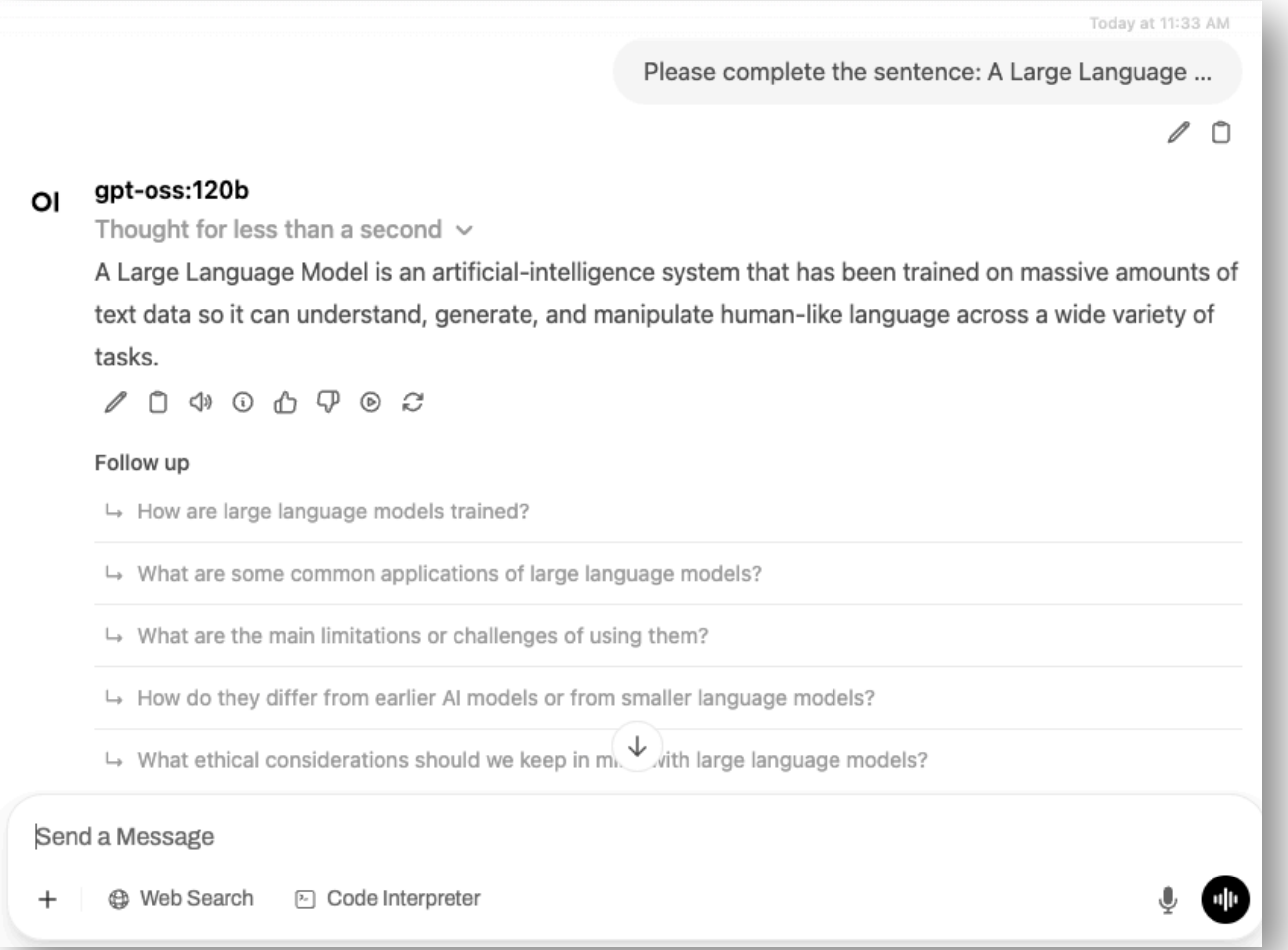
TUTORIALS

1. Quick recap on last week's tutorial

TUTORIALS

1. Quick recap on last week's tutorial
2. Outlook on this week's tutorial
   - Topic:
     Python basics and data types
   - Recommended:
     Complete the Datacamp courses
     - Introduction to Python for Developers
     - Intermediate Python for Developers

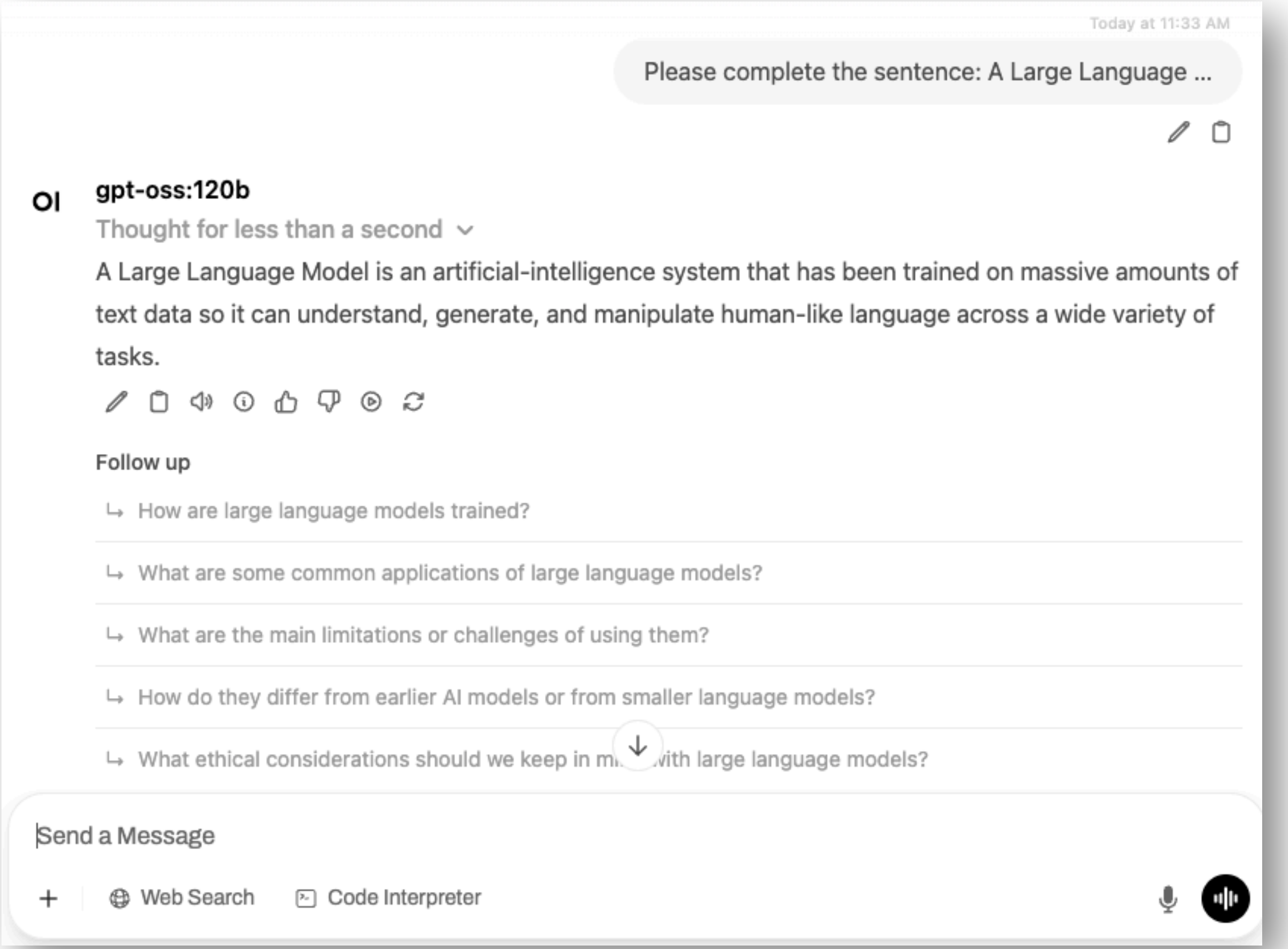
TODAY'S LECTURE



TODAY'S LECTURE

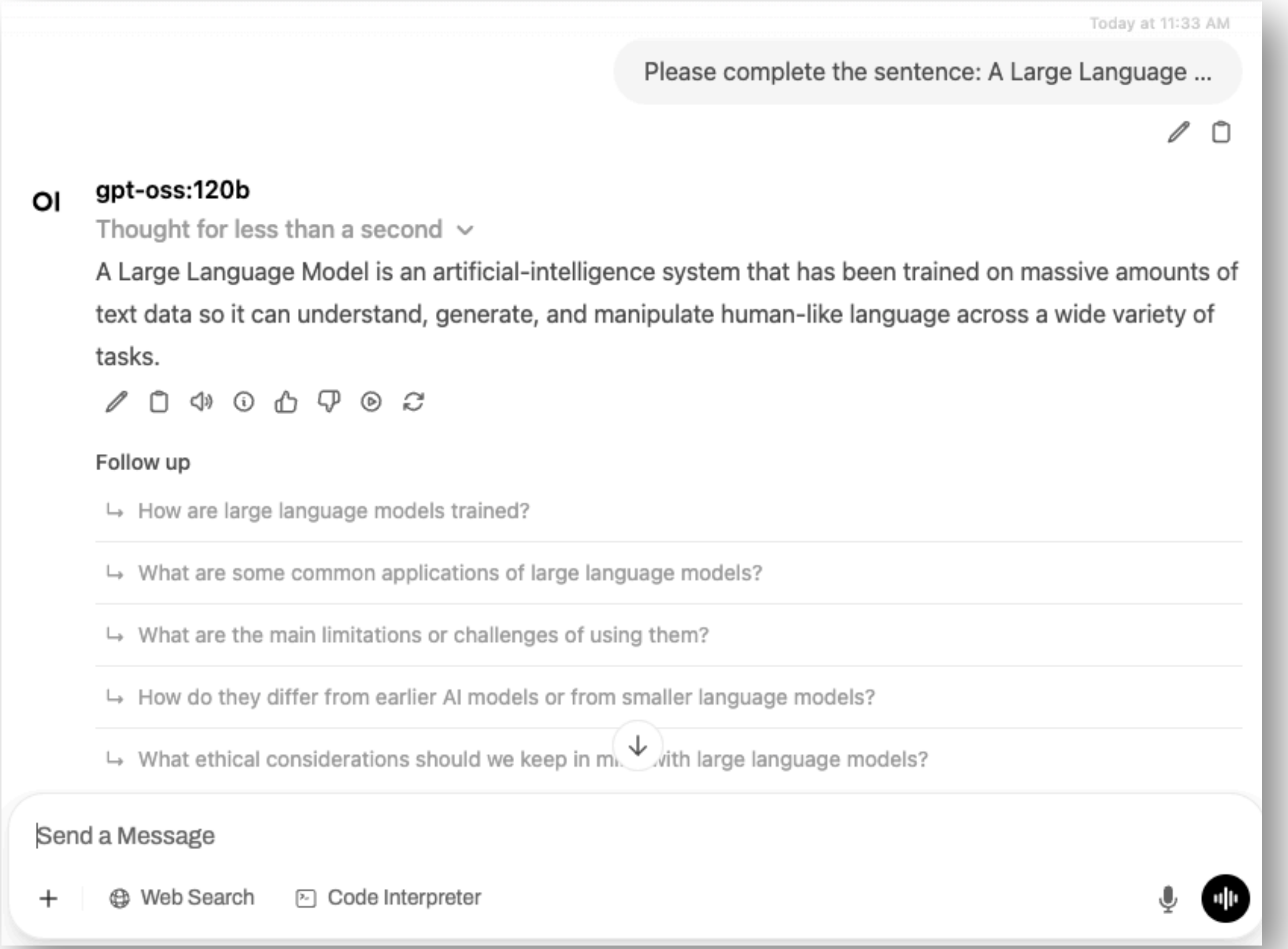
- Central question:

  „How to answer a prompt?"



TODAY’S LECTURE

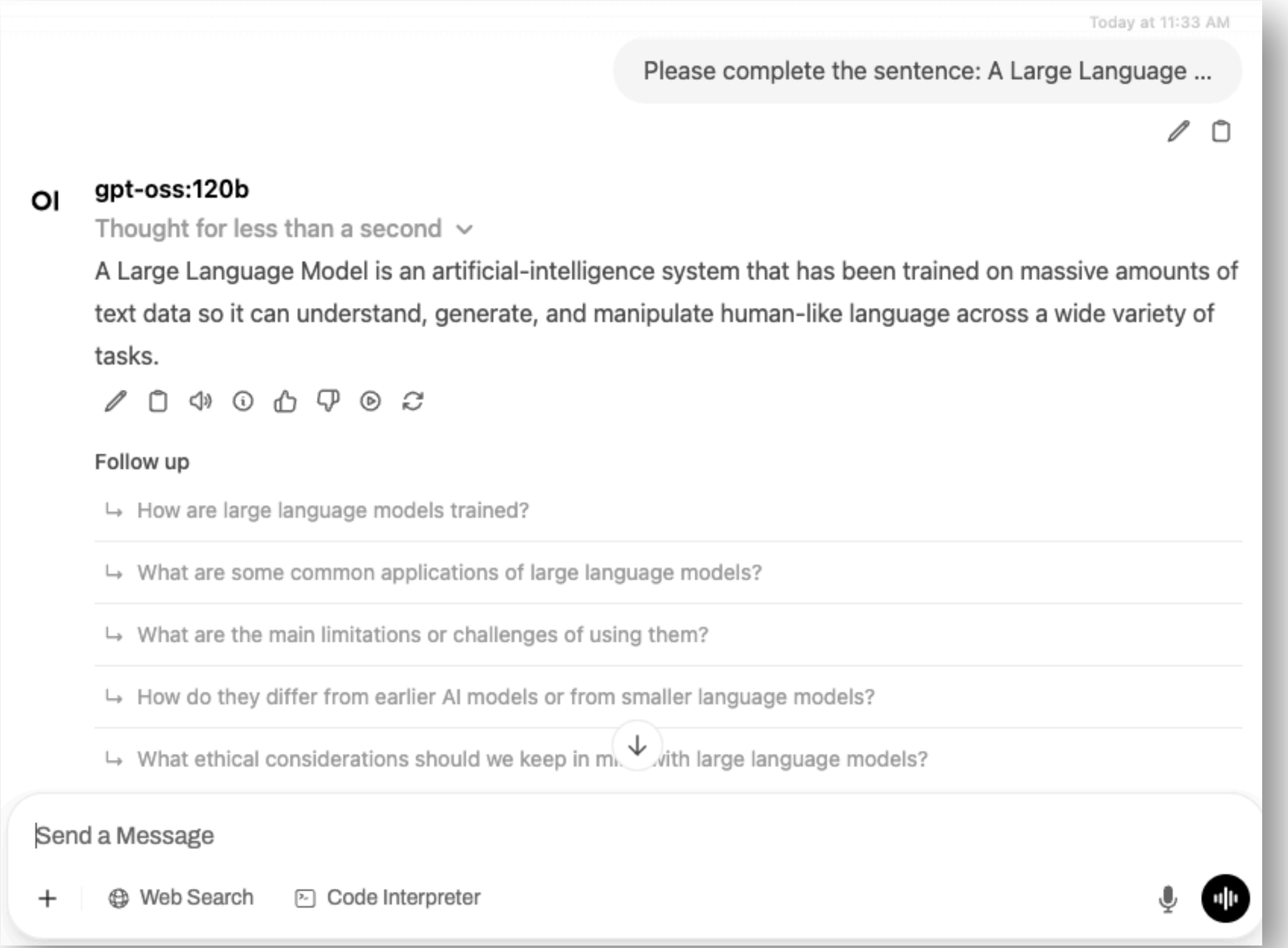
- Central question:
  „How to answer a prompt?“
- Steps and terms
  1. ChatBot
  2. Artificial Intelligence
  3. Large Language Model
  4. Machine Learning



TODAY’S LECTURE

- Central question:
  „How to answer a prompt?“
- Steps and terms
  1. ChatBot
  2. Artificial Intelligence
  3. Large Language Model
  4. Machine Learning

➡Important insights on capabilities and limits of AI



WHY TODAY'S LECTURE?

- Public discourse on LLMs and the future of AI is full of *bullshit* claims

`gpt-oss:120b` asked for„ (faulty) claims about AI“

WHY TODAY'S LECTURE?

- Public discourse on LLMs and the future of AI is full of *bullshit* claims

#	Faulty claim (presented as a quote)
1	“AI will replace all human workers within the next decade.”
2	“When an AI says it ‘feels’ something, it really experiences emotions.”
3	“A single AI system can understand any domain it reads about, just like a human expert.”
4	“If an AI is trained on a massive dataset, it becomes unbiased.”
5	“AI can predict the future with near-perfect accuracy.”
6	“All AI systems are self-learning and need no human supervision after deployment.”
7	“If an AI passes a Turing test, it is truly intelligent.”
8	“AI can replace doctors and make flawless medical diagnoses on its own.”
9	“AI will inevitably become hostile and try to ‘take over the world.’”

`gpt-oss:120b` asked for „(faulty) claims about AI“

WHY TODAY'S LECTURE?

- Public discourse on LLMs and the future of AI is full of *bullshit* claims
- Your future manager reads it and believes most of it. A true story…

#	Faulty claim (presented as a quote)
1	“AI will replace all human workers within the next decade.”
2	“When an AI says it ‘feels’ something, it really experiences emotions.”
3	“A single AI system can understand any domain it reads about, just like a human expert.”
4	“If an AI is trained on a massive dataset, it becomes unbiased.”
5	“AI can predict the future with near-perfect accuracy.”
6	“All AI systems are self-learning and need no human supervision after deployment.”
7	“If an AI passes a Turing test, it is truly intelligent.”
8	“AI can replace doctors and make flawless medical diagnoses on its own.”
9	“AI will inevitably become hostile and try to ‘take over the world.’”

`gpt-oss:120b` asked for „ (faulty) claims about AI“

WHY TODAY'S LECTURE?

- Public discourse on LLMs and the future of AI is full of *bullshit* claims
- Your future manager reads it and believes most of it. A true story…
- In your future jobs, you will have to help people understand what AI, LLMs, and Machine Learning **can** and **cannot** do

#	Faulty claim (presented as a quote)
1	“AI will replace all human workers within the next decade.”
2	“When an AI says it ‘feels’ something, it really experiences emotions.”
3	“A single AI system can understand any domain it reads about, just like a human expert.”
4	“If an AI is trained on a massive dataset, it becomes unbiased.”
5	“AI can predict the future with near-perfect accuracy.”
6	“All AI systems are self-learning and need no human supervision after deployment.”
7	“If an AI passes a Turing test, it is truly intelligent.”
8	“AI can replace doctors and make flawless medical diagnoses on its own.”
9	“AI will inevitably become hostile and try to ‘take over the world.’”

`gpt-oss:120b` asked for „(faulty) claims about AI“

WHY TODAY'S LECTURE?

- Public discourse on LLMs and the future of AI is full of *bullshit* claims
- Your future manager reads it and believes most of it. A true story…
- In your future jobs, you will have to help people understand what AI, LLMs, and Machine Learning **can** and **cannot** do
- Knowing how they work is the only way. We have to **demystify LLMs and the terms around it**.

#	Faulty claim (presented as a quote)
1	"AI will replace all human workers within the next decade."
2	"When an AI says it 'feels' something, it really experiences emotions."
3	"A single AI system can understand any domain it reads about, just like a human expert."
4	"If an AI is trained on a massive dataset, it becomes unbiased."
5	"AI can predict the future with near-perfect accuracy."
6	"All AI systems are self-learning and need no human supervision after deployment."
7	"If an AI passes a Turing test, it is truly intelligent."
8	"AI can replace doctors and make flawless medical diagnoses on its own."
9	"AI will inevitably become hostile and try to 'take over the world.'"

`gpt-oss:120b` asked for „(faulty) claims about AI"

BACK TO THE TOPIC

Respond to a prompt.

THE PROBLEM

THE PROBLEM

- Input

THE PROBLEM

- Input
  - Any type of (textual) prompt.
  - Freely written by a human

THE PROBLEM

- Input
  - Any type of (textual) prompt.
  - Freely written by a human
- Response

THE PROBLEM

- Input
  - Any type of (textual) prompt.
  - Freely written by a human
- Response
  - Any type of textual response
  - Automatically generated by the „system“

THE PROBLEM

- Input
  - Any type of (textual) prompt.
  - Freely written by a human
- Response
  - Any type of textual response
  - Automatically generated by the „system“
- Difficulties

Difficulties?

THE PROBLEM

Difficulties?

- Input
  - Any type of (textual) prompt.
  - Freely written by a human
- Response
  - Any type of textual response
  - Automatically generated by the „system“
- Difficulties
  - No prior knowledge about input, topic, …

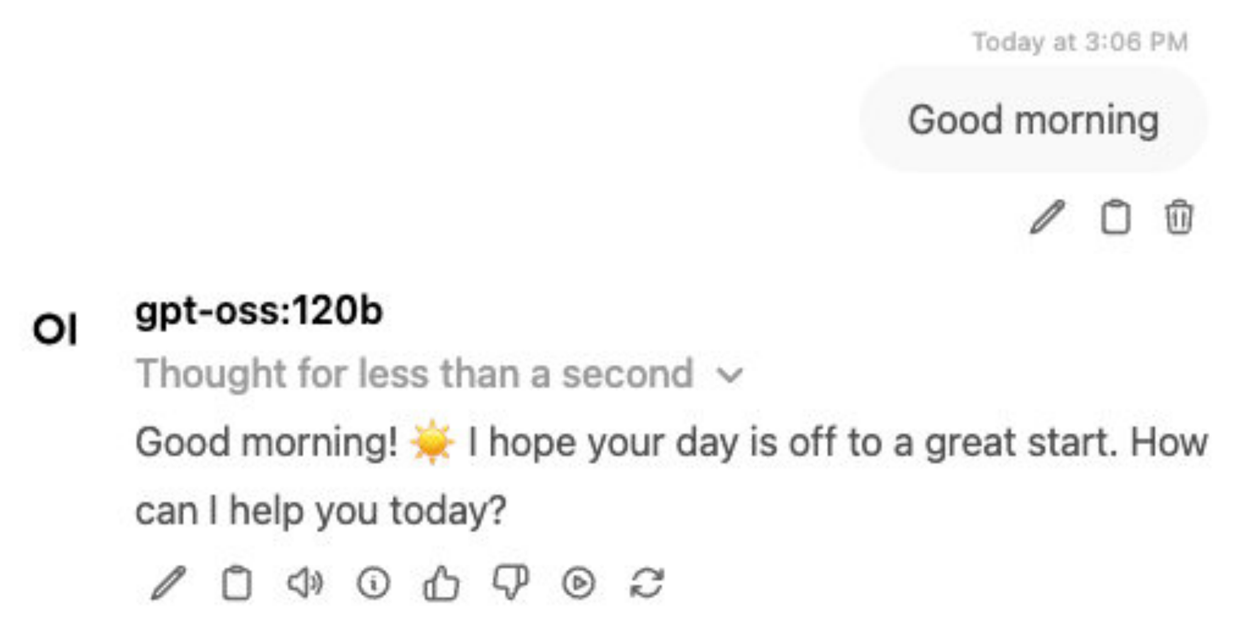
THE PROBLEM

Difficulties?

- Input
  - Any type of (textual) prompt.
  - Freely written by a human
- Response
  - Any type of textual response
  - Automatically generated by the „system“
- Difficulties
  - No prior knowledge about input, topic, …
  - *High quality* of response expected

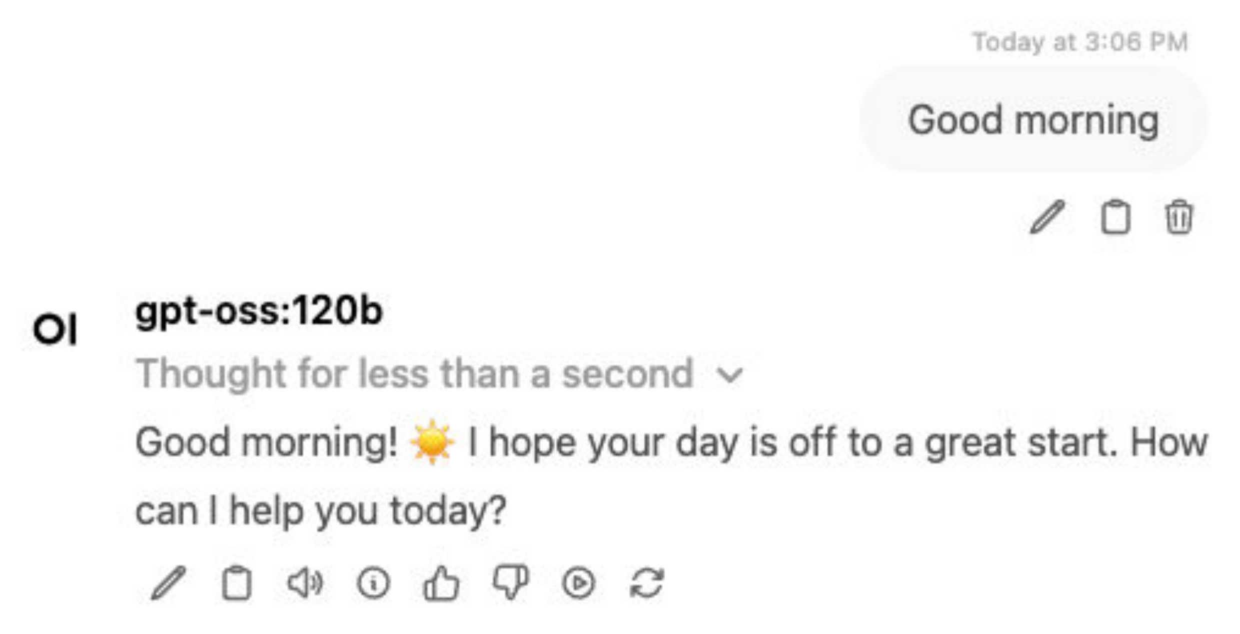
A FIRST TRY

A FIRST TRY


Today at 3:06 PM
Good morning
gpt-oss:120b
Thought for less than a second
Good morning! ☀️ I hope your day is off to a great start. How can I help you today?

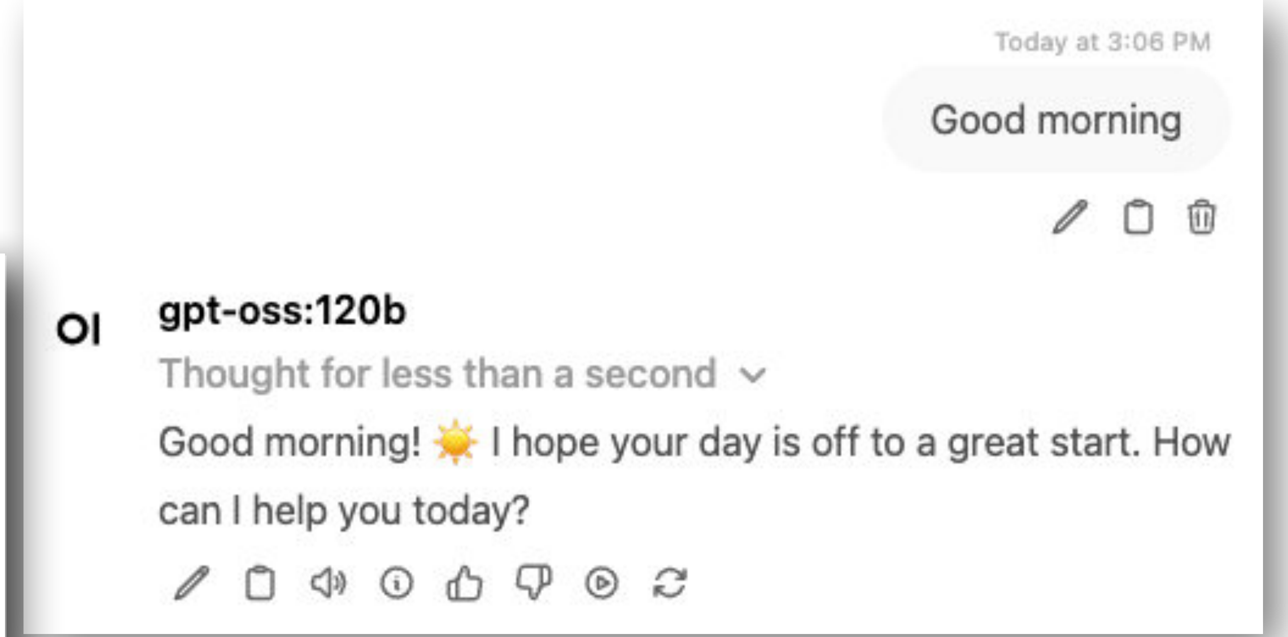
A FIRST TRY


Goddag
gpt-oss:120b Today at 7:20 PM
Thought for less than a second
Goddag! Hvordan kan jeg hjælpe dig i dag?

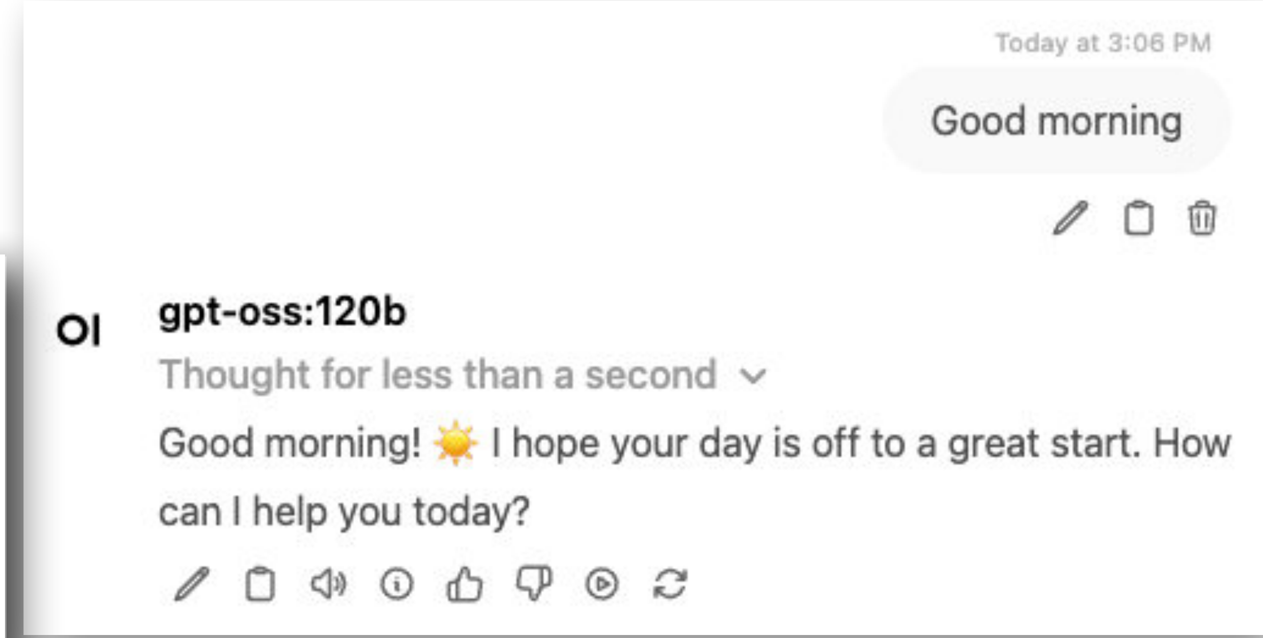

Today at 3:06 PM
Good morning
gpt-oss:120b
Thought for less than a second
Good morning! ☀️ I hope your day is off to a great start. How can I help you today?

A FIRST TRY


Goddag
gpt-oss:120b Today at 7:20 PM
Thought for less than a second
"User says "Goddag" which is Danish/Norwegian/Swedish for "Good day". Likely they are greeting. We respond in same language. Maybe ask how we can help. "
Goddag! Hvordan kan jeg hjælpe dig i dag?


Today at 3:06 PM
Good morning
gpt-oss:120b
Thought for less than a second
Good morning! ☀️ I hope your day is off to a great start. How can I help you today?

A FIRST TRY




```
if prompt == „Good morning“:
  respond(„Good morning! ☀️ I hope your
  day is off to a great start. How can I
  help you today?“)
else if prompt == „Goddag“:
  respond(„Goddag! Hvordan kan jeg hjælpe
  dig i dag?“)
else if …
```

ARTIFICIAL INTELLIGENCE

Arrange the terms: ChatBot, Artificial Intelligence, Large Language Model

CHATBOT

CHATBOT

- Responds to prompts

CHATBOT

- Responds to prompts
- More than just *generate text*

CHATBOT

Thought for 1 seconds

> "The user asks: "What is Aarhus BSS?" Likely they refer to Aarhus Business School? Actually Aarhus BSS stands for Aarhus Business School? Let's think: Aarhus BSS is part of Aarhus University, Denmark. BSS stands for Business and Social Sciences. It's a faculty within Aarhus University offering programs in business, economics, social sciences. So answer: it's a faculty, its history, programs, research, campuses, etc. Provide overview. Maybe also mention that BSS stands for Business and Social Sciences, includes Business School, Department of Business and Management, Department of Economics and Business Economics, etc. Also mention location, international orientation, rankings. Provide details. "

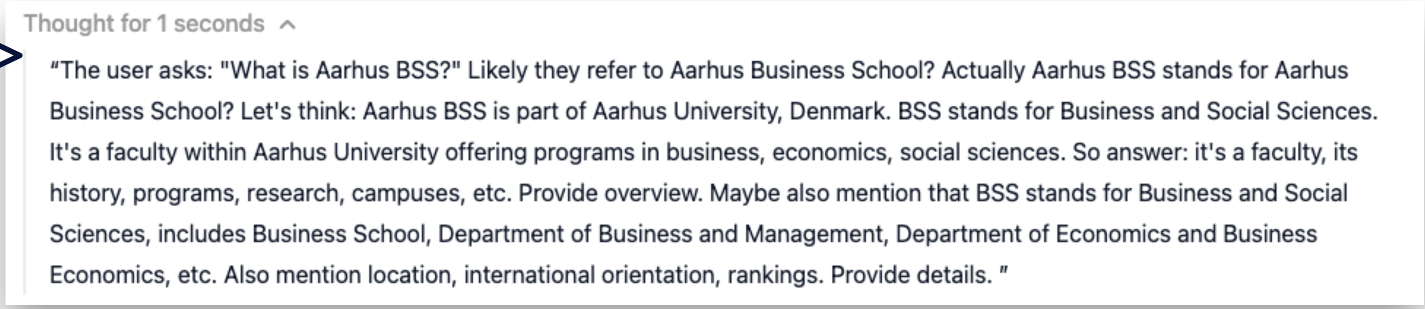
- Responds to prompts
- More than just *generate text*
    1. Analyzes prompt and decides about next step → „Thinking“

CHATBOT

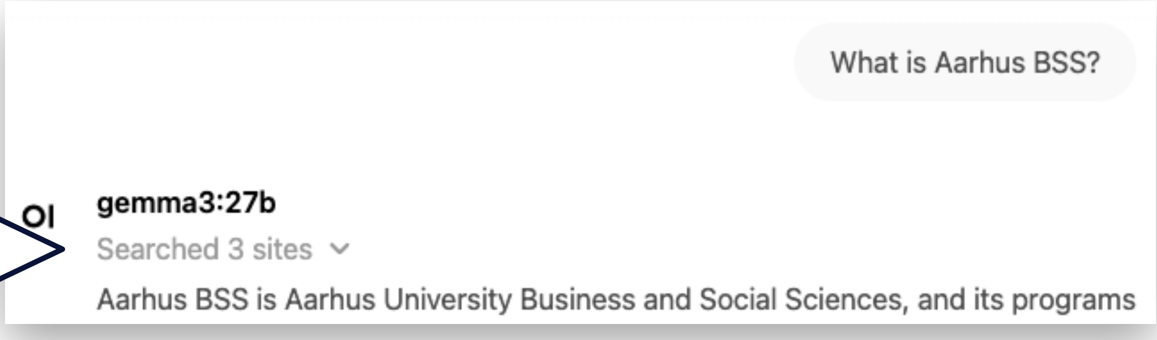
> Thought for 1 seconds
>
> “The user asks: "What is Aarhus BSS?" Likely they refer to Aarhus Business School? Actually Aarhus BSS stands for Aarhus Business School? Let's think: Aarhus BSS is part of Aarhus University, Denmark. BSS stands for Business and Social Sciences. It's a faculty within Aarhus University offering programs in business, economics, social sciences. So answer: it's a faculty, its history, programs, research, campuses, etc. Provide overview. Maybe also mention that BSS stands for Business and Social Sciences, includes Business School, Department of Business and Management, Department of Economics and Business Economics, etc. Also mention location, international orientation, rankings. Provide details. ”

- Responds to prompts
- More than just *generate text*
    1. Analyzes prompt and decides about next step → „Thinking“
    2. Possibly does a web search

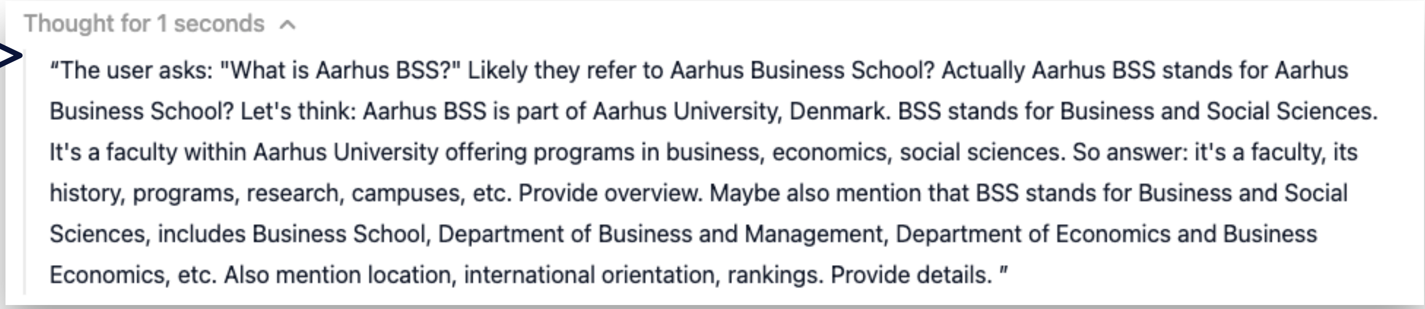
CHATBOT



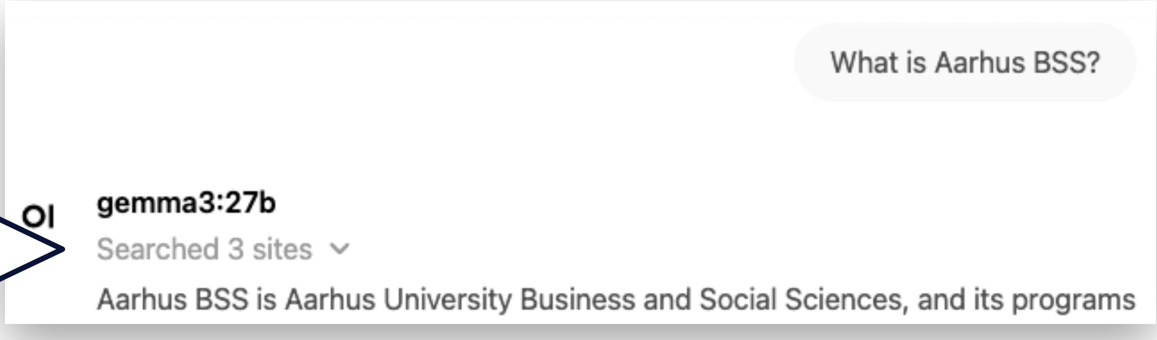
- Responds to prompts
- More than just *generate text*
  1. Analyzes prompt and decides about next step → „Thinking“
  2. Possibly does a web search
  3. Analyzes results of web search



CHATBOT

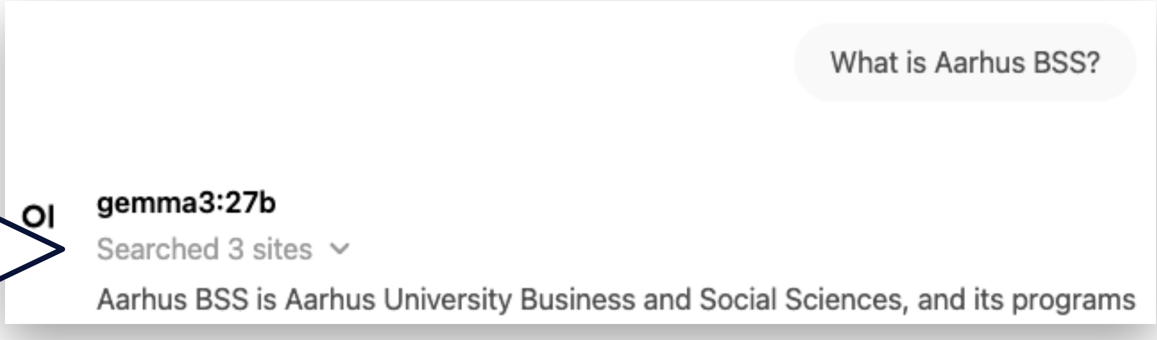


- Responds to prompts
- More than just *generate text*
    1. Analyzes prompt and decides about next step → „Thinking"
    2. Possibly does a web search
    3. Analyzes results of web search
    4. Generates response



CHATBOT

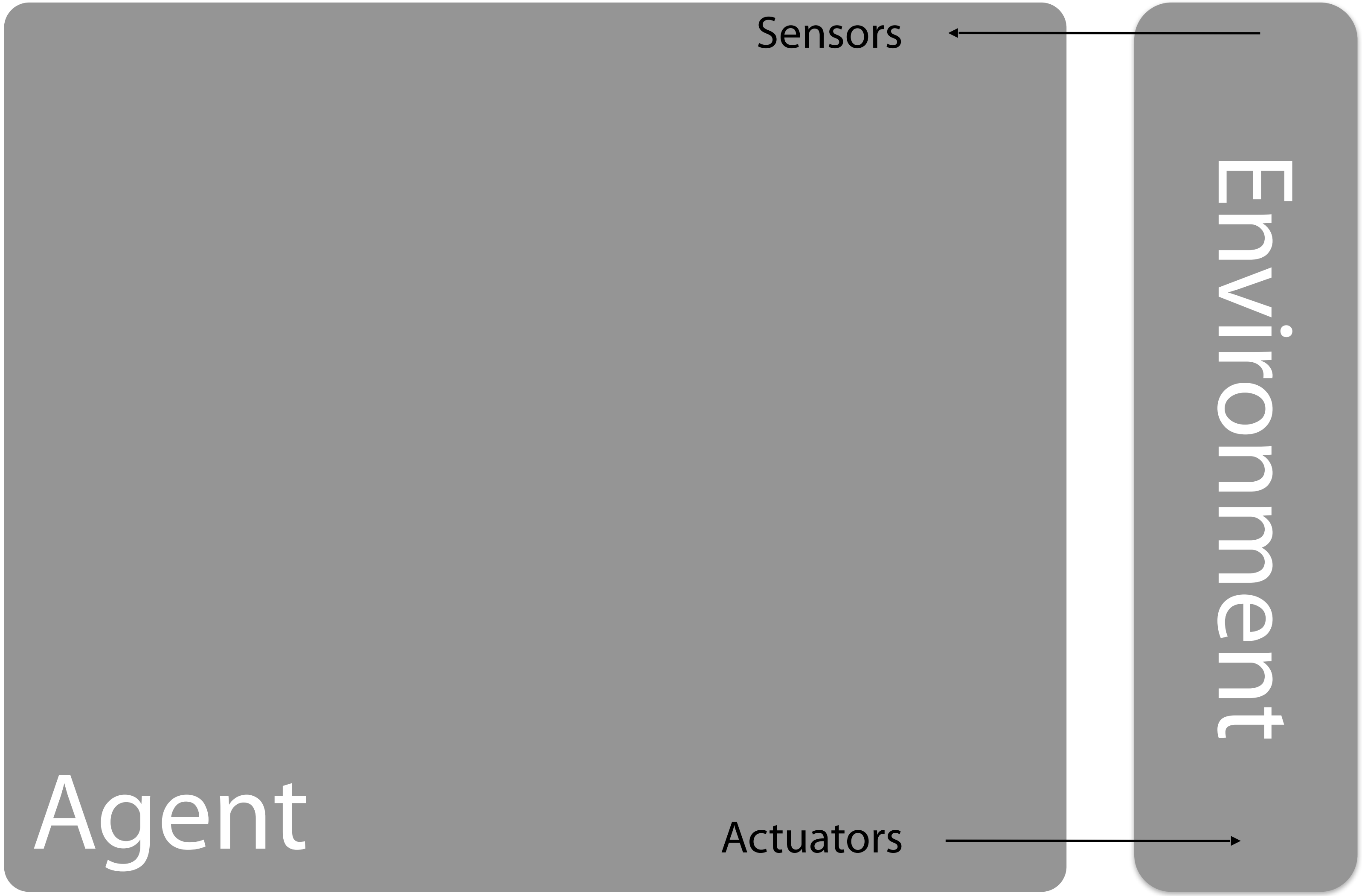
- Responds to prompts
- More than just *generate text*
    1. Analyzes prompt and decides about next step → „Thinking“
    2. Possibly does a web search
    3. Analyzes results of web search
    4. Generates response
- Backend by a Large Language Model, e.g., GPT, LLama, Gemma, …



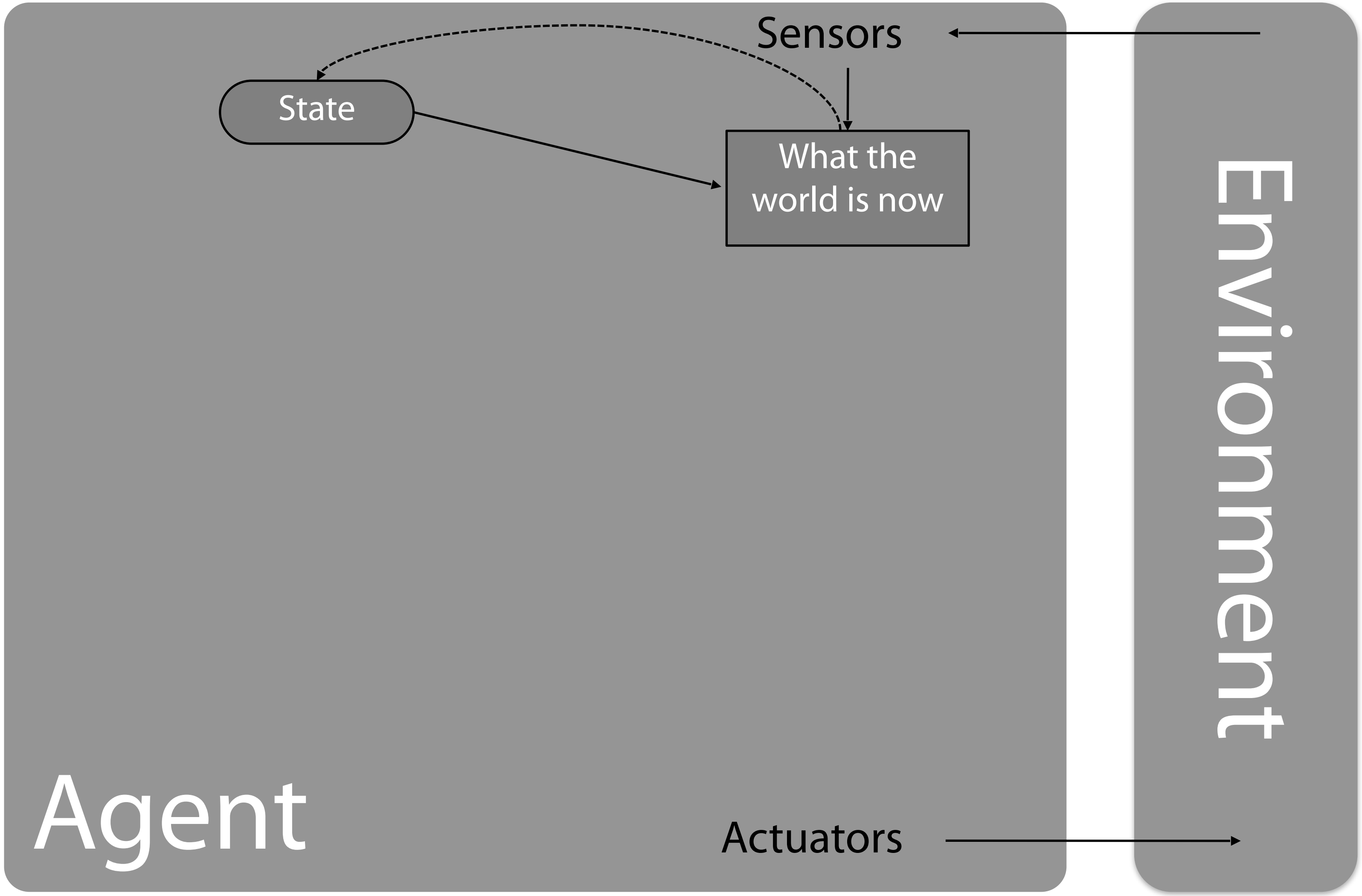
ChatBot	Models
ChatGPT	• GPT 5 • GPT 4o • …
Gemini	• 2.5 Flash • 2.5 Pro • …

ARTIFICIAL INTELLIGENT AGENTS

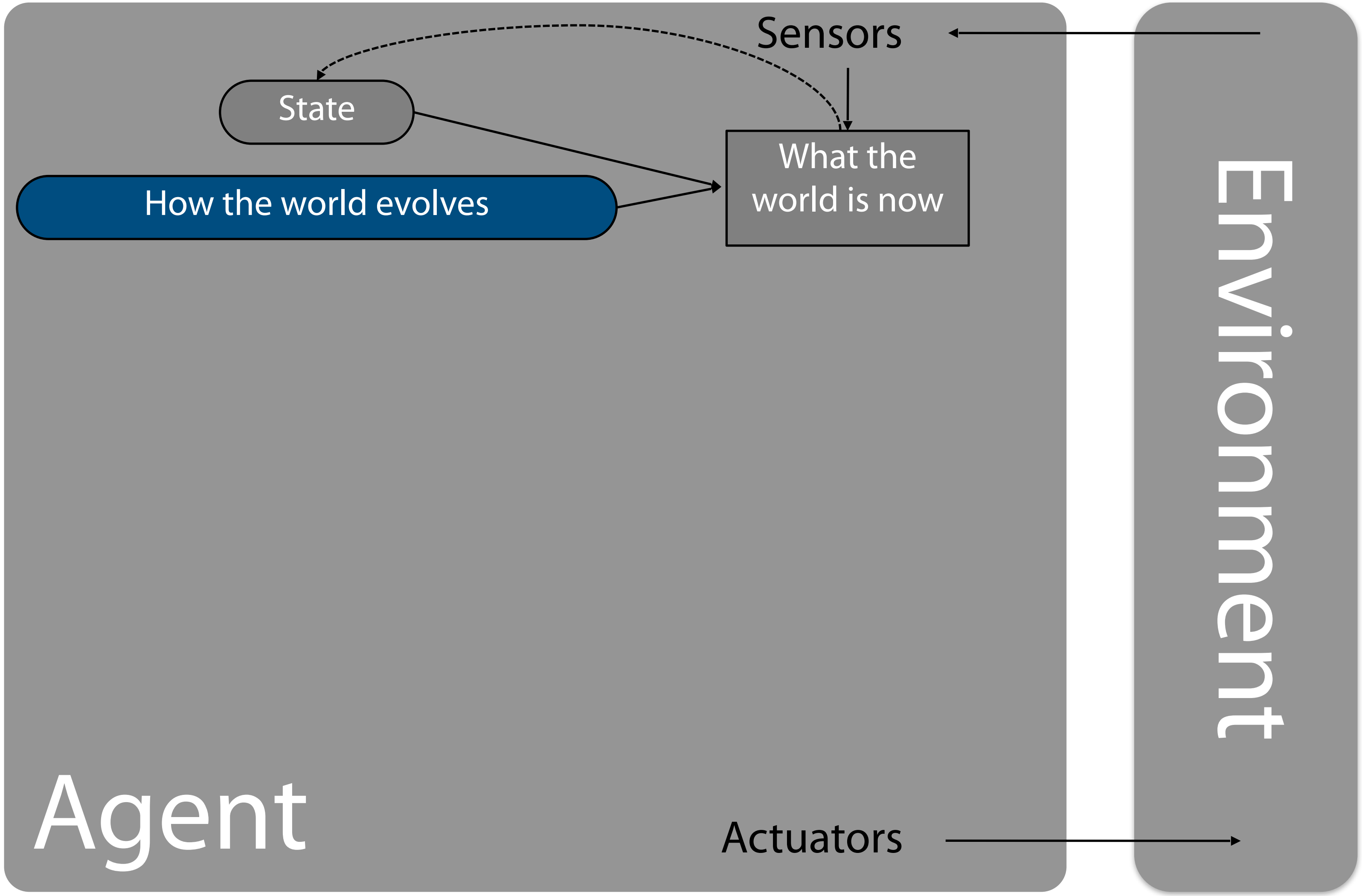
ARTIFICIAL INTELLIGENT AGENTS



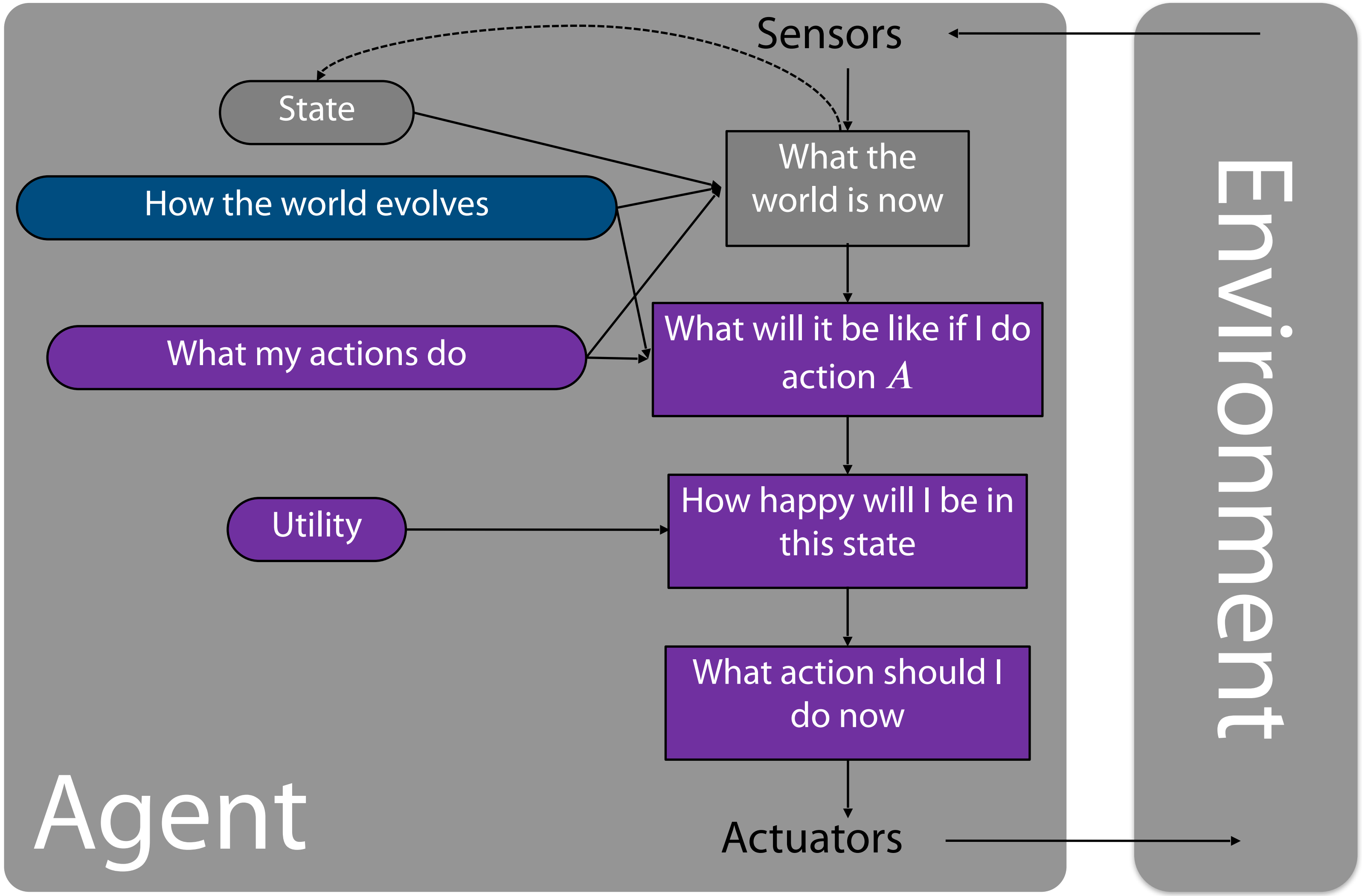
ARTIFICIAL INTELLIGENT AGENTS



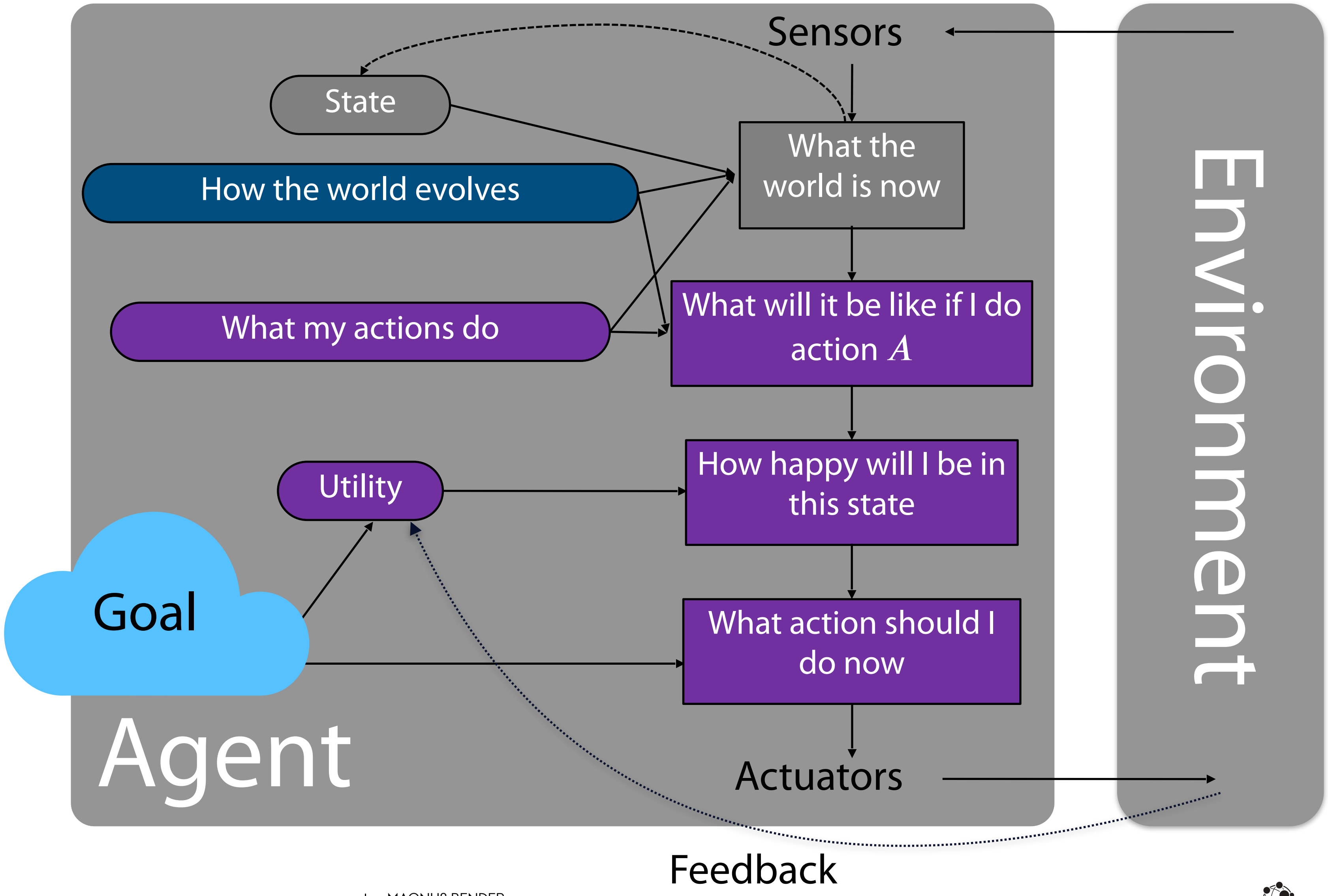
ARTIFICIAL INTELLIGENT AGENTS



ARTIFICIAL INTELLIGENT AGENTS

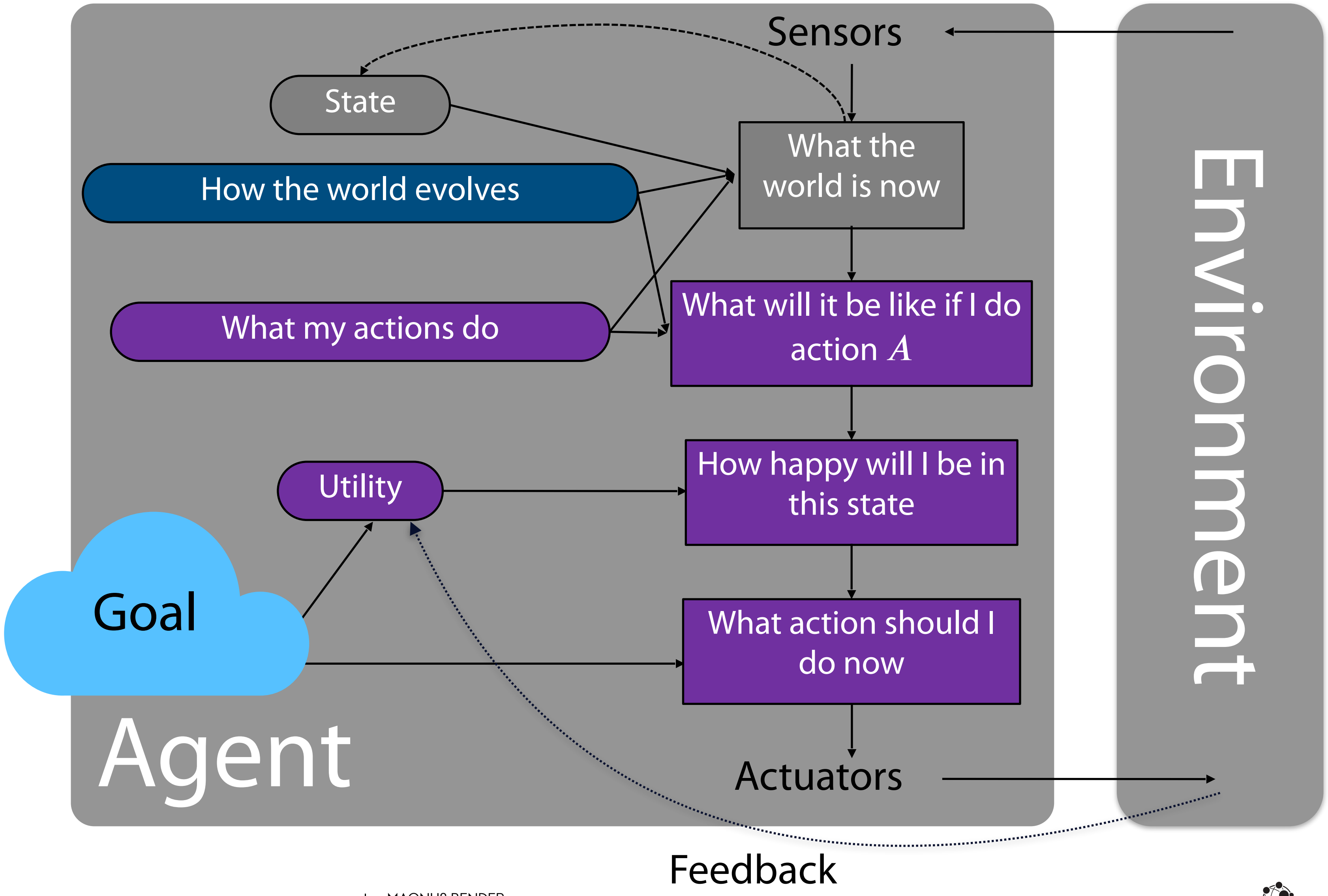


ARTIFICIAL INTELLIGENT AGENTS



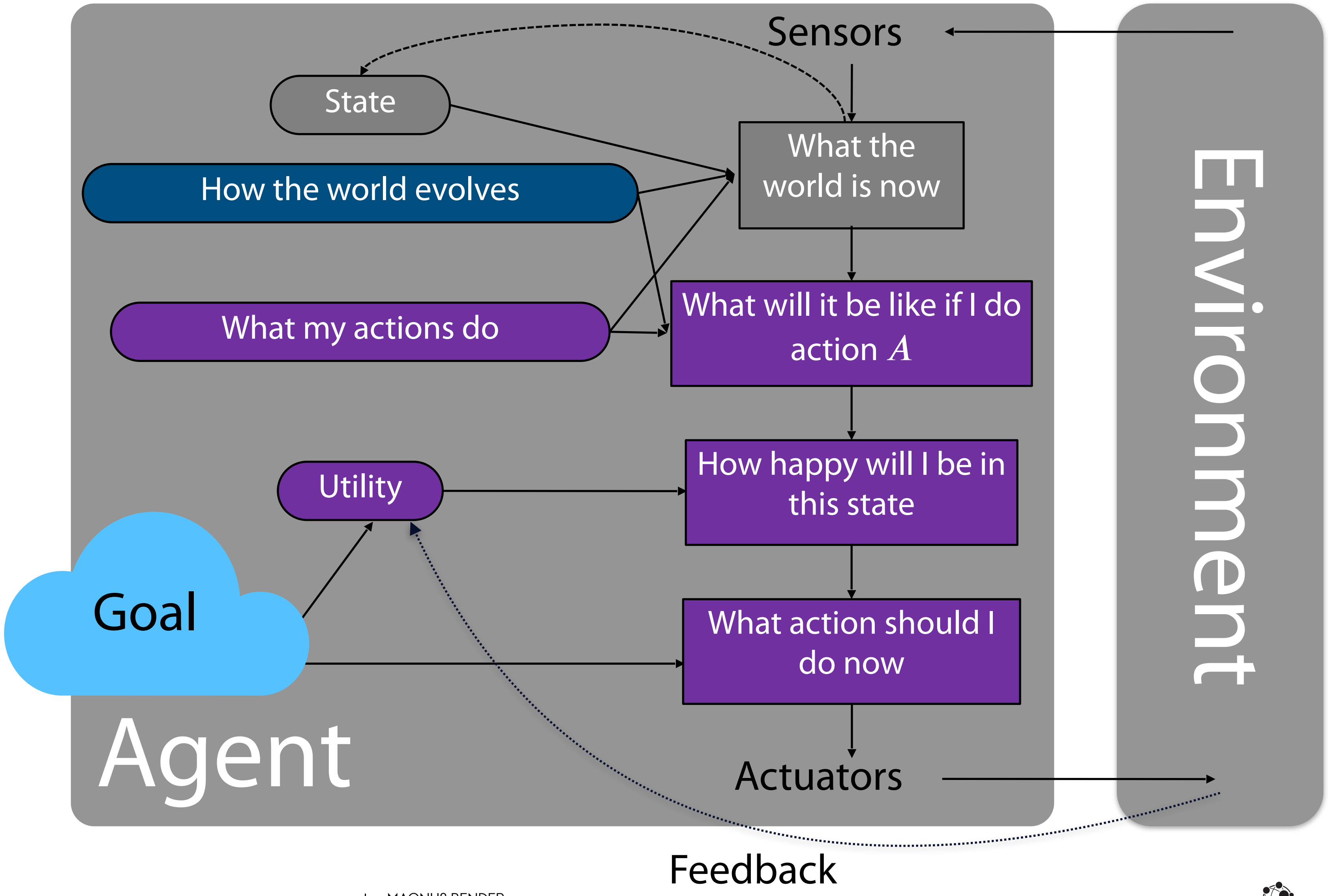
ARTIFICIAL INTELLIGENT AGENTS

- Intelligent systems, but not necessarily *intelligent* in a human sense



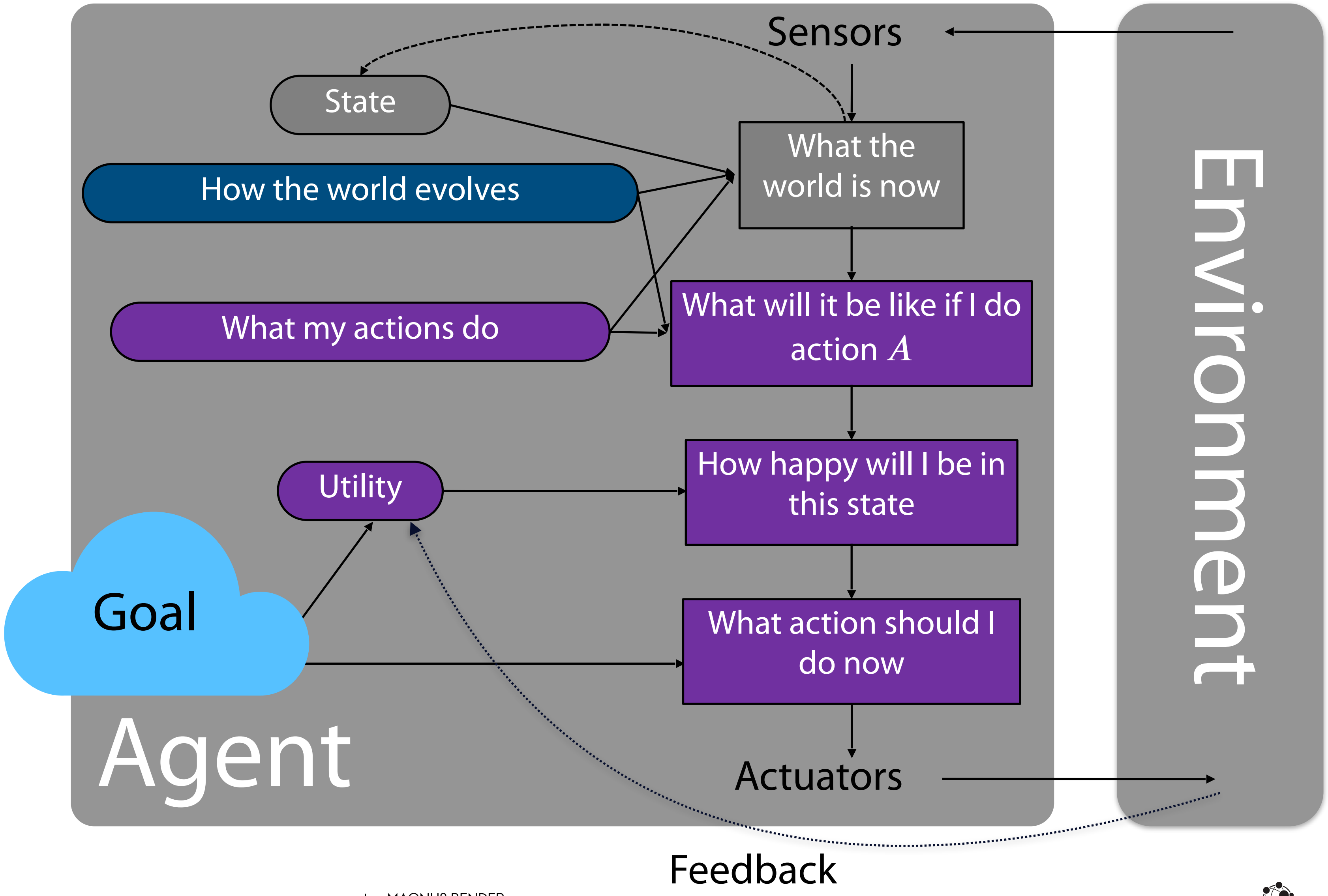
ARTIFICIAL INTELLIGENT AGENTS

- Intelligent systems, but not necessarily *intelligent* in a human sense
- Agents



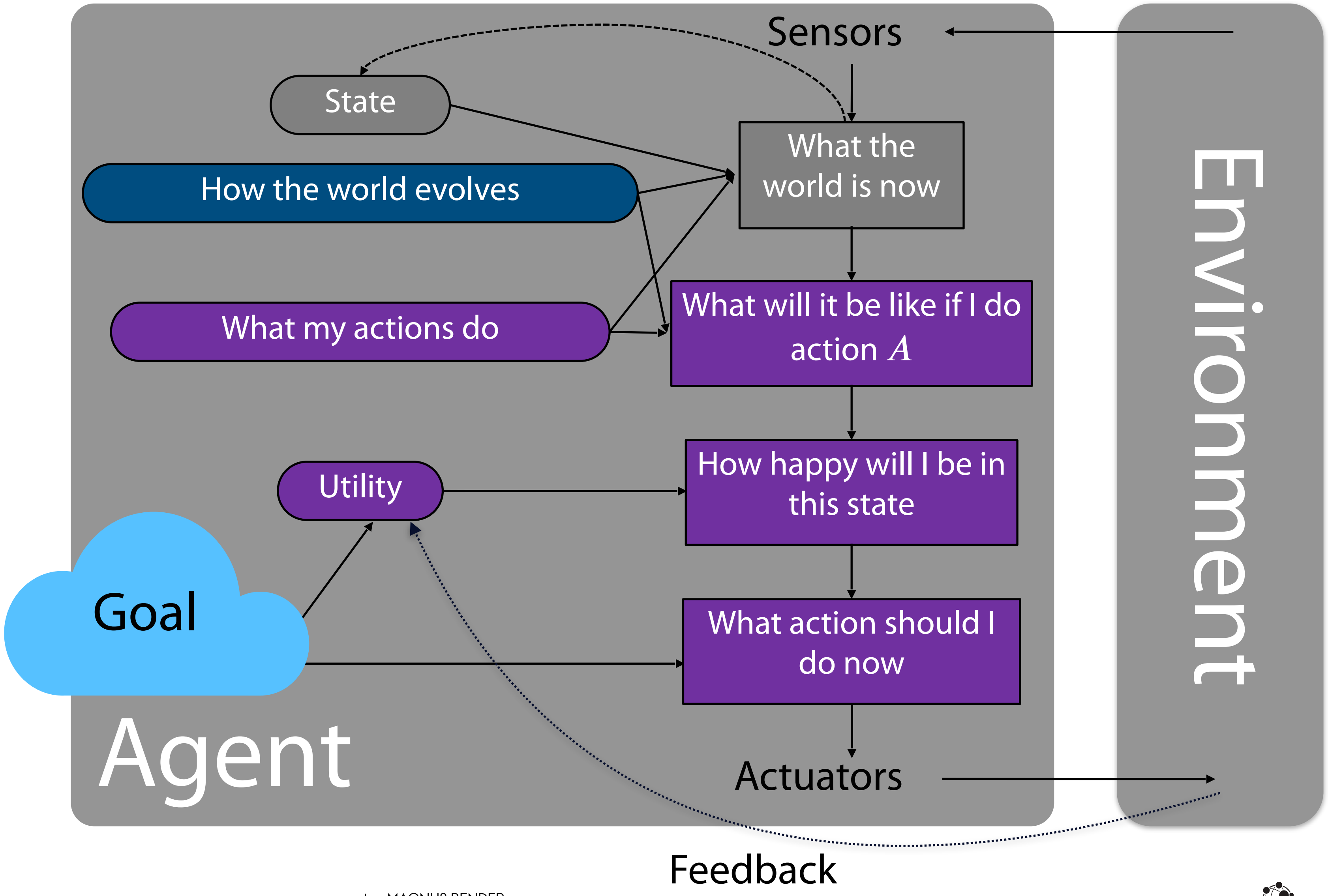
ARTIFICIAL INTELLIGENT AGENTS

- Intelligent systems, but not necessarily *intelligent* in a human sense
- Agents
  - ... have goals



ARTIFICIAL INTELLIGENT AGENTS

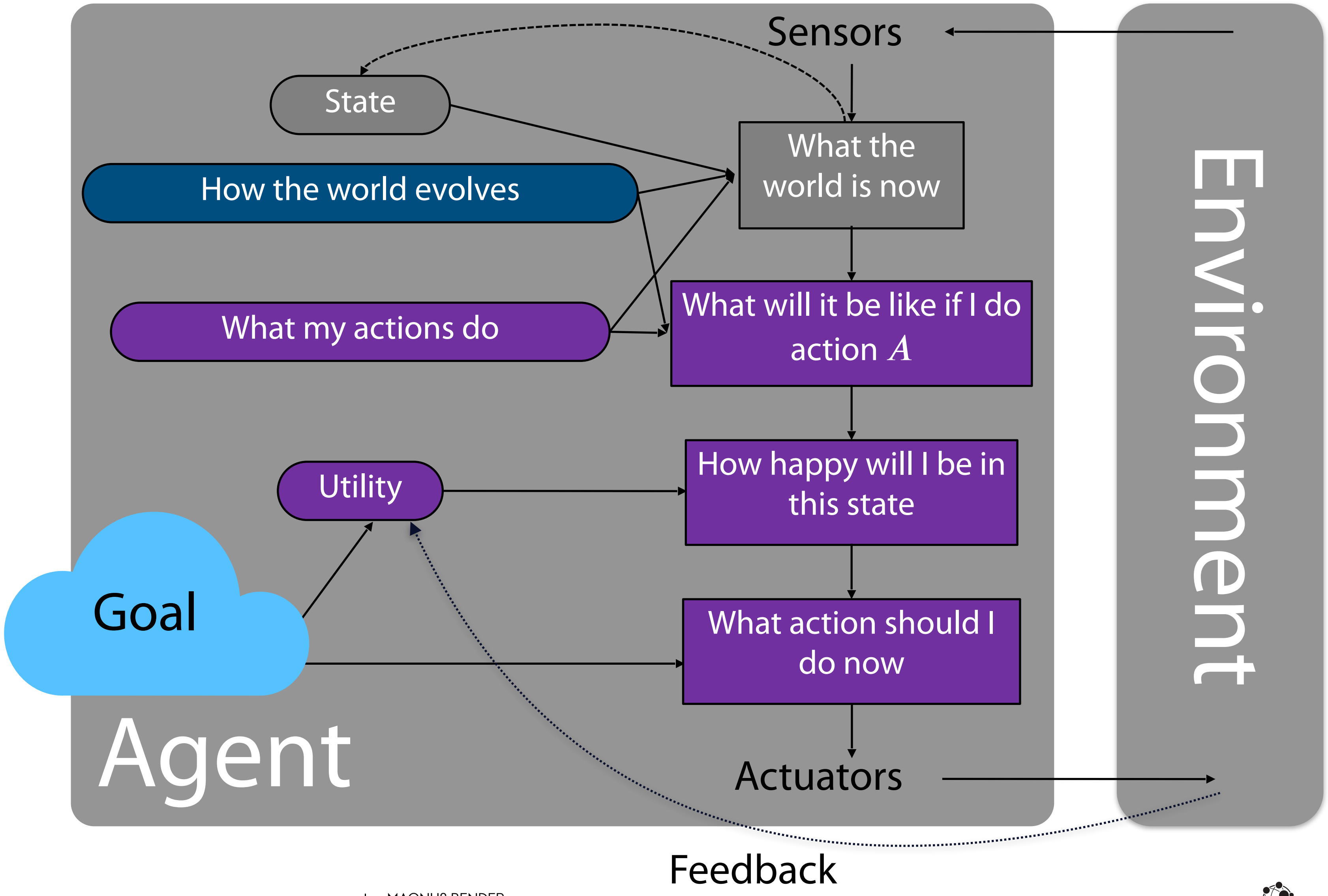
- Intelligent systems, but not necessarily *intelligent* in a human sense
- Agents
  - ... have goals
  - ... have a perception of their environment (sensors)



AARHUS BSS
DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

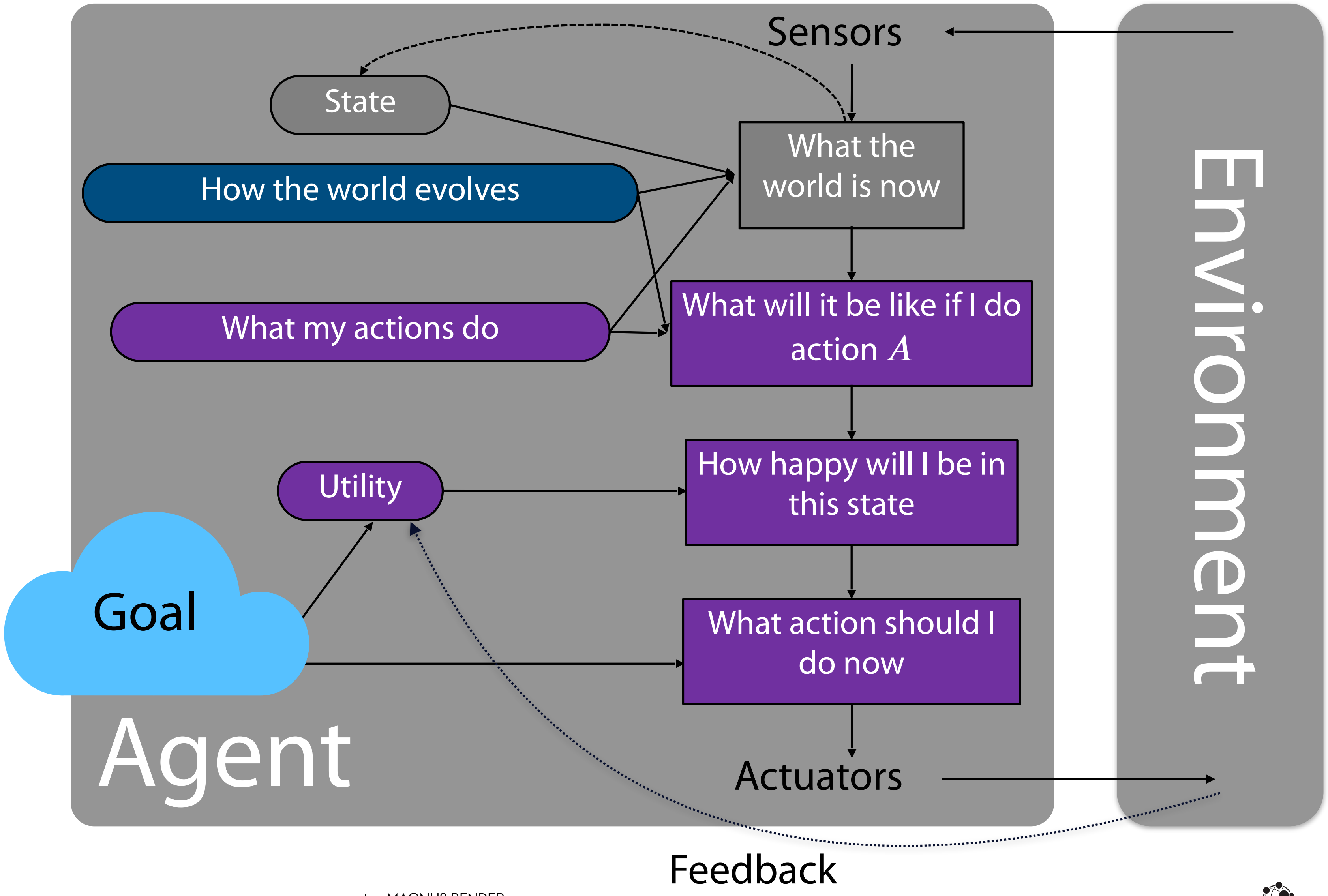
ARTIFICIAL INTELLIGENT AGENTS

- Intelligent systems, but not necessarily *intelligent* in a human sense
- Agents
  - … have goals
  - … have a perception of their environment (sensors)
  - … can change their environment (actuators)



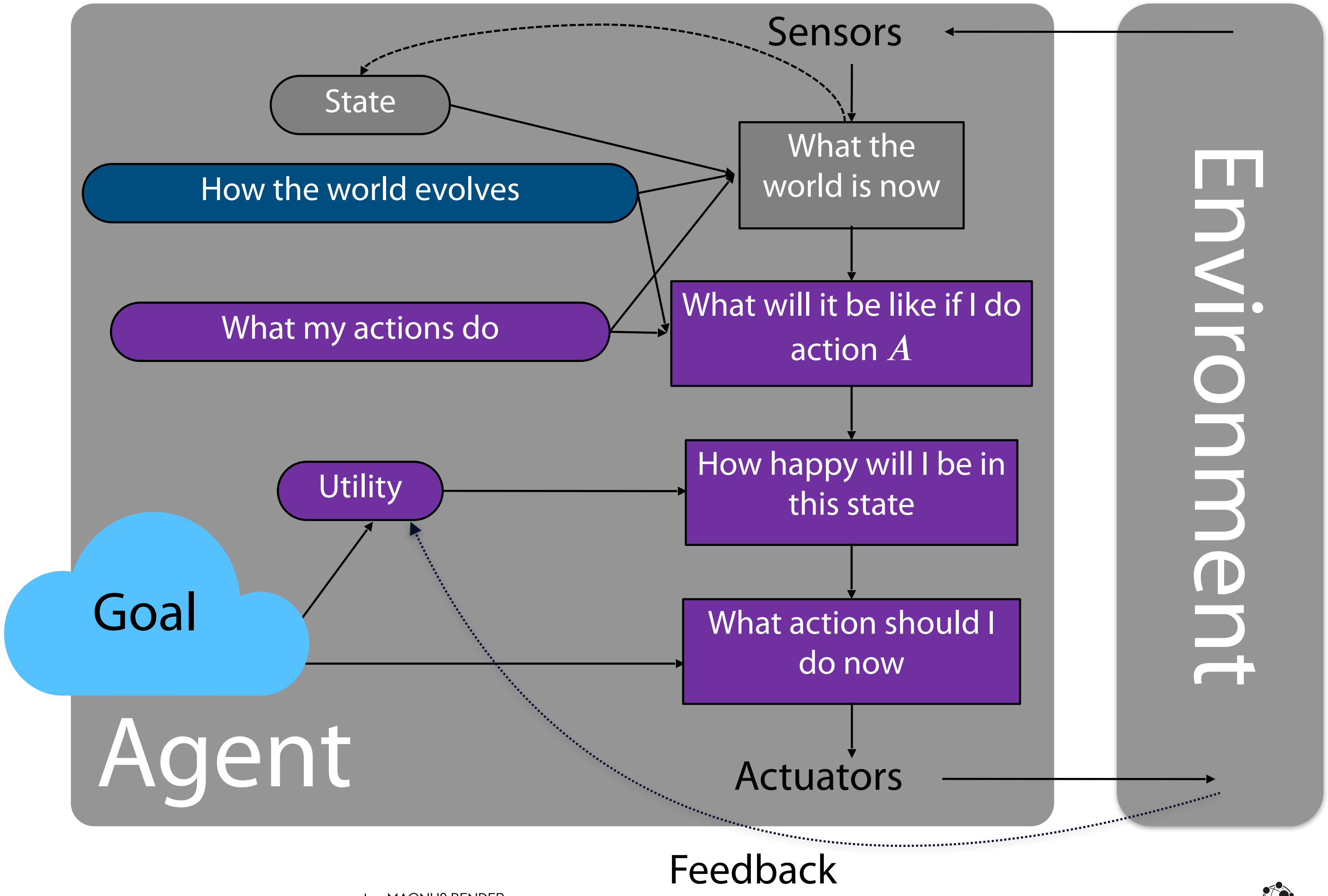
ARTIFICIAL INTELLIGENT AGENTS

- Intelligent systems, but not necessarily *intelligent* in a human sense
- Agents
  - … have goals
  - … have a perception of their environment (sensors)
  - … can change their environment (actuators)
  - … plan their actions



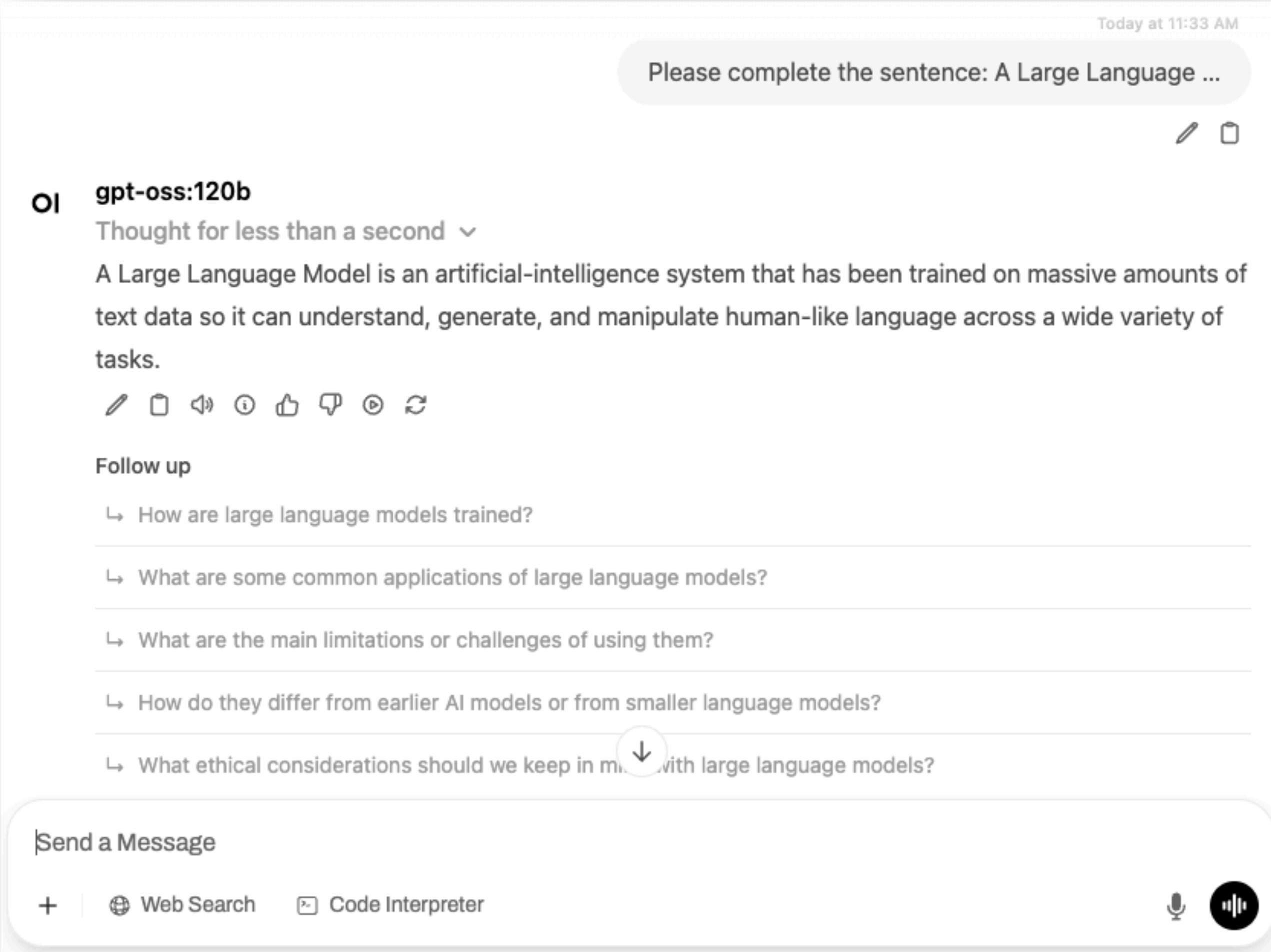
ARTIFICIAL INTELLIGENT AGENTS

- Intelligent systems, but not necessarily *intelligent* in a human sense
- Agents
  - ... have goals
  - ... have a perception of their environment (sensors)
  - ... can change their environment (actuators)
  - ... plan their actions
  - ... update theirs goals → learn during runtime



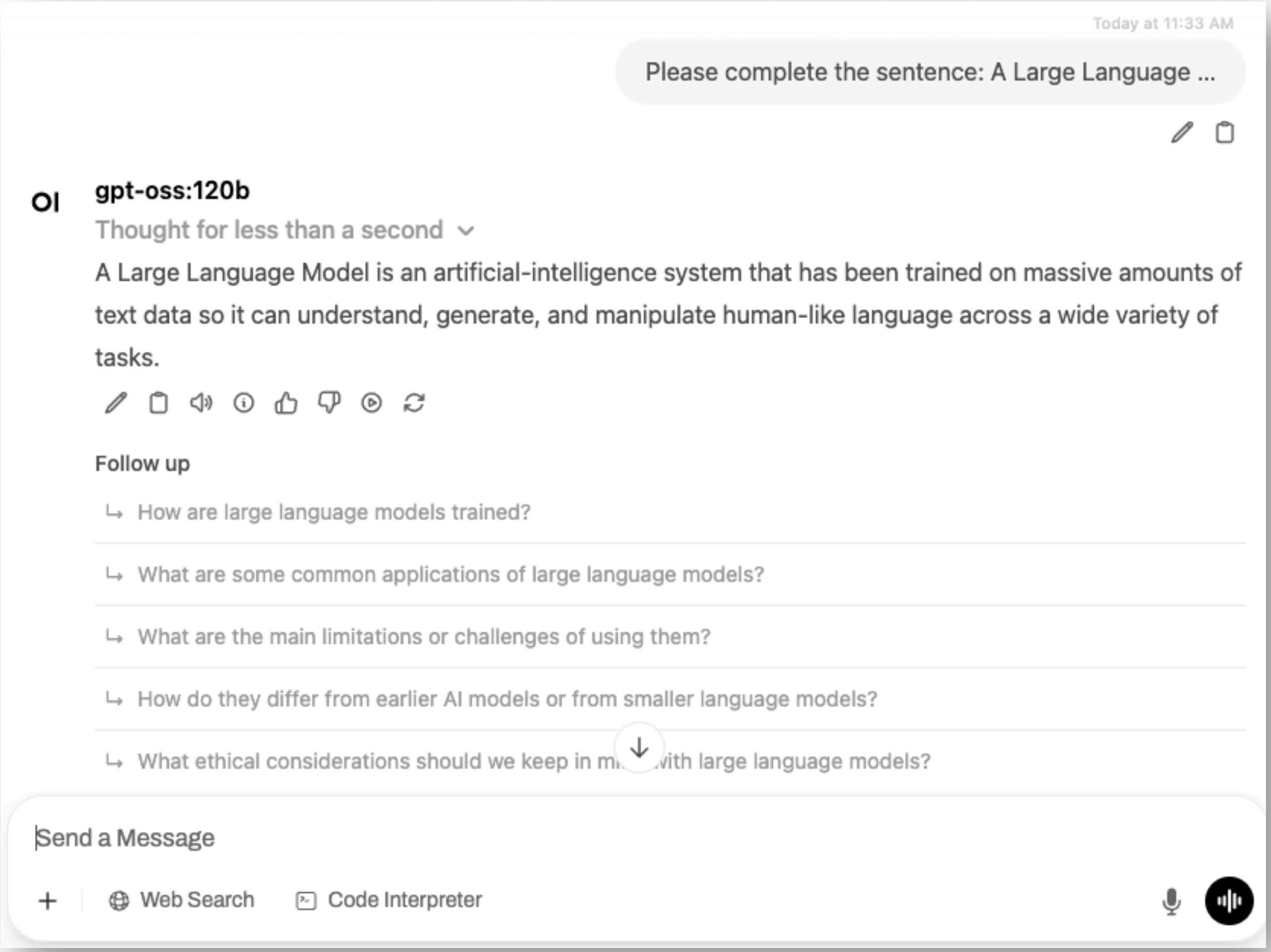
HOW TO ANSWER A PROMPT?

- Prompt
  - „Please complete the sentence: A Large Language ...“



HOW TO ANSWER A PROMPT?

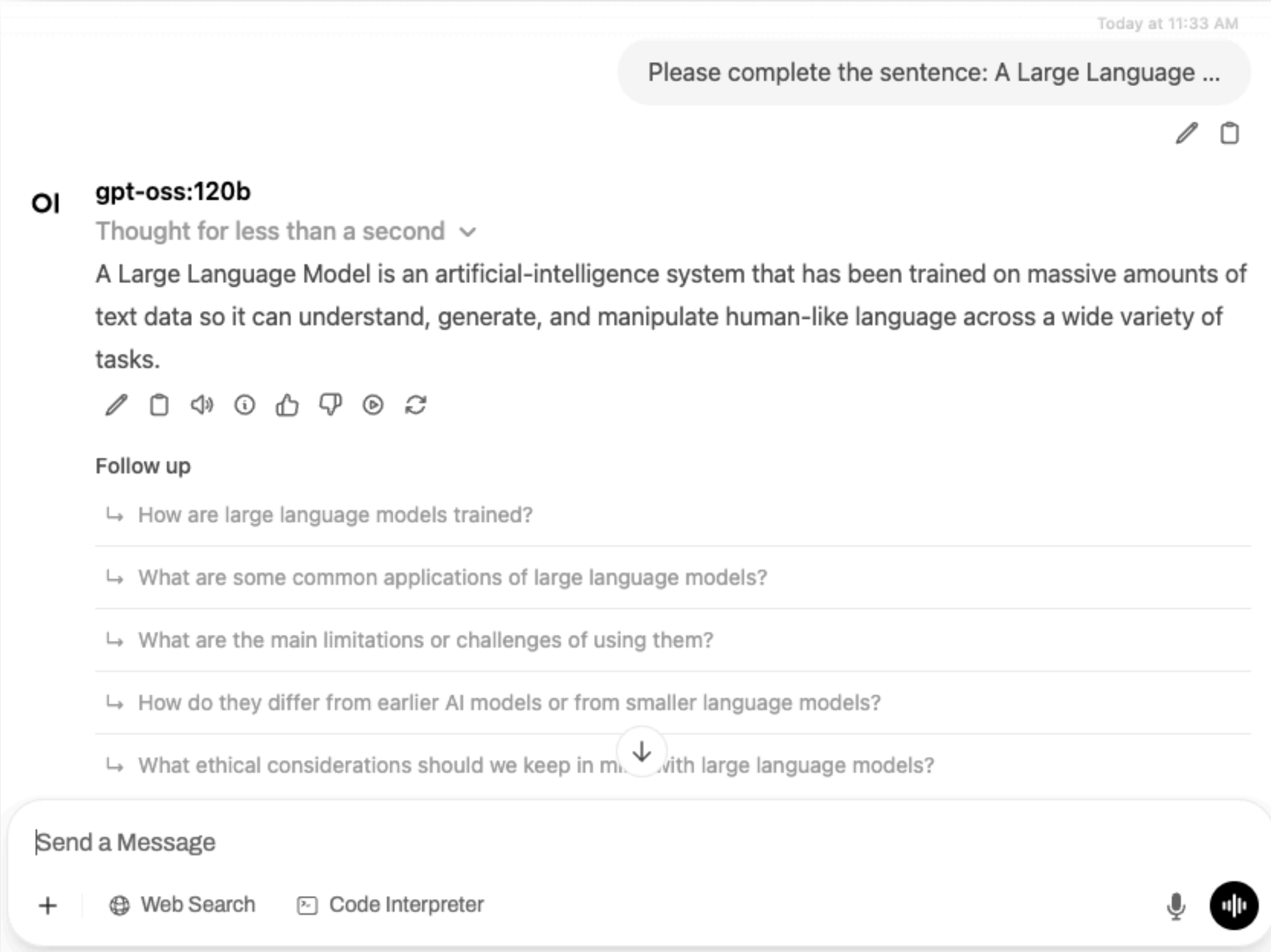
- Prompt
  - „Please complete the sentence: A Large Language …“
- Consists of two parts:
  1. Task: „Please complete the sentence:“
  2. Data: „A Large Language …“



HOW TO ANSWER A PROMPT?

- Prompt
  - „Please complete the sentence: A Large Language …“
- Consists of two parts:
  1. Task: „Please complete the sentence:“
  2. Data: „A Large Language …“

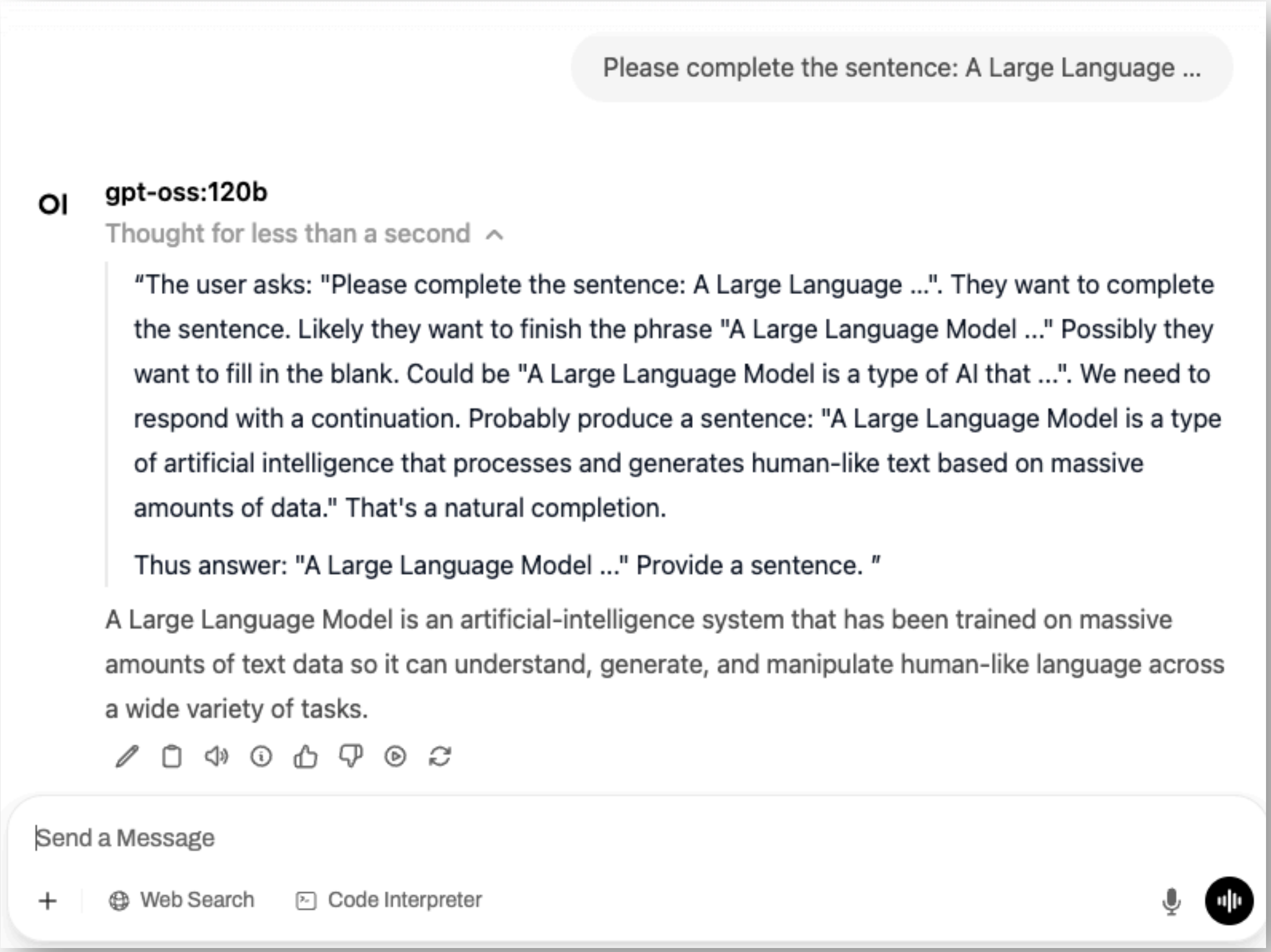
➡ Analysis of prompt to identify task and execution of task on data can be done with the help of an LLM



HOW TO ANSWER A PROMPT?

- Prompt
  - „Please complete the sentence: A Large Language …“
- Consists of two parts:
  1. Task: „Please complete the sentence:“
  2. Data: „A Large Language …“

➡ Analysis of prompt to identify task and execution of task on data can be done with the help of an LLM



LARGE LANGUAGE MODELS

Hands on the field of Natural Language Processing

DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

10 SEPTEMBER 2025

MAGNUS BENDER
ASSISTANT PROFESSOR

PREDICT *REASONABLE* WORDS

PREDICT *REASONABLE* WORDS

Let's guess words!

„The red ____ rolled down the hill.“

PREDICT *REASONABLE* WORDS

Let’s guess words!

„The red ____ rolled down the hill.“

Menti „Word Cloud“

PREDICT *REASONABLE* WORDS

Let's guess words!

„The red ____ rolled down the hill.“

- What do we infer from context?
  - It must be a noun, referencing a rollable object
  - It is probably something that is typical of this situation.
- What could ___ be, but is probably not??

Menti „Word Cloud“

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let’s guess words again! (heads up: more tricky!)

„The cat drank the milk because it was ___.“

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let’s guess words again! (heads up: more tricky!)

„The cat drank the milk because it was ___.“

Menti „Word Cloud“

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let’s guess words again! (heads up: more tricky!)

„The cat drank the milk because it was ___.“

- Popular choices:
  - „hungry“, „thirsty“
  - „delicious“, „cold“

Menti „Word Cloud“

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let’s guess words again! (heads up: more tricky!)

„The cat drank the milk because it was ___.“

- Popular choices:
  - „hungry“, „thirsty“
  - „delicious“, „cold“
- The definite article „it“ can refer to:
  - „the cat“
  - „the milk“

Menti „Word Cloud“

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let’s guess words again! (heads up: more tricky!)

„The cat drank the milk because it was ___.“

- Popular choices:
  - „hungry“, „thirsty“
  - „delicious“, „cold“
- The definite article „it“ can refer to:
  - „the cat“
  - „the milk“

Menti „Word Cloud“

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let’s guess words! (heads up: more tricky!)

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let's guess words! (heads up: more tricky!)

Menti
„Ranking“

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let's guess words! (heads up: more tricky!)

The **cat** drank the milk because it was **hungry**, because it had been …

The cat drank the **milk** because it was **delicious**, because it had been …

Menti
„Ranking"

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let's guess words! (heads up: more tricky!)

The **cat** drank the milk because it was **hungry**, because it had been ...

... out hunting all night.

The cat drank the **milk** because it was **delicious**, because it had been ...

... in the fridge all night.

Menti „Ranking“

WHY ARE LLMS SO GOOD? "SELF ATTENTION"

Let's guess words! (heads up: more tricky!)

The **cat** drank the milk because it was **hungry**, because it had been …

… out hunting all night.

➡ Cause for cat being hungry

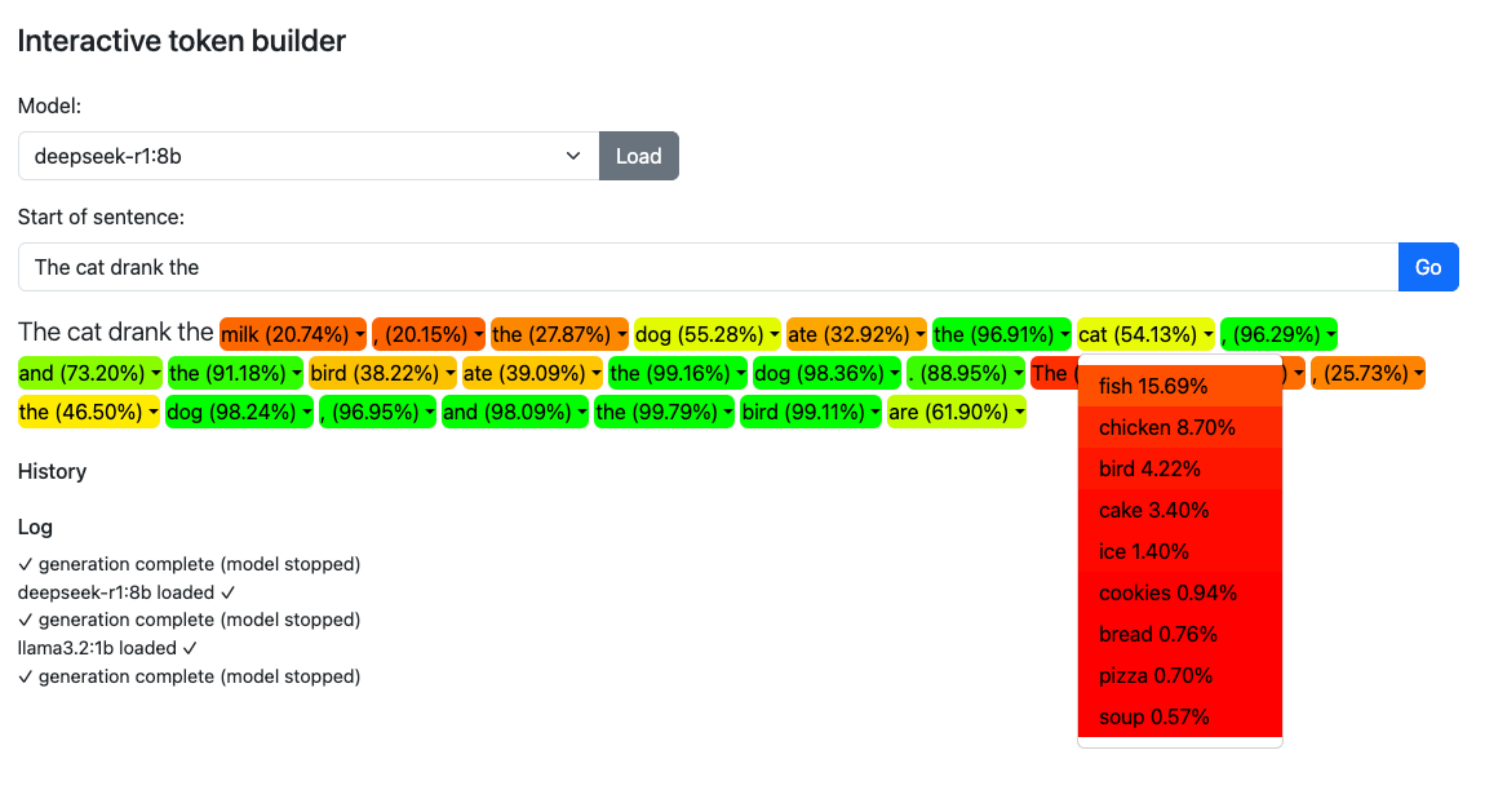
The cat drank the **milk** because it was **delicious**, because it had been …

… in the fridge all night.

➡ Cause for milk being delicious.

Menti
„Ranking"

DEMO: COMPLETE THE SENTENCE

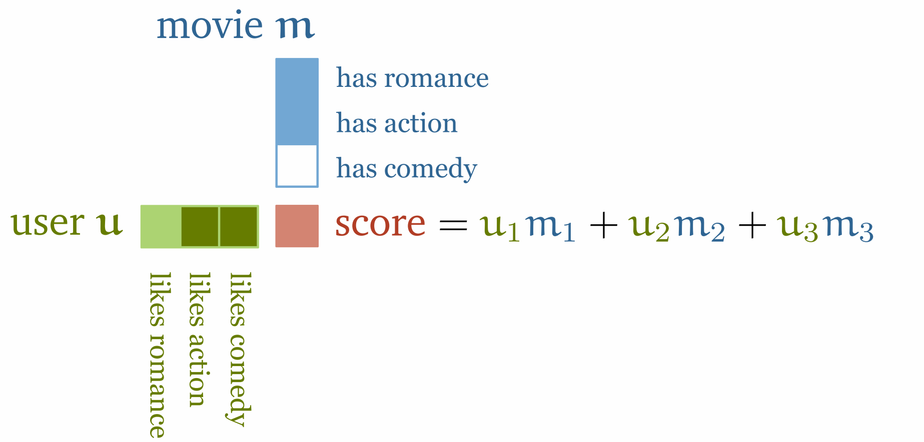


SELF-ATTENTION

Neural Machine Translation by Jointly Learning to Align and Translate; Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio, **2014**

https://peterbloem.nl/blog/transformers

SELF-ATTENTION

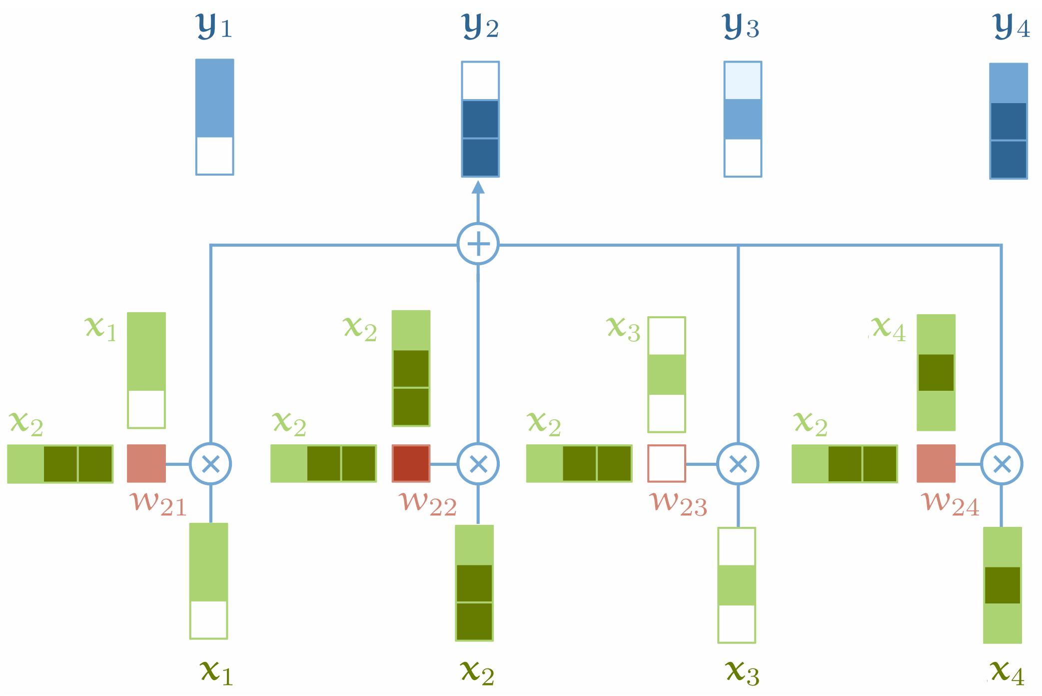


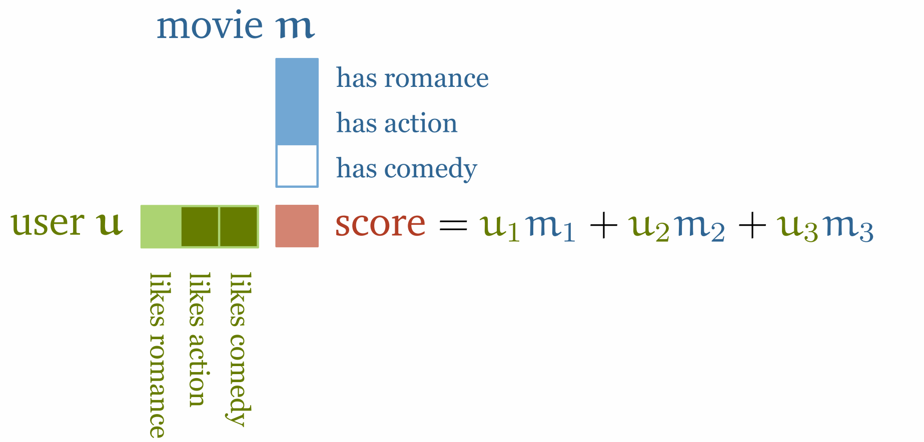
Neural Machine Translation by Jointly Learning to Align and Translate; Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio, **2014**

https://peterbloem.nl/blog/transformers

SELF-ATTENTION

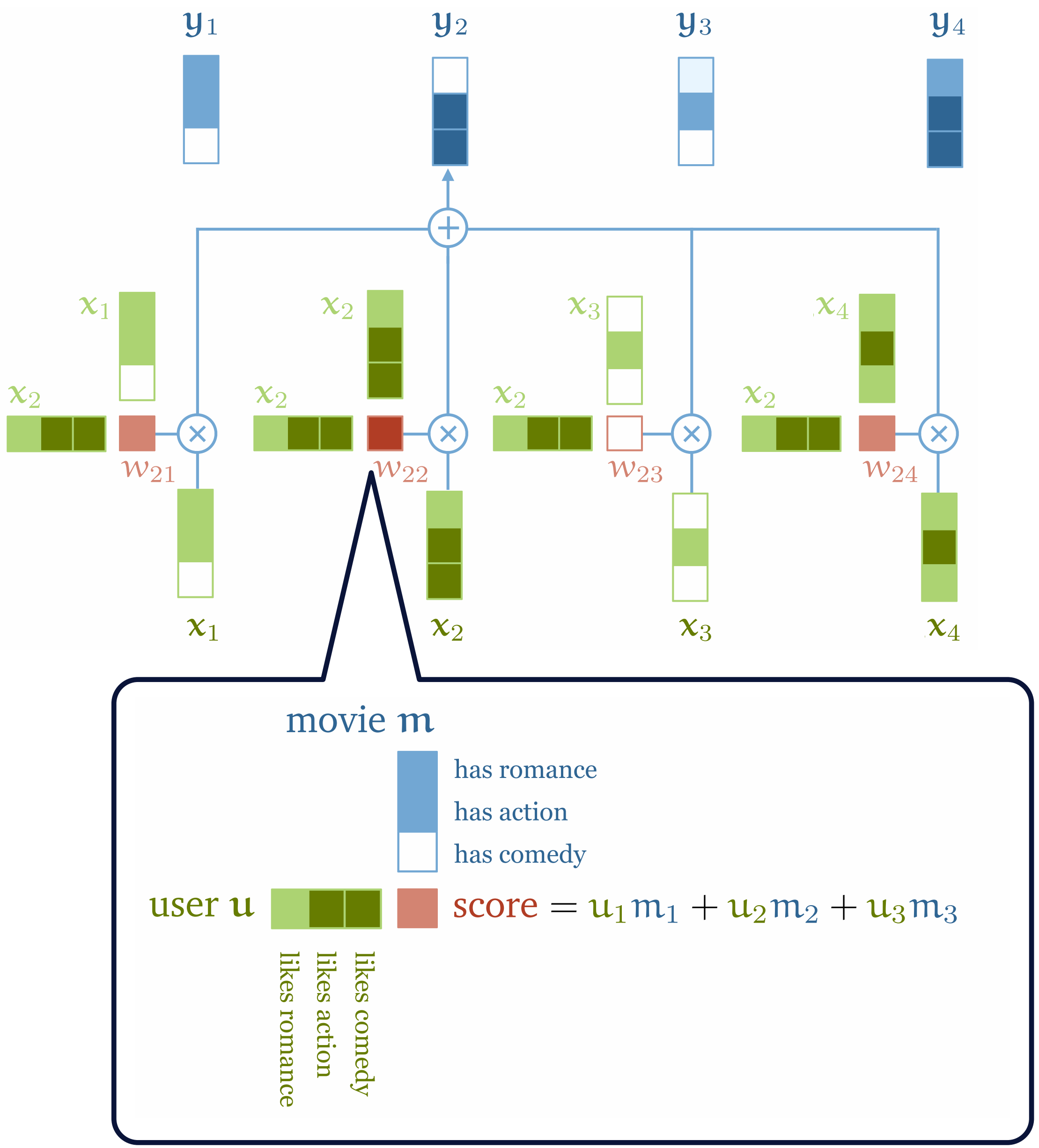
- Transformation of one sequence of vectors $x_1, x_2, \ldots, x_i$ (representing words) to another sequence of vectors $y_1, y_2, \ldots, y_i$





SELF-ATTENTION

- Transformation of one sequence of vectors $x_1, x_2, ..., x_i$ (representing words) to another sequence of vectors $y_1, y_2, ..., y_i$

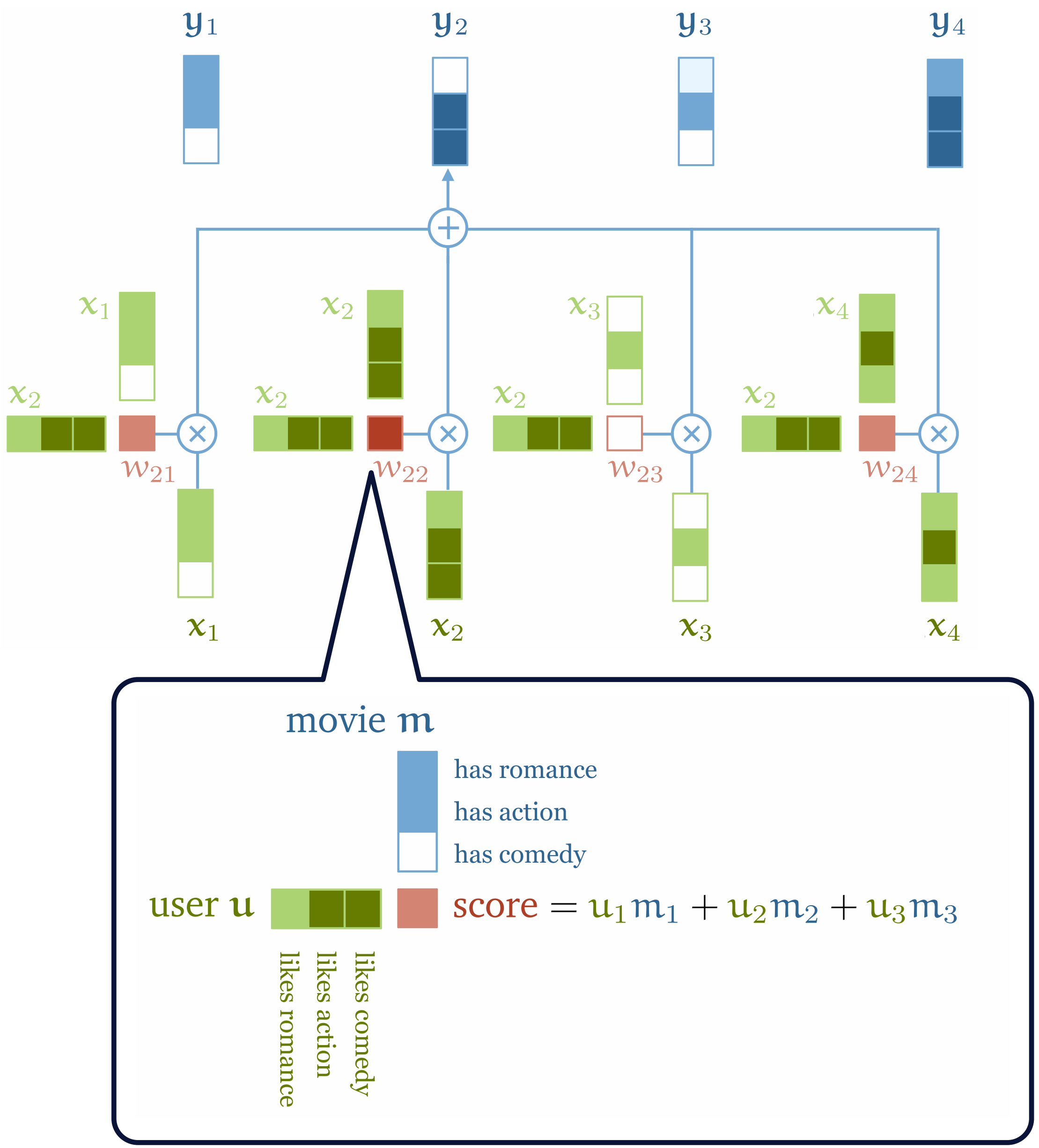


Neural Machine Translation by Jointly Learning to Align and Translate; Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio, **2014**

https://peterbloem.nl/blog/transformers

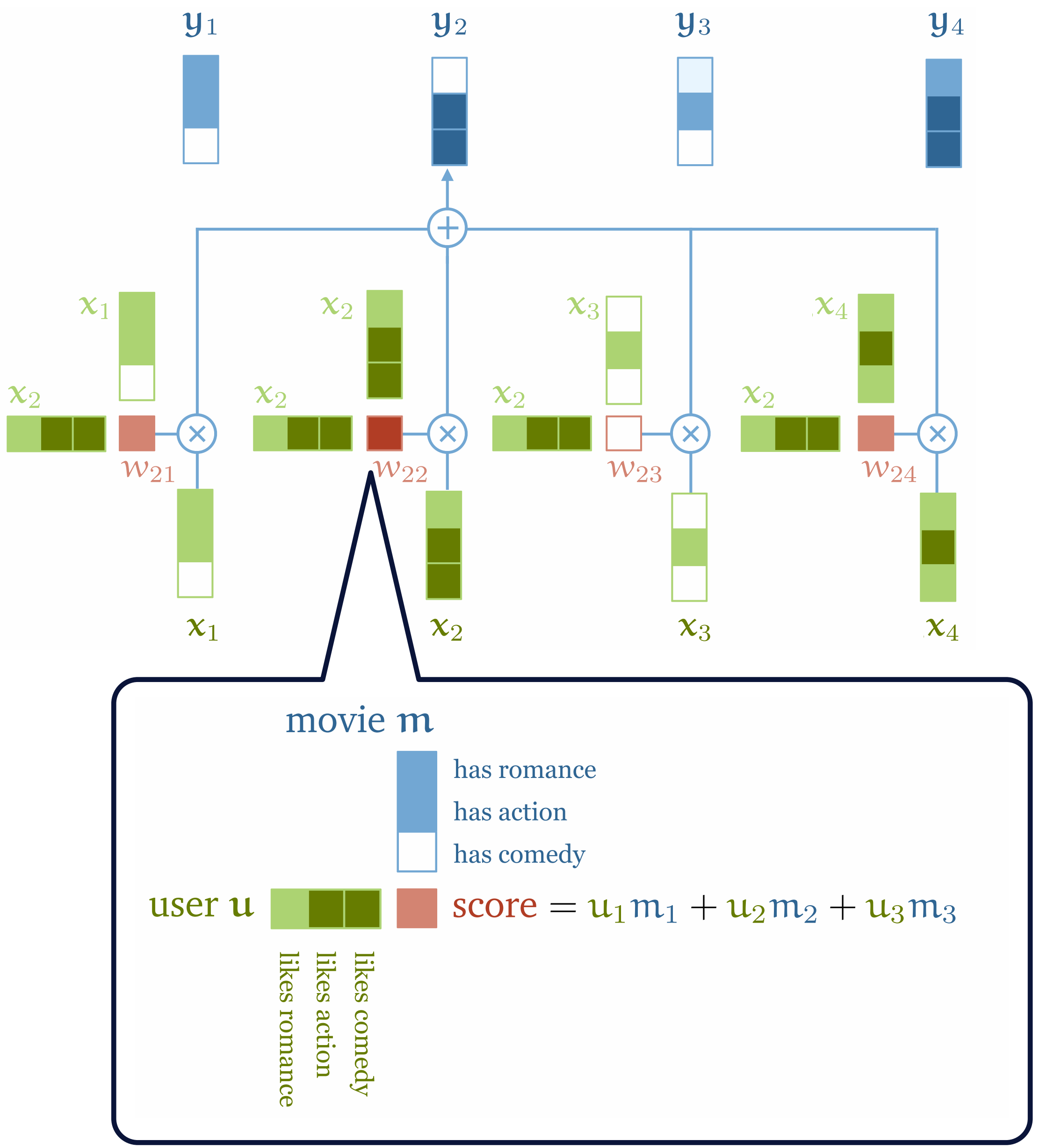
SELF-ATTENTION

- Transformation of one sequence of vectors $x_1, x_2, \ldots, x_i$ (representing words) to another sequence of vectors $y_1, y_2, \ldots, y_i$
  - Emphasize important parts



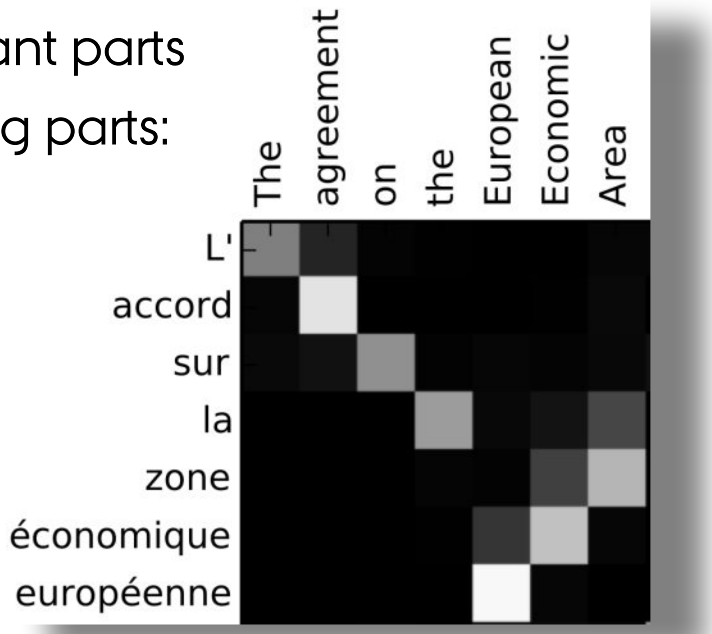
SELF-ATTENTION

- Transformation of one sequence of vectors $x_1, x_2, \ldots, x_i$ (representing words) to another sequence of vectors $y_1, y_2, \ldots, y_i$
  - Emphasize important parts
  - Fade-out less important parts

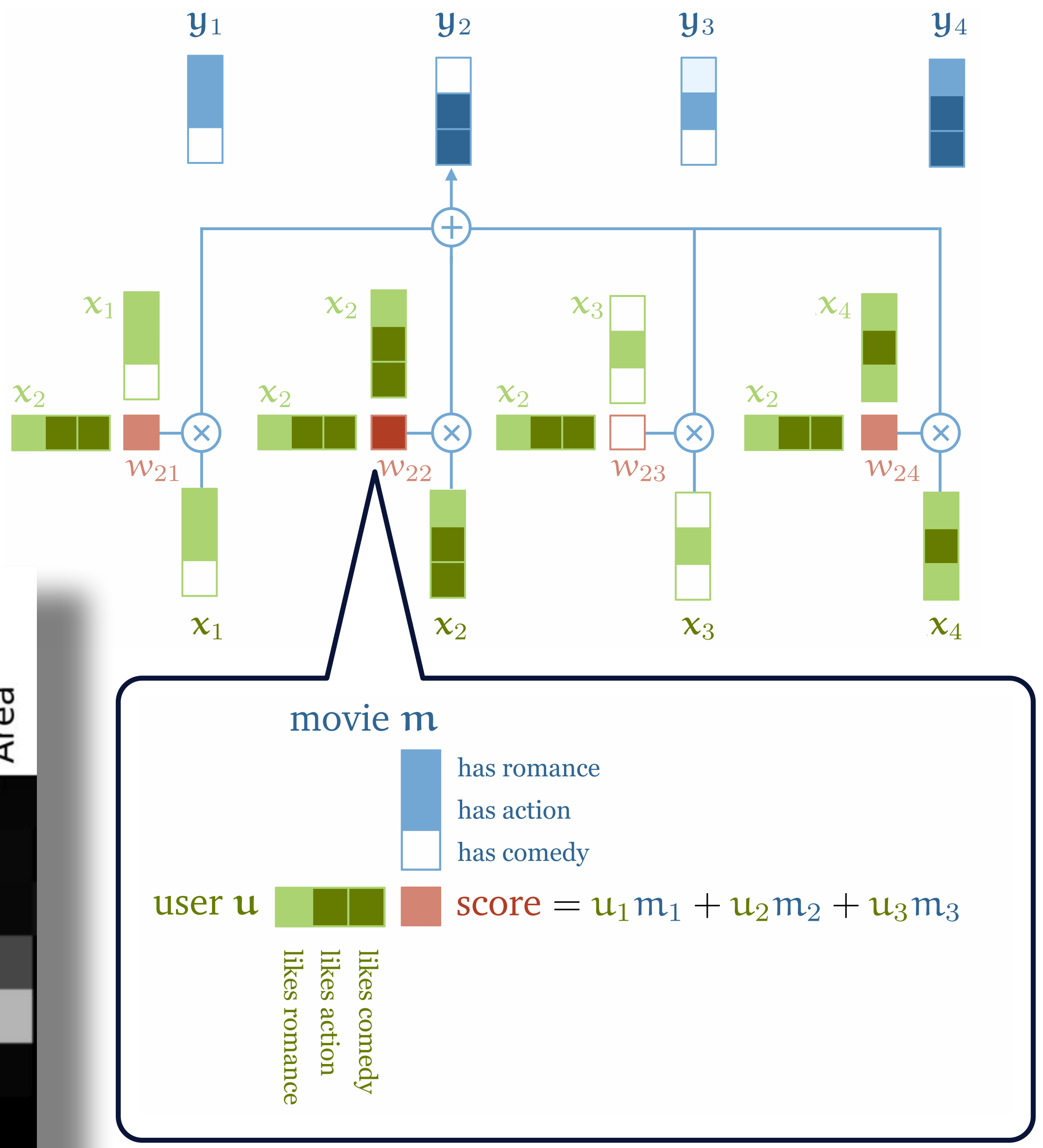


SELF-ATTENTION

- Transformation of one sequence of vectors $x_1, x_2, \ldots, x_i$ (representing words) to another sequence of vectors $y_1, y_2, \ldots, y_i$
  - Emphasize important parts
  - Fade-out less important parts
  - Identify corresponding parts:



Neural Machine Translation by Jointly Learning to Align and Translate; Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio, **2014**



https://peterbloem.nl/blog/transformers

INTERMEDIATE SUMMARY: CONDITIONAL PROBABILITIES, GIVEN SURROUNDING WORDS

We can think of context as giving us

INTERMEDIATE SUMMARY: CONDITIONAL PROBABILITIES, GIVEN SURROUNDING WORDS

We can think of context as giving us

a **possibility** space

- Of all the possible words in the world, which could be there?
- More importantly, which cannot?

INTERMEDIATE SUMMARY: CONDITIONAL PROBABILITIES, GIVEN SURROUNDING WORDS

We can think of context as giving us

a **possibility** space

- Of all the possible words in the world, which could be there?
- More importantly, which cannot?

and a **probable** space:

- Which words are the more likely?

INTERMEDIATE SUMMARY: CONDITIONAL PROBABILITIES, GIVEN SURROUNDING WORDS

We can think of context as giving us

a **possibility** space

- Of all the possible words in the world, which could be there?
- More importantly, which cannot?

and a **probable** space:

- Which words are the more likely?

Formally, these are all probabilities (where the impossible ones have a probability of 0.0)

INTERMEDIATE SUMMARY: CONDITIONAL PROBABILITIES, GIVEN SURROUNDING WORDS

We can think of context as giving us

a **possibility** space

- Of all the possible words in the world, which could be there?
- More importantly, which cannot?

and a **probable** space:

- Which words are the more likely?

Internally an LLMs use vectors representing words.
How to get these?

Formally, these are all probabilities (where the impossible ones have a probability of 0.0)

TRANSFORMER LANGUAGE MODELS

Let's get technical: The Transformer architecture of LLMs

AARHUS BSS
DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

MAGNUS BENDER
ASSISTANT PROFESSOR

WORDS AS VECTORS

Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean „Efficient Estimation of Word Representations in Vector Space", **2013**

WORDS AS VECTORS

- Represent the meaning of words in vectors

Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean „Efficient Estimation of Word Representations in Vector Space", **2013**

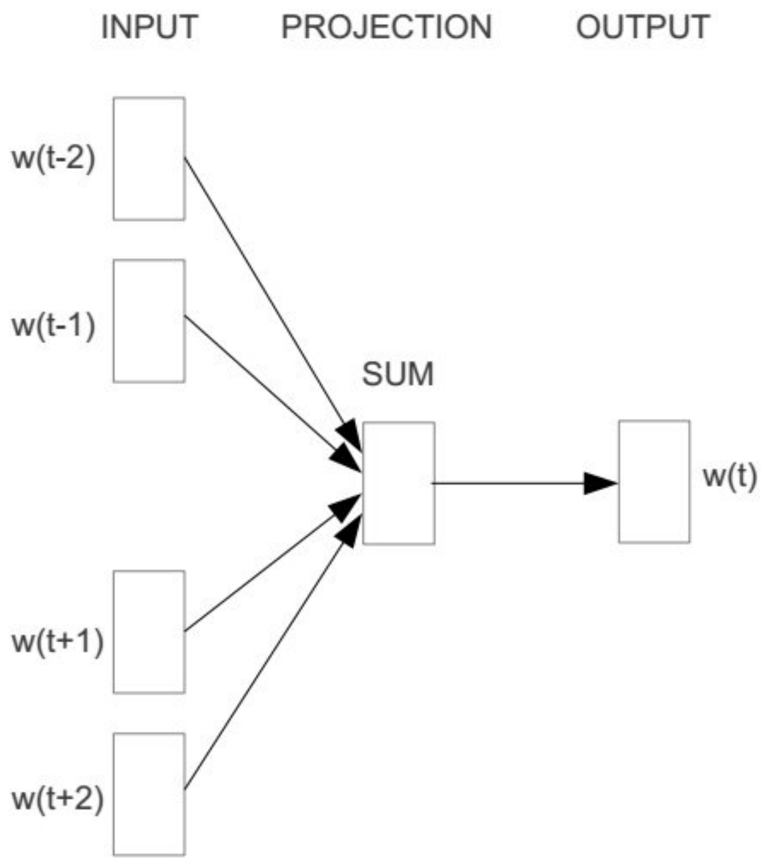
WORDS AS VECTORS

- Represent the meaning of words in vectors
- Example:

Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean „Efficient Estimation of Word Representations in Vector Space", **2013**

WORDS AS VECTORS

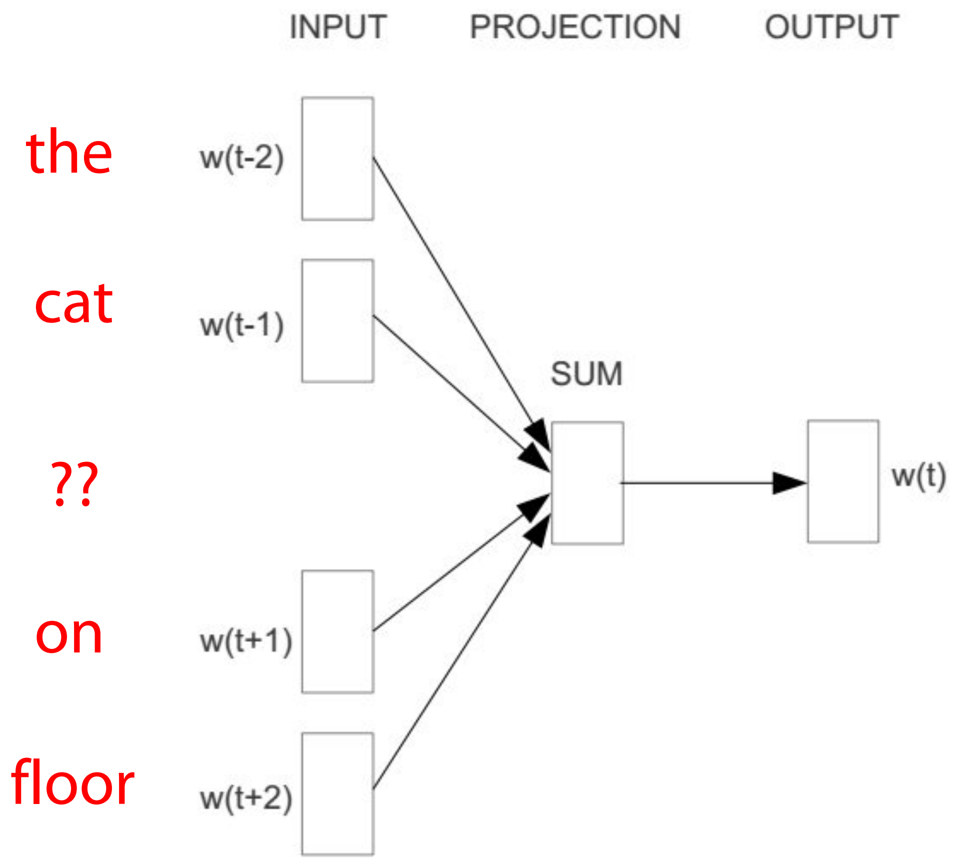
- Represent the meaning of words in vectors
- Example:
  - Continuous Bag of Words (CBOW) from the so-called Word2Vec approach



Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean „Efficient Estimation of Word Representations in Vector Space“, **2013**

WORDS AS VECTORS

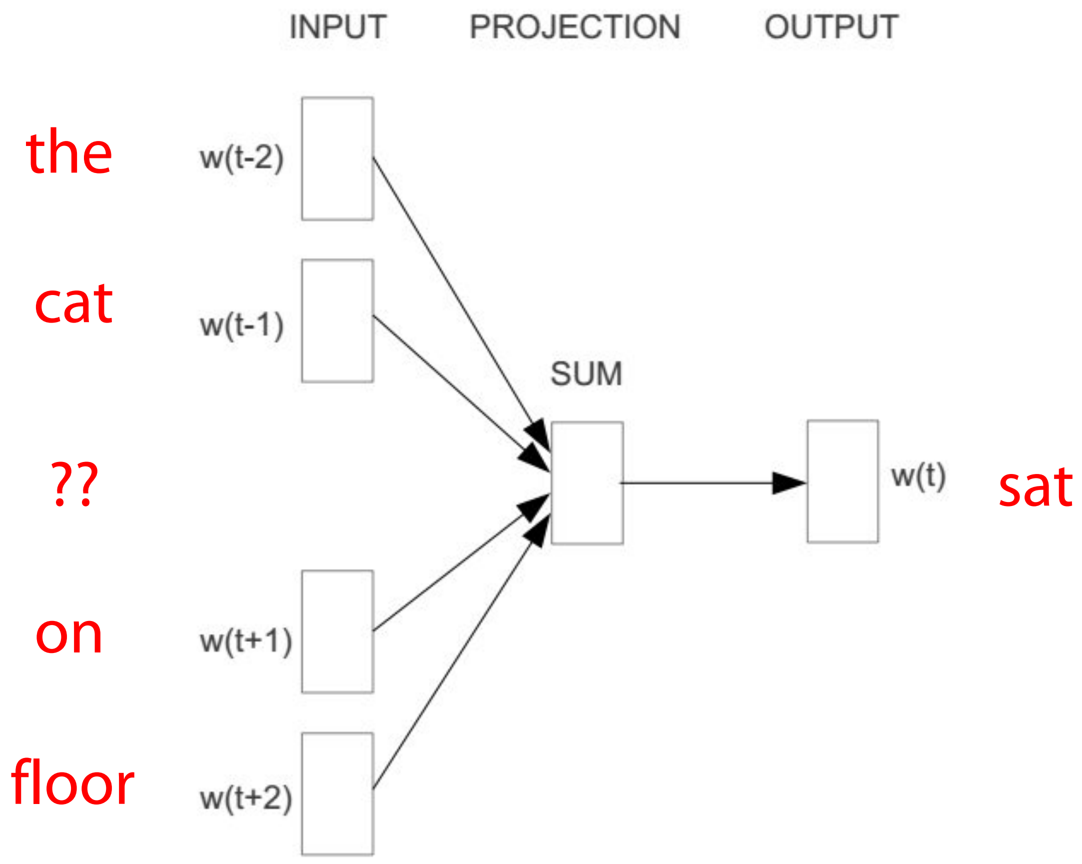
- Represent the meaning of words in vectors
- Example:
  - Continuous Bag of Words (CBOW) from the so-called Word2Vec approach
  - Use a window of words to predict the middle word



Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean „Efficient Estimation of Word Representations in Vector Space“, **2013**

WORDS AS VECTORS

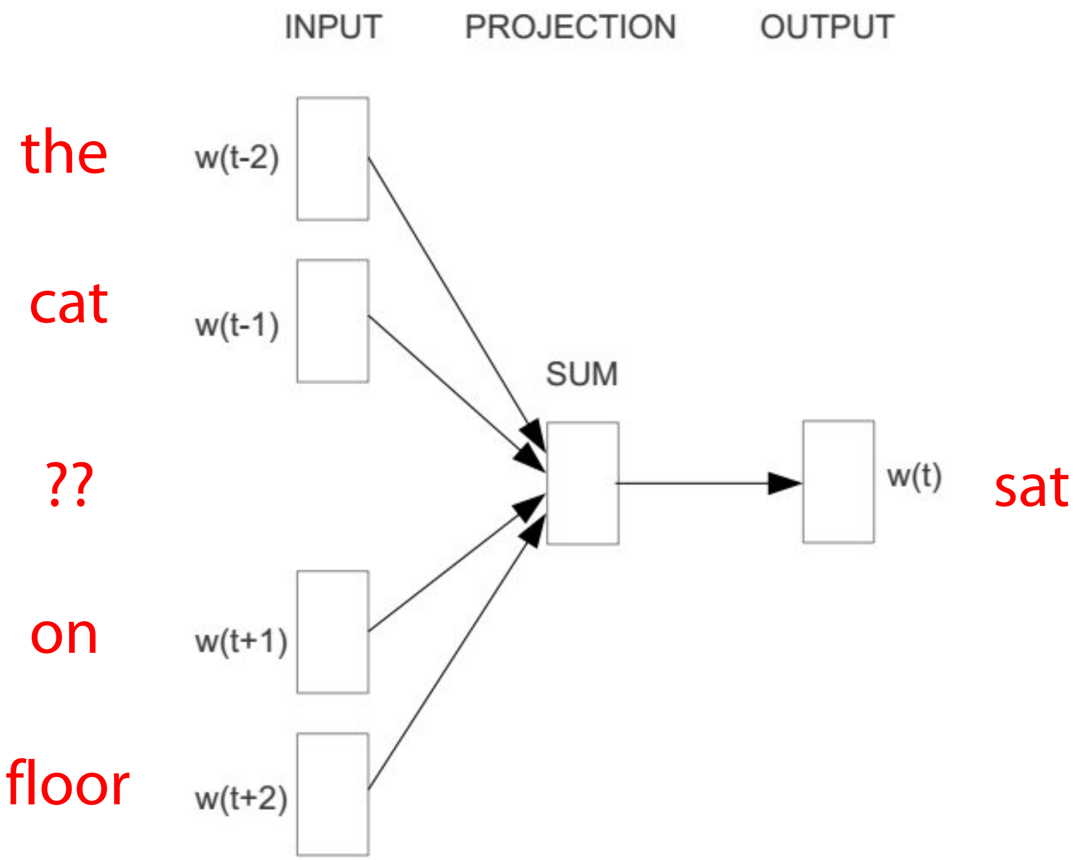
- Represent the meaning of words in vectors
- Example:
  - Continuous Bag of Words (CBOW) from the so-called Word2Vec approach
  - Use a window of words to predict the middle word



Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean „Efficient Estimation of Word Representations in Vector Space", **2013**

WORDS AS VECTORS

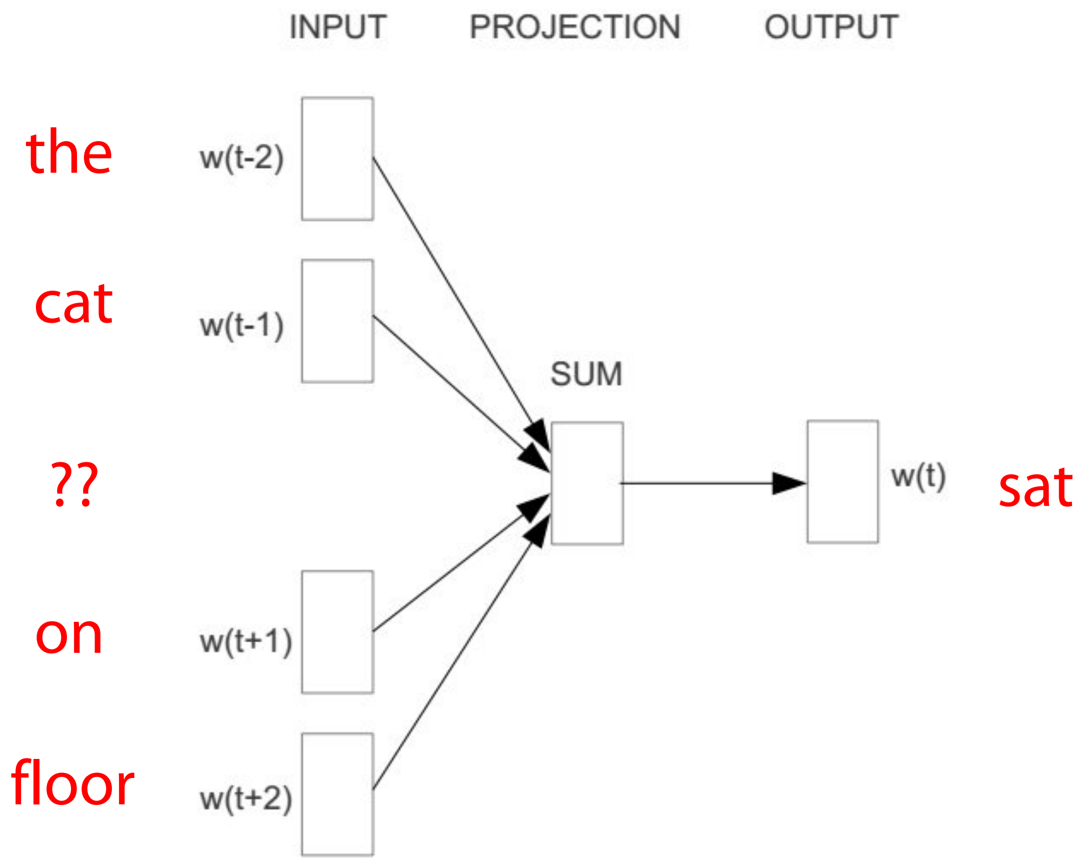
- Represent the meaning of words in vectors
- Example:
  - Continuous Bag of Words (CBOW) from the so-called Word2Vec approach
  - Use a window of words to predict the middle word
- Generally:
  - A vector representing a word (or token) is called an embedding



Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean „Efficient Estimation of Word Representations in Vector Space", **2013**

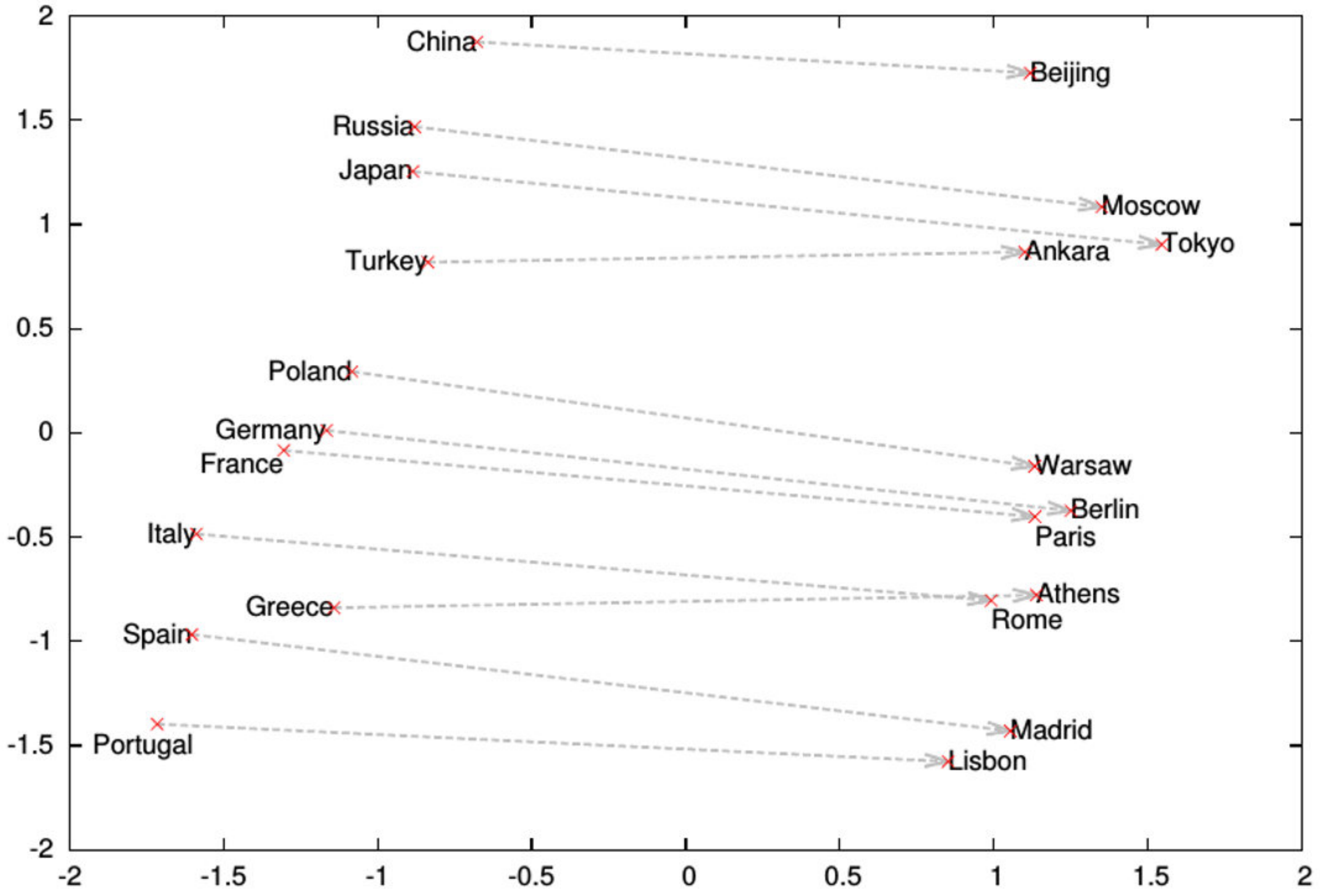
WORDS AS VECTORS

- Represent the meaning of words in vectors
- Example:
  - Continuous Bag of Words (CBOW) from the so-called Word2Vec approach
  - Use a window of words to predict the middle word
- Generally:
  - A vector representing a word (or token) is called an embedding
  - Embeddings are trained/ crafted to represent a words as vectors useful for the current use-case!



Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean „Efficient Estimation of Word Representations in Vector Space", **2013**

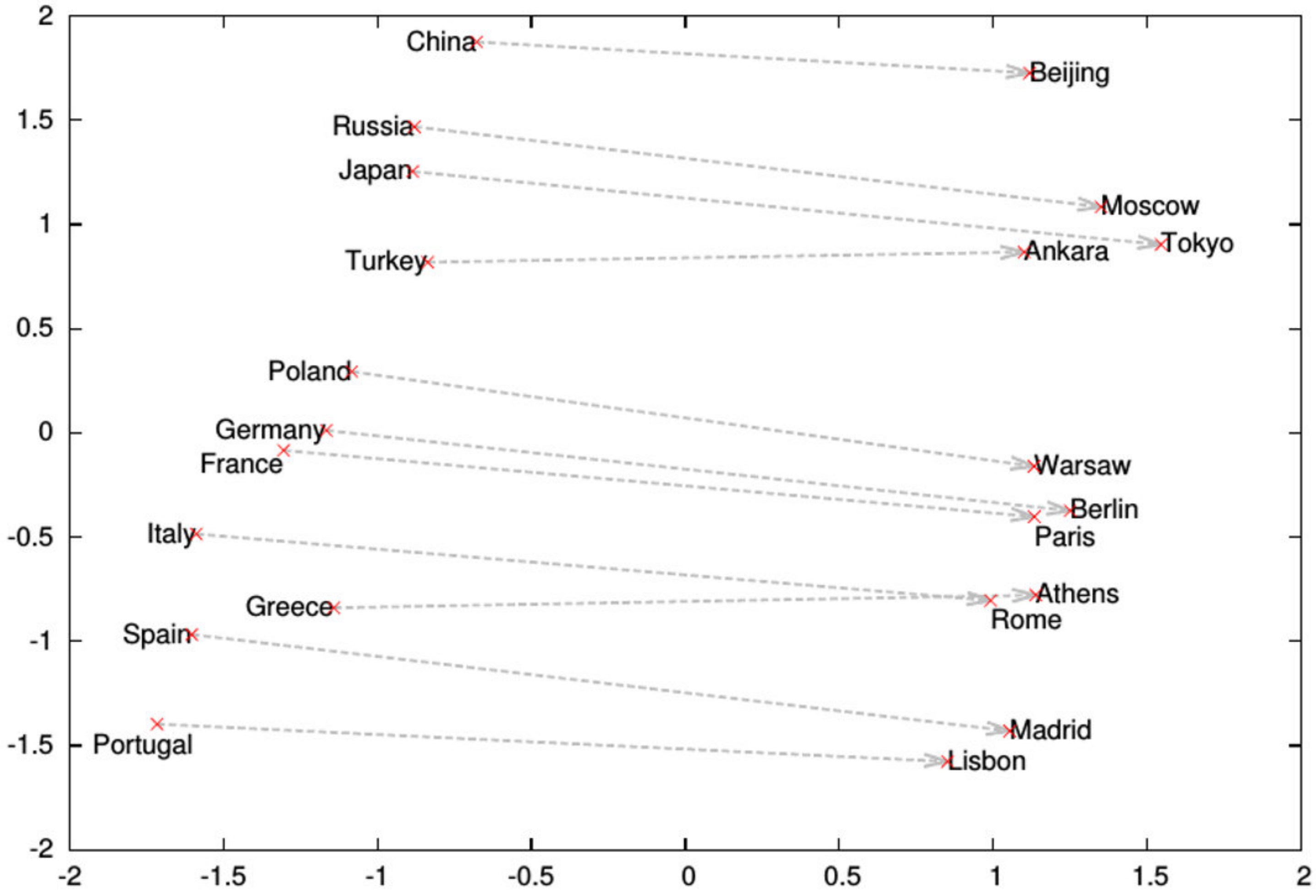
EXAMPLE: WORD2VEC



https://opensource.googleblog.com/2013/08/learning-meaning-behind-words.html

EXAMPLE: WORD2VEC

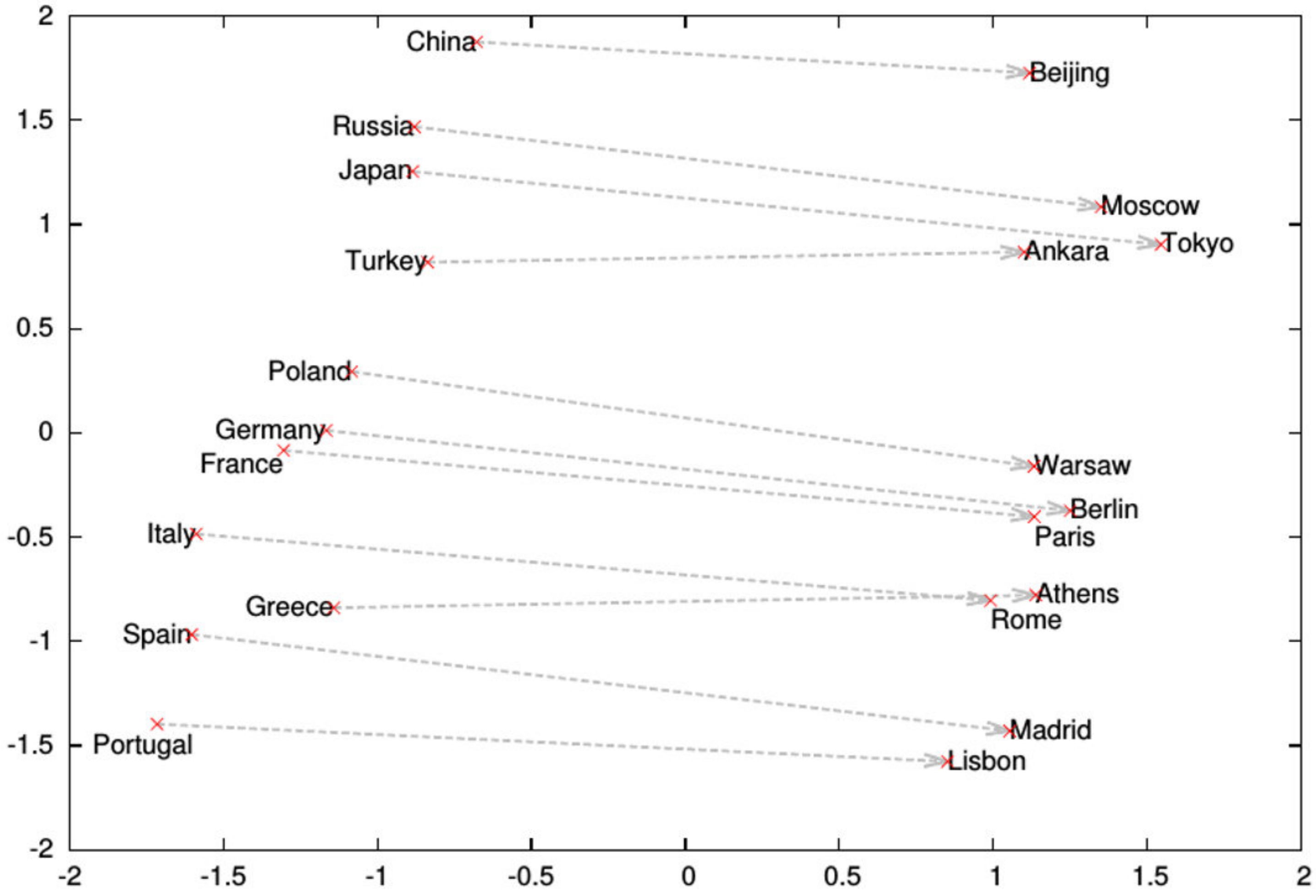
- Known vectors (positions) for:



https://opensource.googleblog.com/2013/08/learning-meaning-behind-words.html

EXAMPLE: WORD2VEC

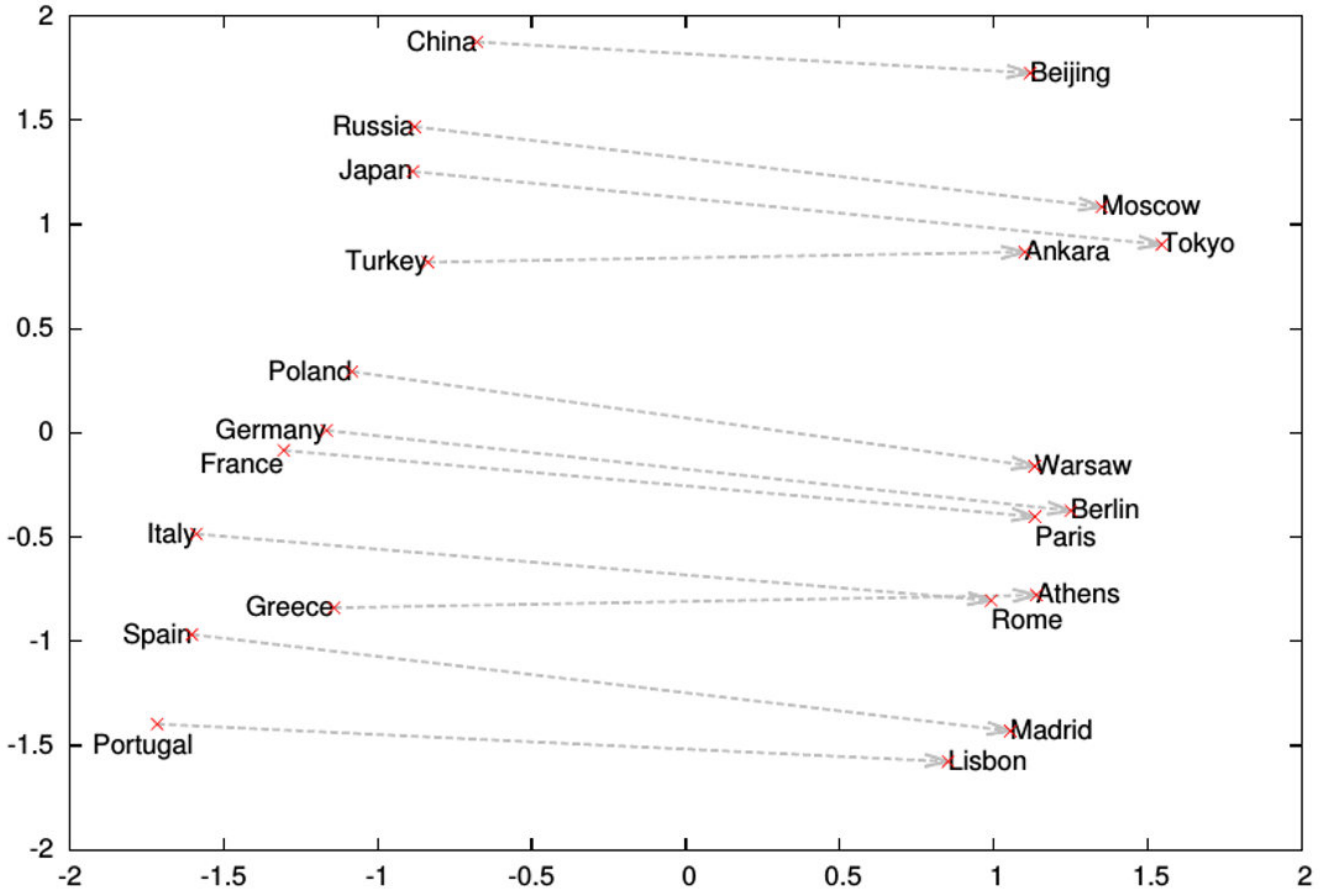
- Known vectors (positions) for:
  - Man
  - Woman
  - King



https://opensource.googleblog.com/2013/08/learning-meaning-behind-words.html

EXAMPLE: WORD2VEC

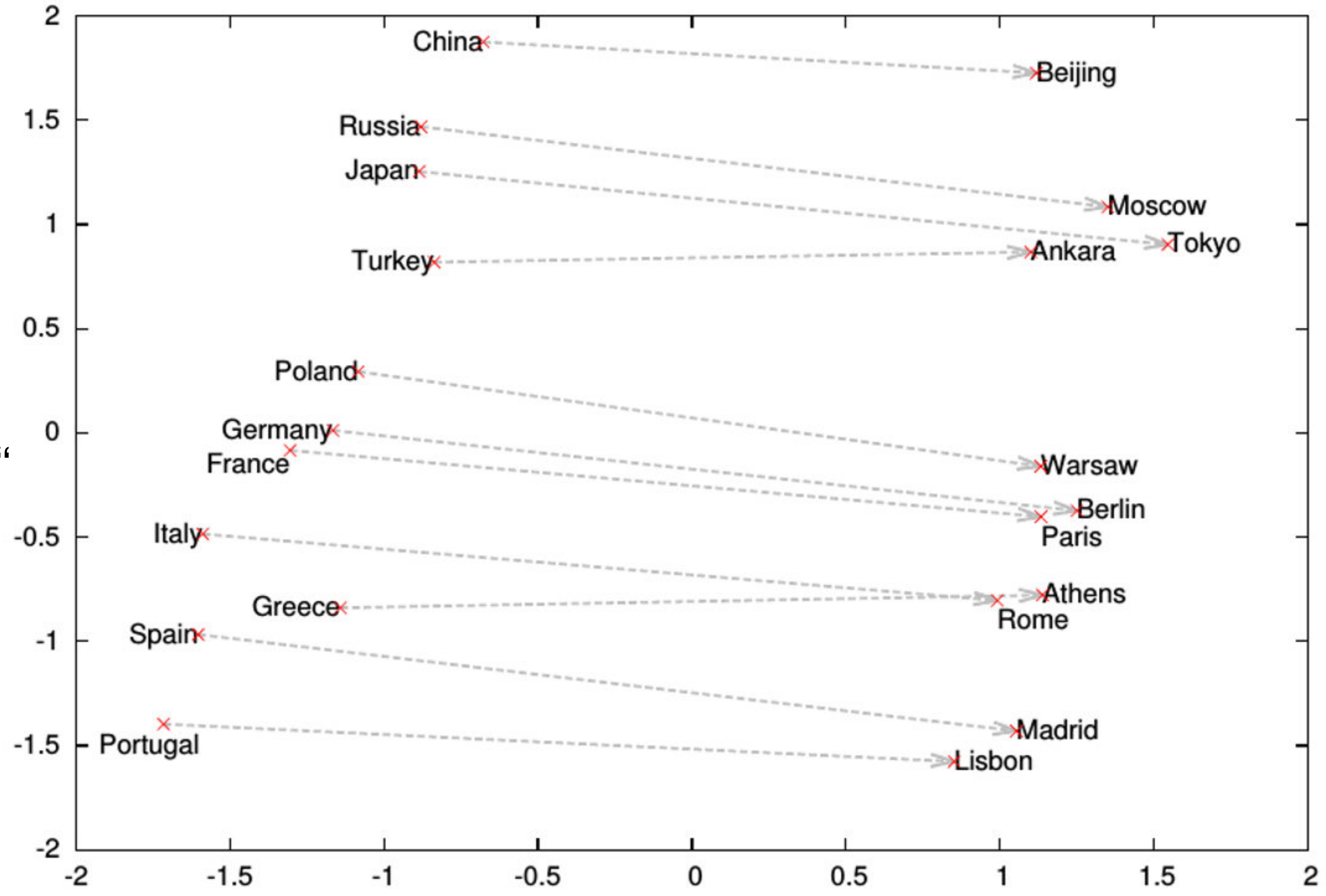
- Known vectors (positions) for:
  - Man
  - Woman
  - King
- Known vectors (differences) for:
  - Man → Woman



https://opensource.googleblog.com/2013/08/learning-meaning-behind-words.html

EXAMPLE: WORD2VEC

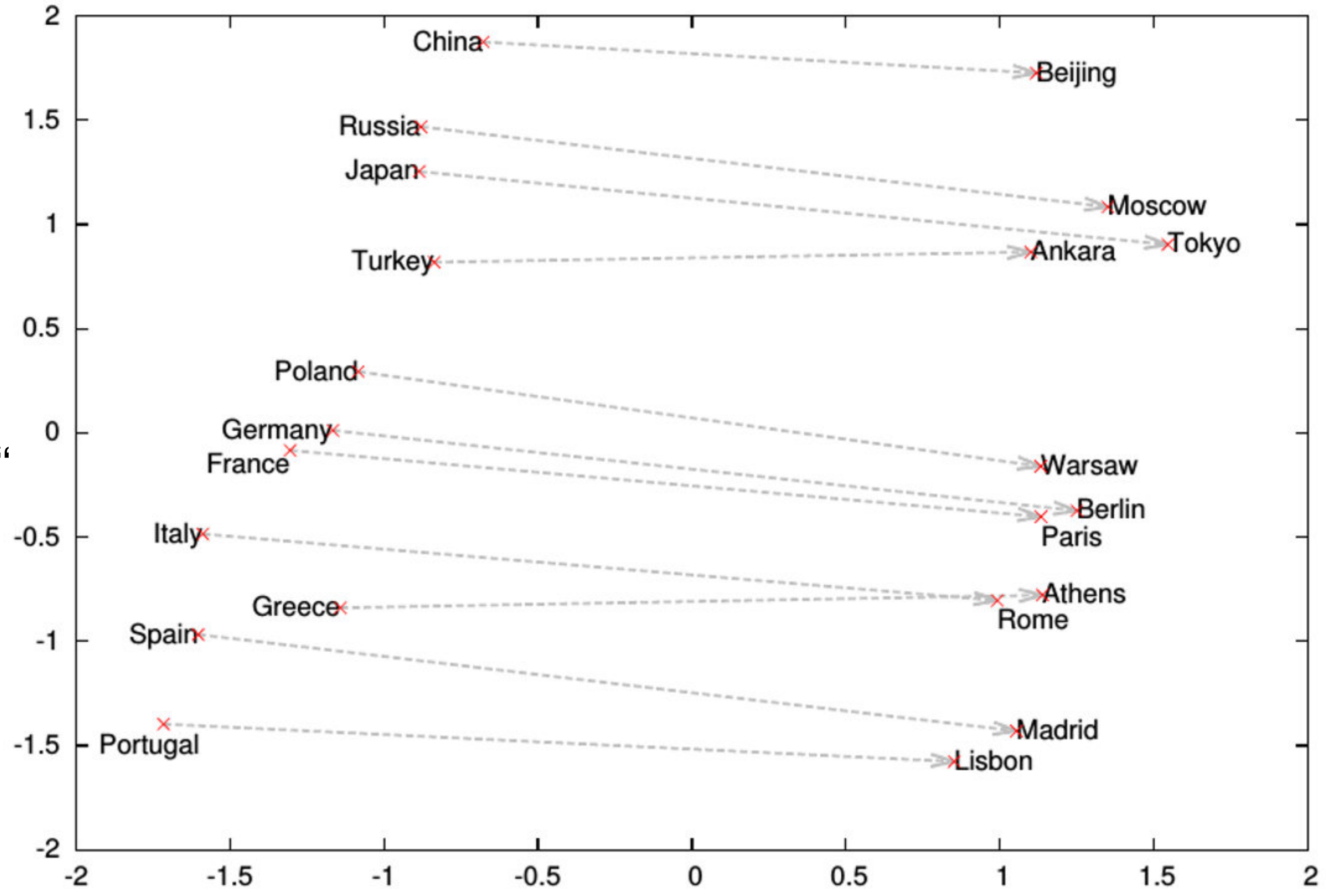
- Known vectors (positions) for:
  - Man
  - Woman
  - King
- Known vectors (differences) for:
  - Man → Woman
- What is the corresponding word for „King“ in the relation „Man → Woman“
  - King → ?



https://opensource.googleblog.com/2013/08/learning-meaning-behind-words.html

EXAMPLE: WORD2VEC

- Known vectors (positions) for:
  - Man
  - Woman
  - King
- Known vectors (differences) for:
  - Man → Woman
- What is the corresponding word for „King“ in the relation „Man → Woman“
  - King → ?
  - King - Man + Woman = Queen

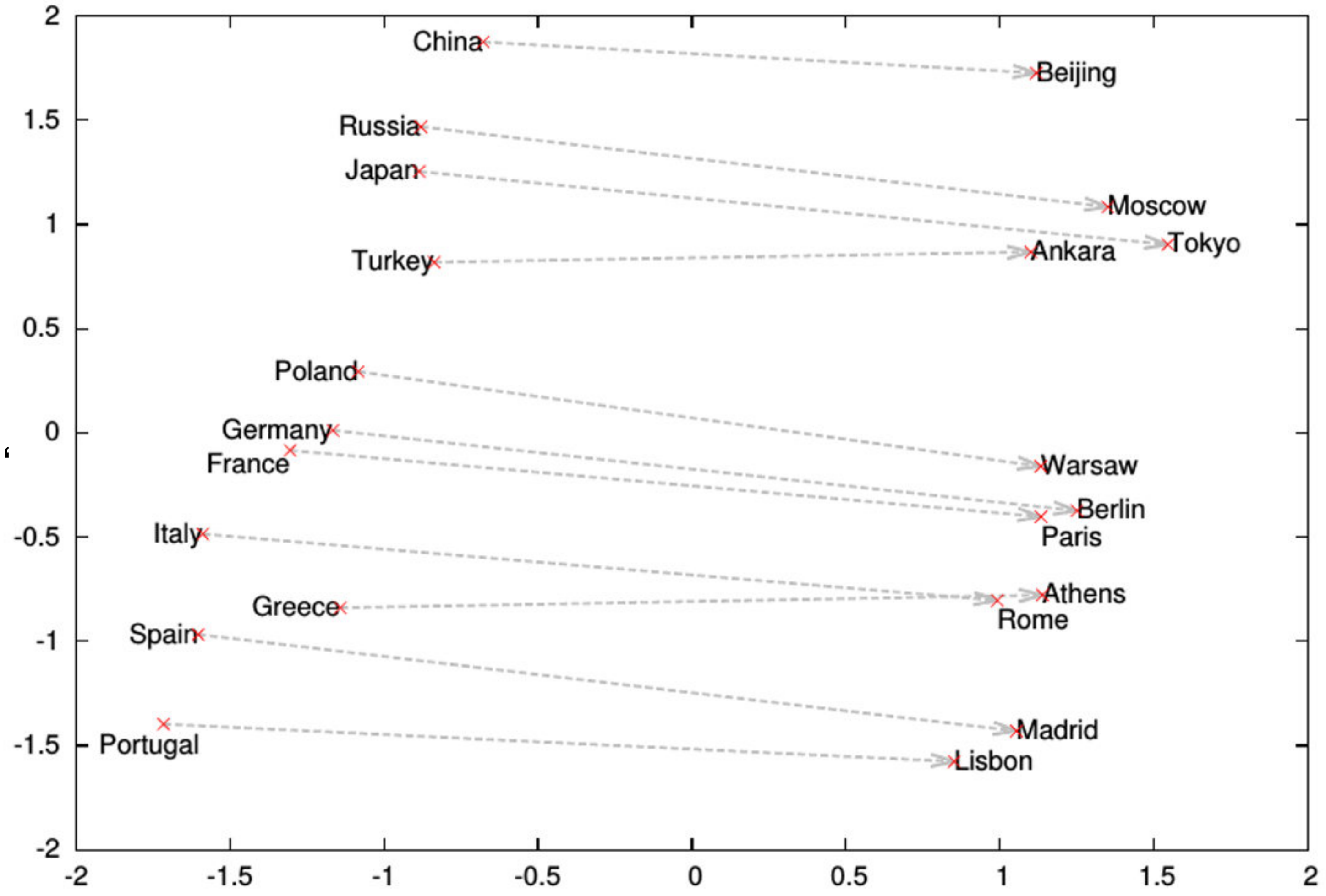


https://opensource.googleblog.com/2013/08/learning-meaning-behind-words.html

EXAMPLE: WORD2VEC

- Known vectors (positions) for:
  - Man
  - Woman
  - King
- Known vectors (differences) for:
  - Man → Woman
- What is the corresponding word for „King“ in the relation „Man → Woman“
  - King → ?
  - King - Man + Woman = Queen

Representing words by vectors in a multi-dimensional vector space



https://opensource.googleblog.com/2013/08/learning-meaning-behind-words.html

LLMS DO THE SAME

- We said „A word is represented by a vector“

LLMS DO THE SAME

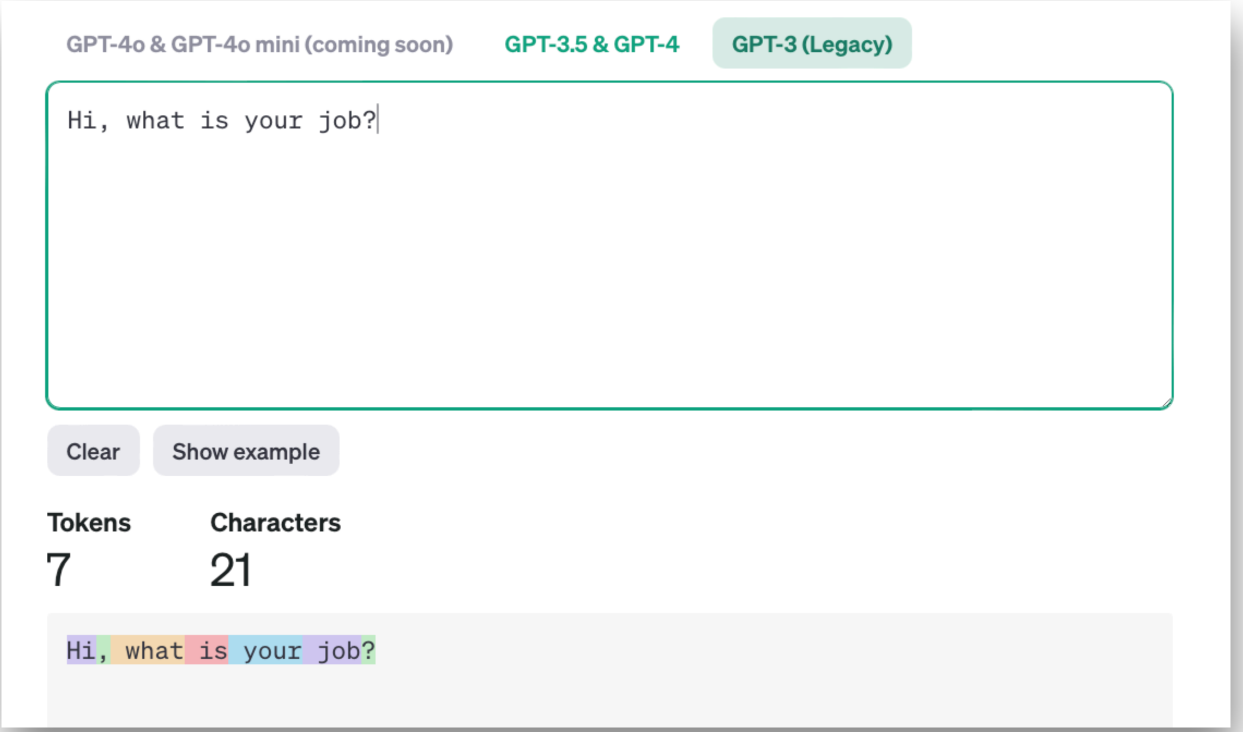
- We said „A word is represented by a vector“
- But what to do with special names or very uncommon words

LLMS DO THE SAME

- We said „A word is represented by a vector“
- But what to do with special names or very uncommon words
  - Less common „barrel“
  - Name „Aarhus“
  - Typo „cllass“

LLMS DO THE SAME

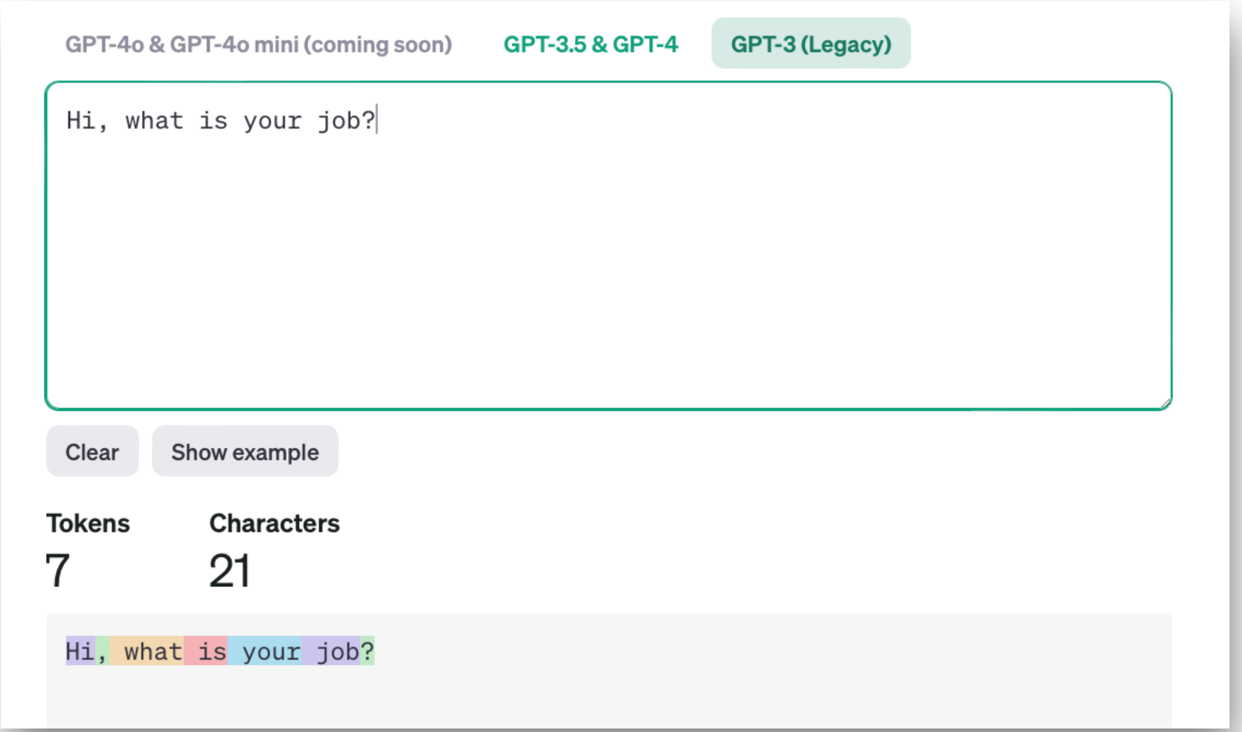
- We said „A word is represented by a vector“
- But what to do with special names or very uncommon words
  - Less common „barrel“
  - Name „Aarhus“
  - Typo „cllass“
- Words are split into tokens
  - Each token may represent a full word, a part of it, or only one character or symbol.



https://platform.openai.com/tokenizer

LLMS DO THE SAME

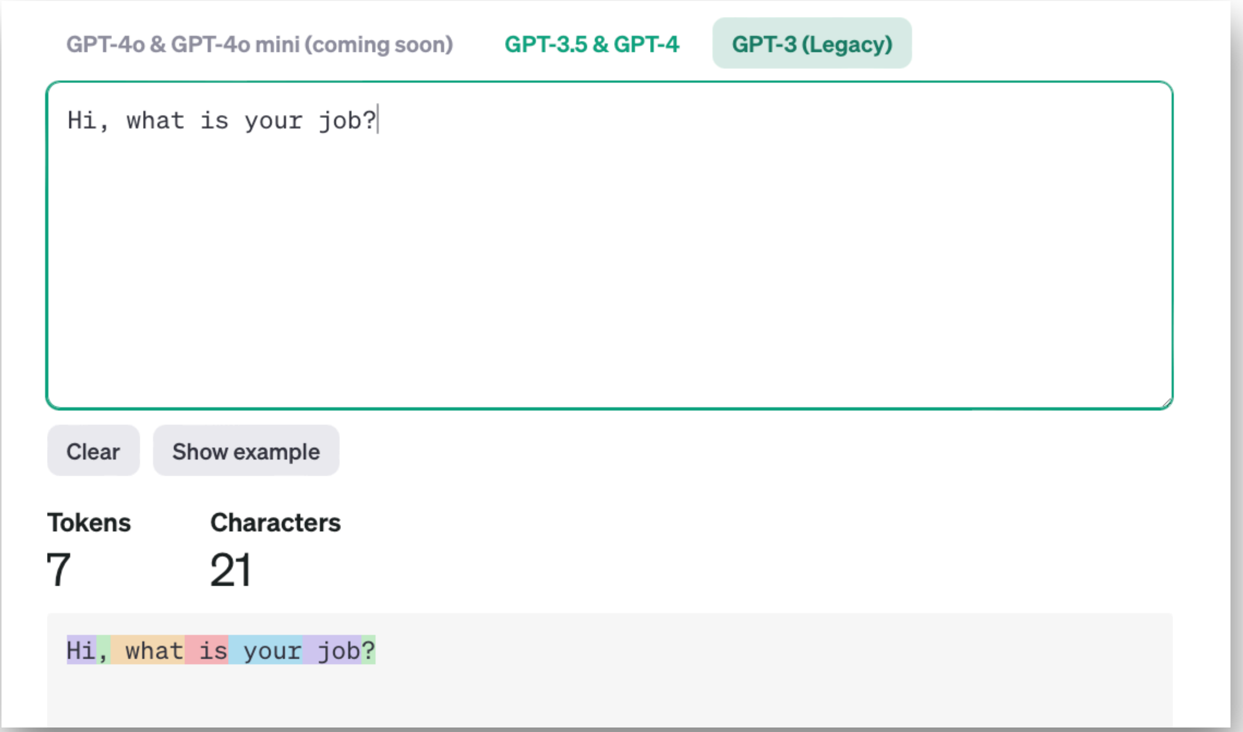
- We said „A word is represented by a vector“
- But what to do with special names or very uncommon words
  - Less common „barrel“
  - Name „Aarhus“
  - Typo „cllass“
- Words are split into tokens
  - Each token may represent a full word, a part of it, or only one character or symbol.
- LLM is able to get any sequence of text as input and to create any sequence, too.



https://platform.openai.com/tokenizer

LLMS DO THE SAME

- We said „A word is represented by a vector“
- But what to do with special names or very uncommon words
  - Less common „barrel“
  - Name „Aarhus“
  - Typo „cllass“
- Words are split into tokens
  - Each token may represent a full word, a part of it, or only one character or symbol.
- LLM is able to get any sequence of text as input and to create any sequence, too.
  - The outputs are again vectors representing tokens wich are transformed to their corresponding word or character.

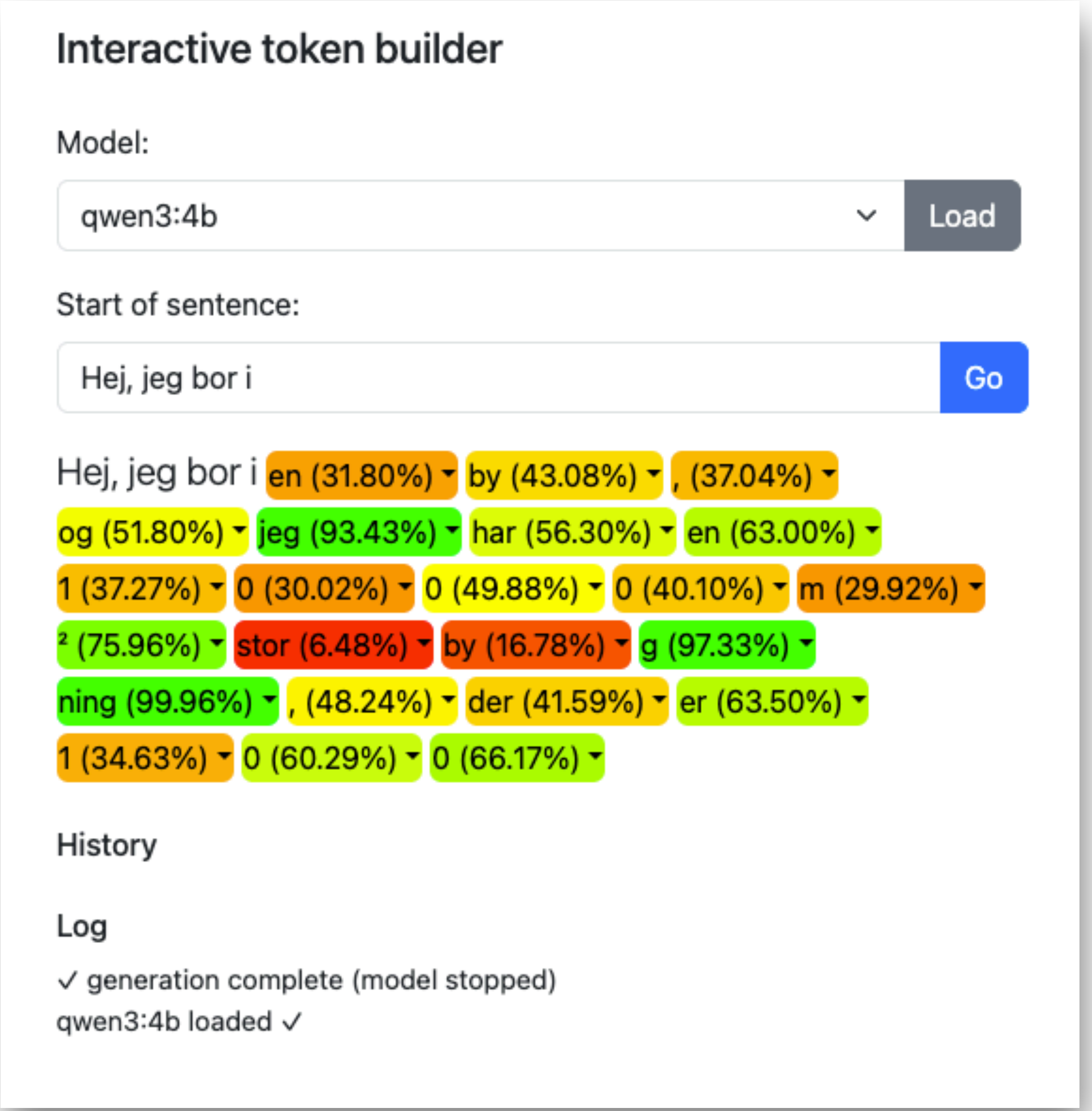


https://platform.openai.com/tokenizer

DEMO: TOKEN

„Hej, jeg bor i en by, og jeg har en 1-0-0-m-² stor by-g-ning, ...“

- No token representing the word „bygning“
- Model groups multiple tokens, representing substrings to form the word „bygning“

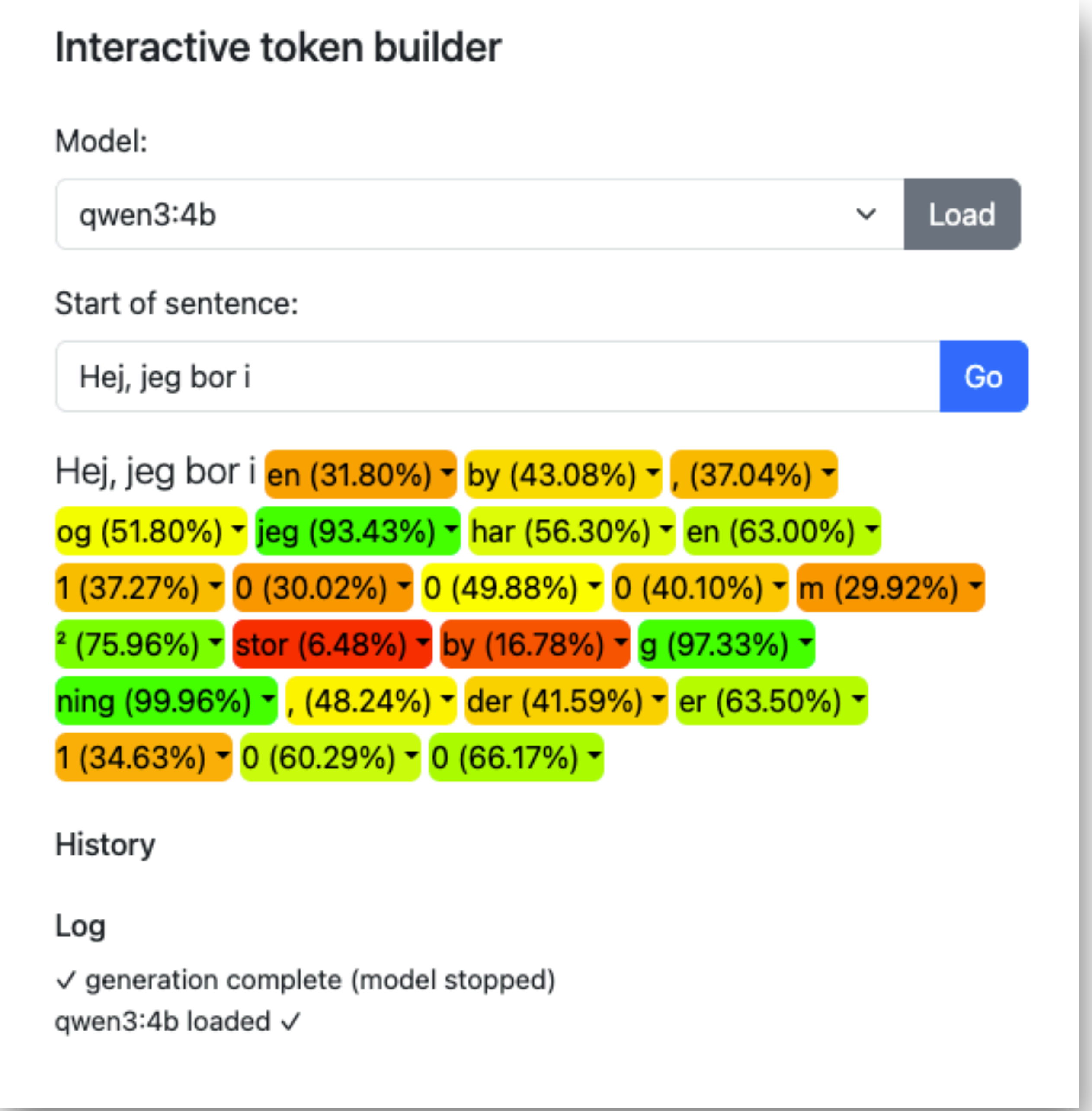


DEMO: TOKEN

„Hej, jeg bor i en by, og jeg har en 1-0-0-m-² stor by-g-ning, …“

- No token representing the word „bygning“
- Model groups multiple tokens, representing substrings to form the word „bygning“

➡ Thus, possible to represent any word by letters and efficient to represent common words by their token.



TRANSFORMER BASED LLM

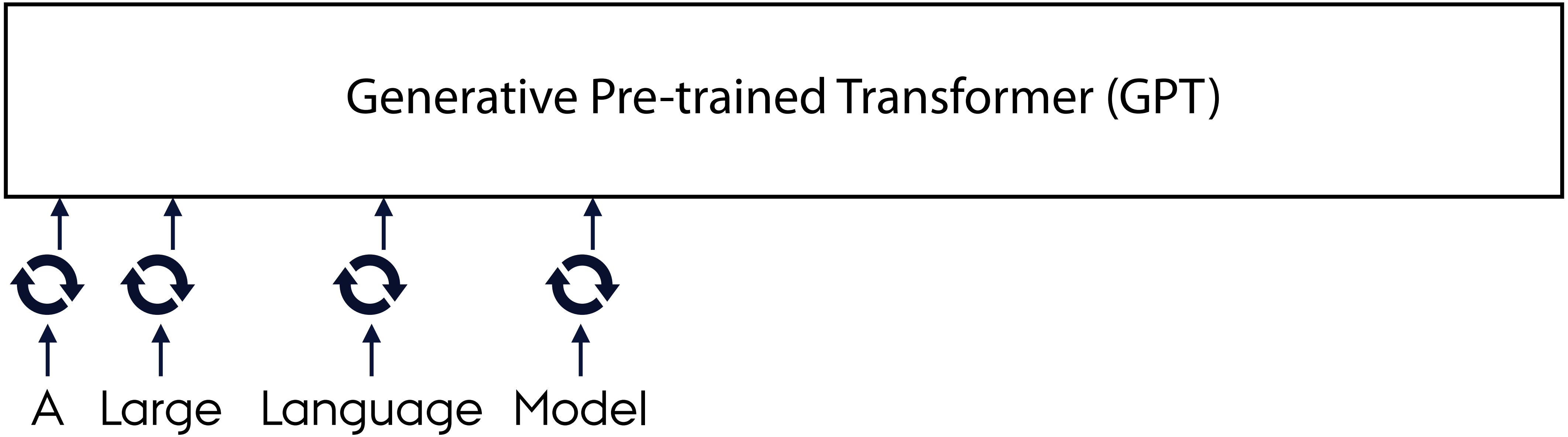
A Large Language Model

TRANSFORMER BASED LLM

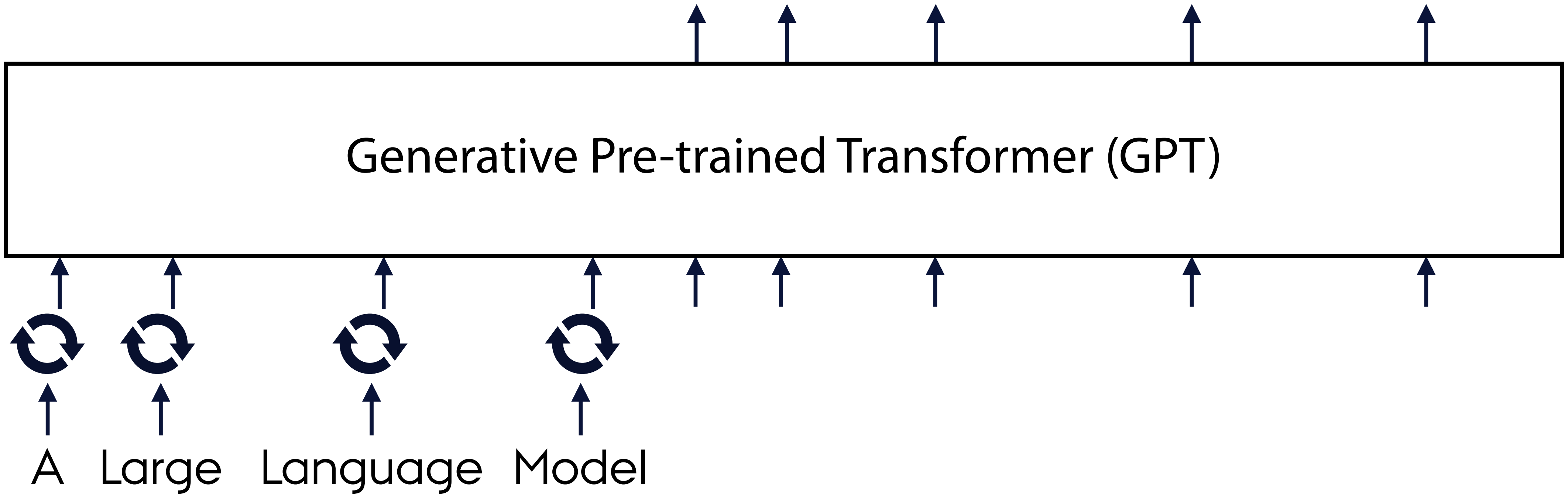
Generative Pre-trained Transformer (GPT)

A Large Language Model

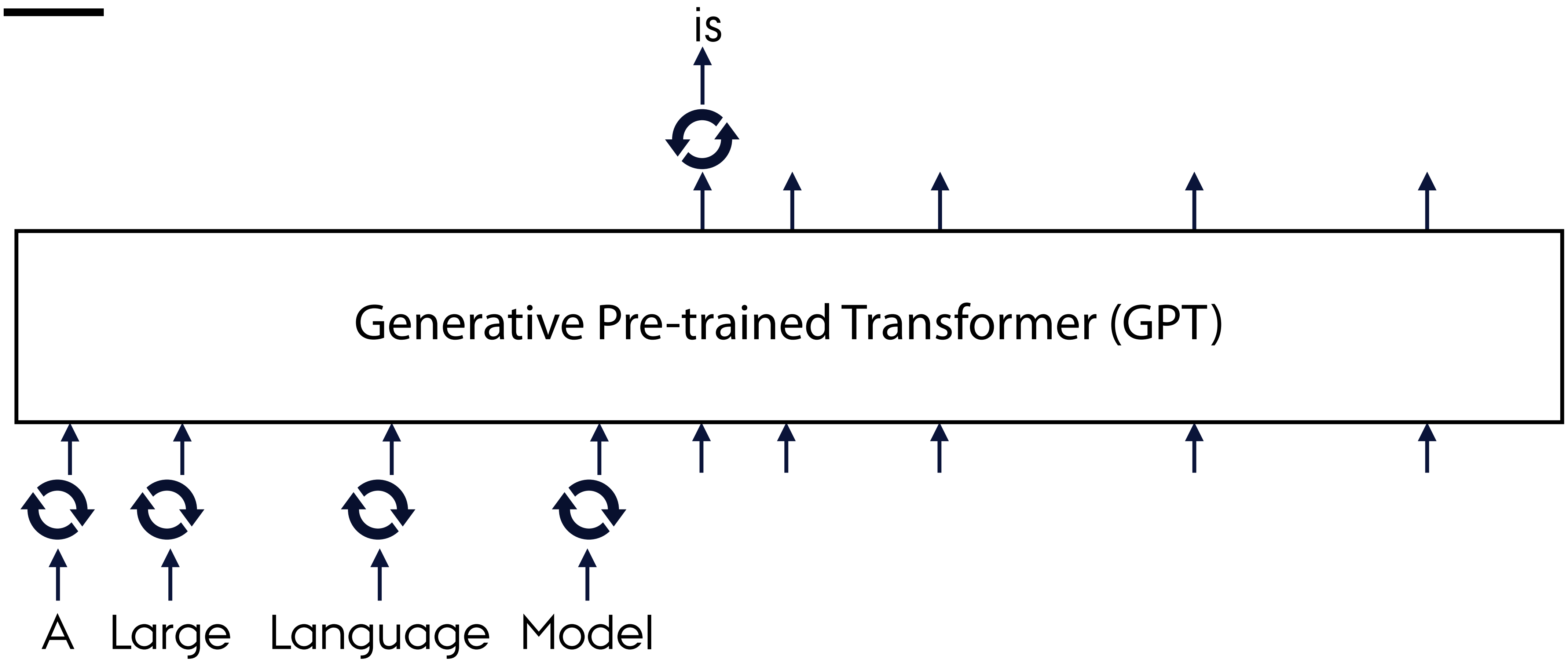
TRANSFORMER BASED LLM



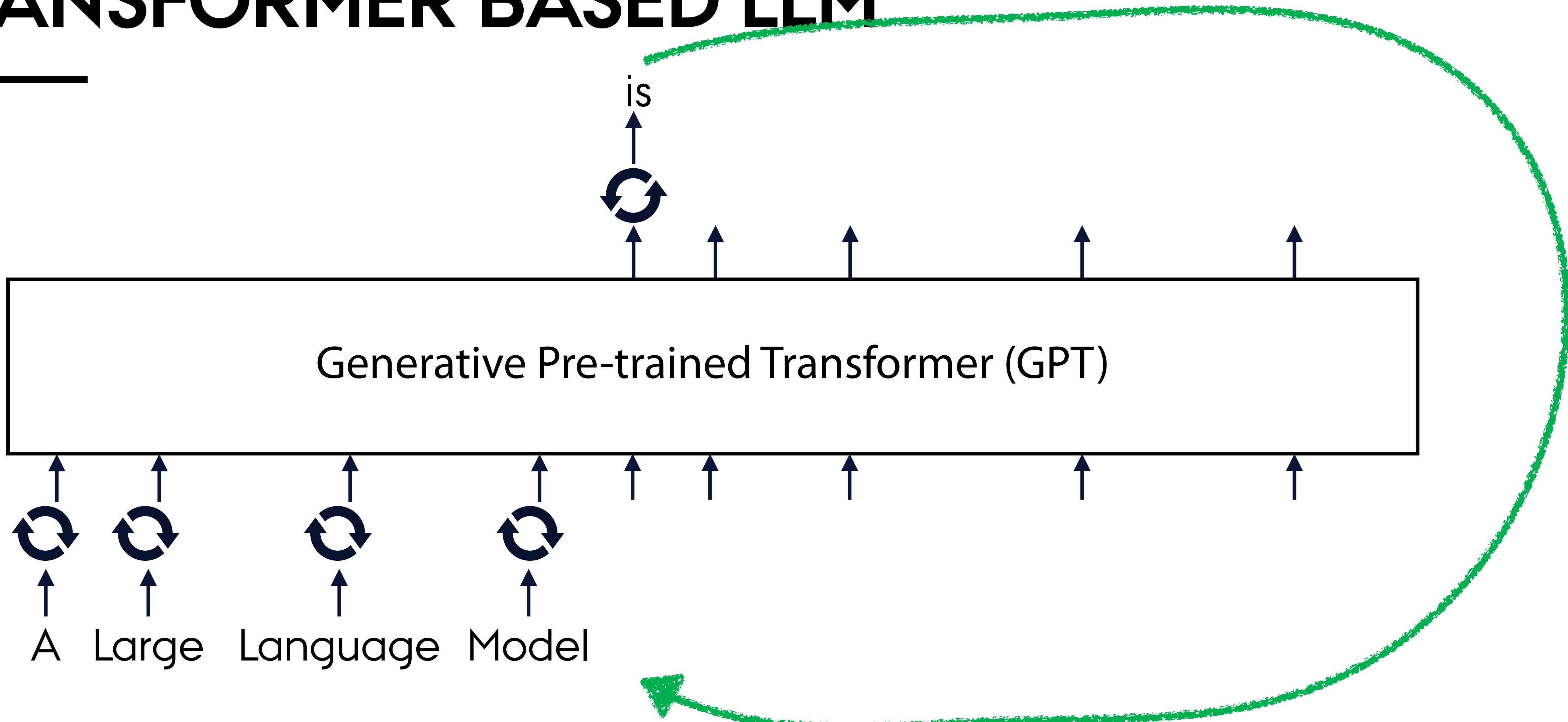
TRANSFORMER BASED LLM



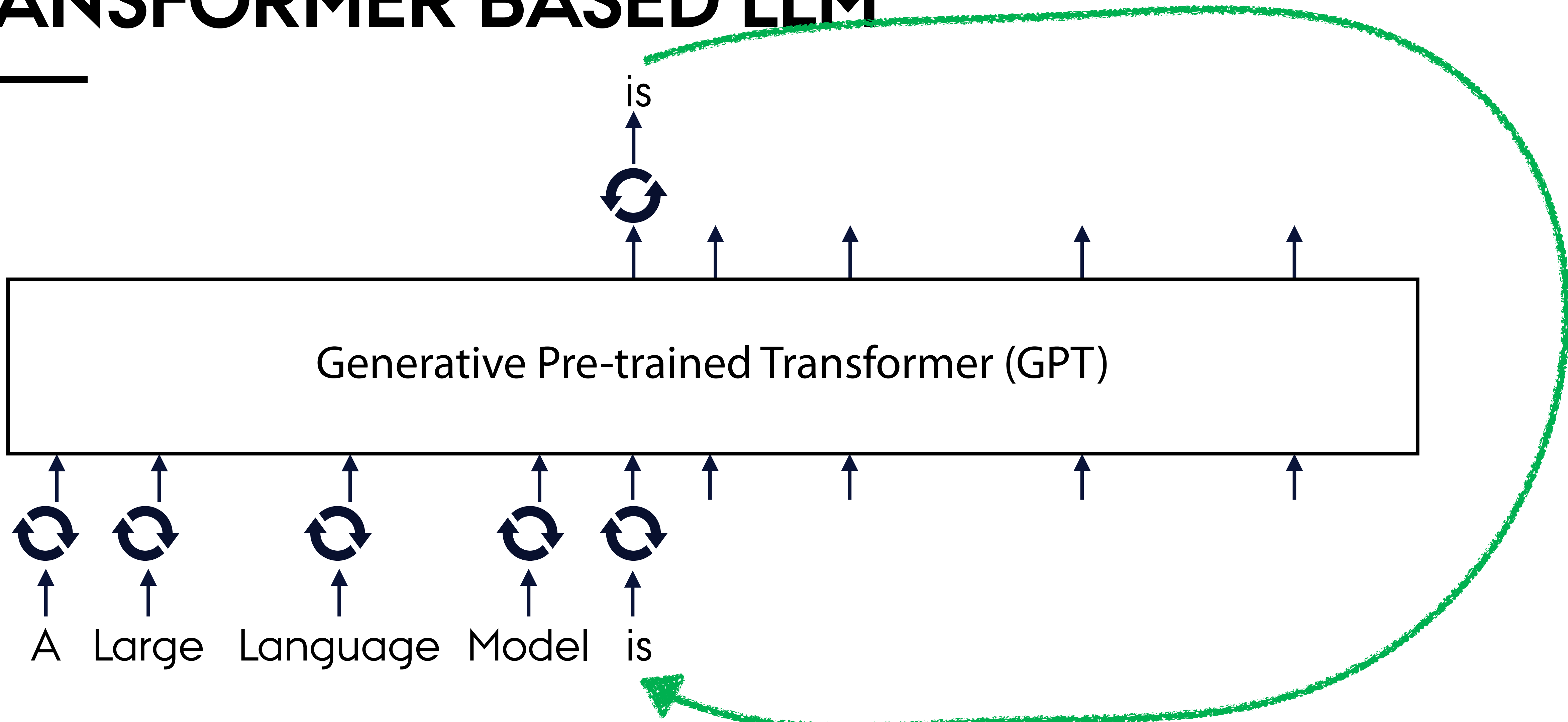
TRANSFORMER BASED LLM



TRANSFORMER BASED LLM

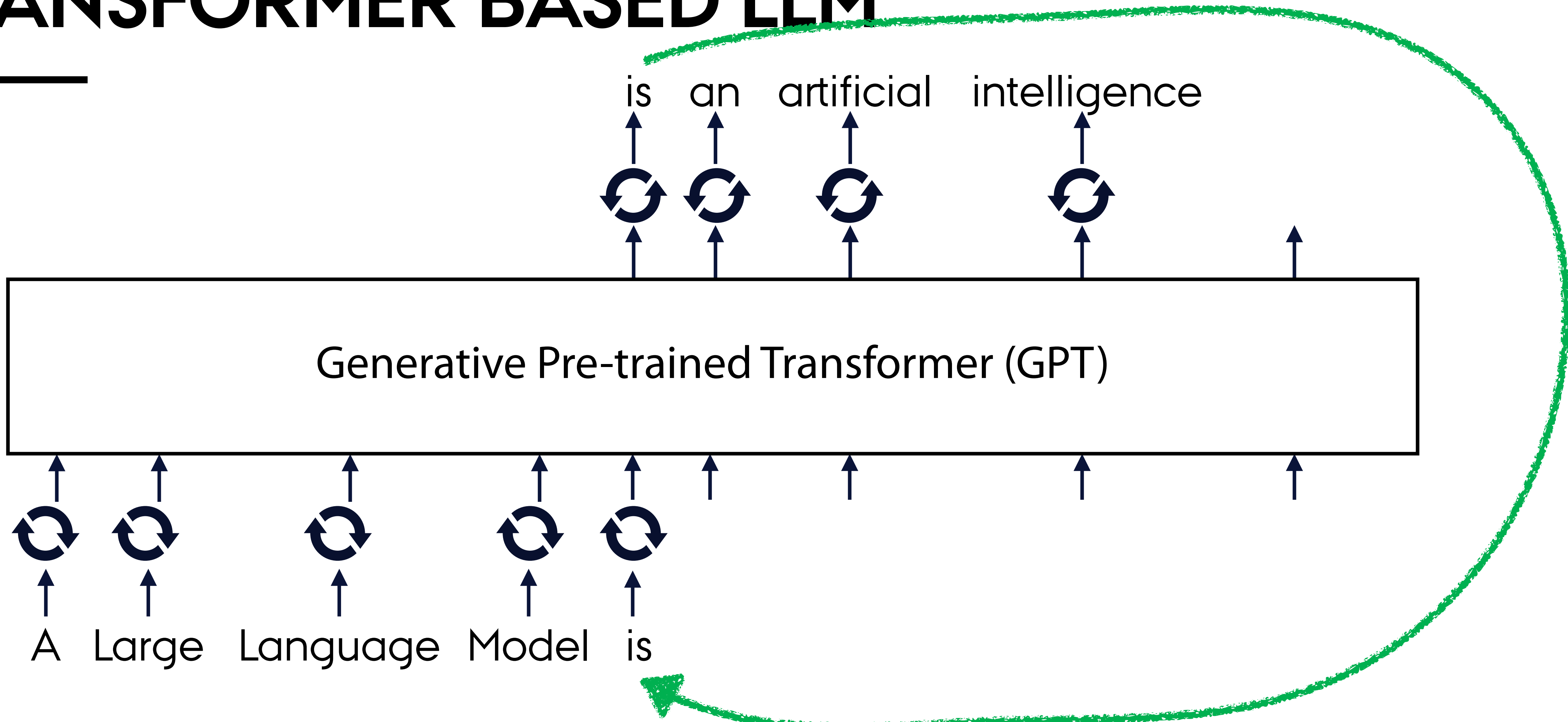


TRANSFORMER BASED LLM

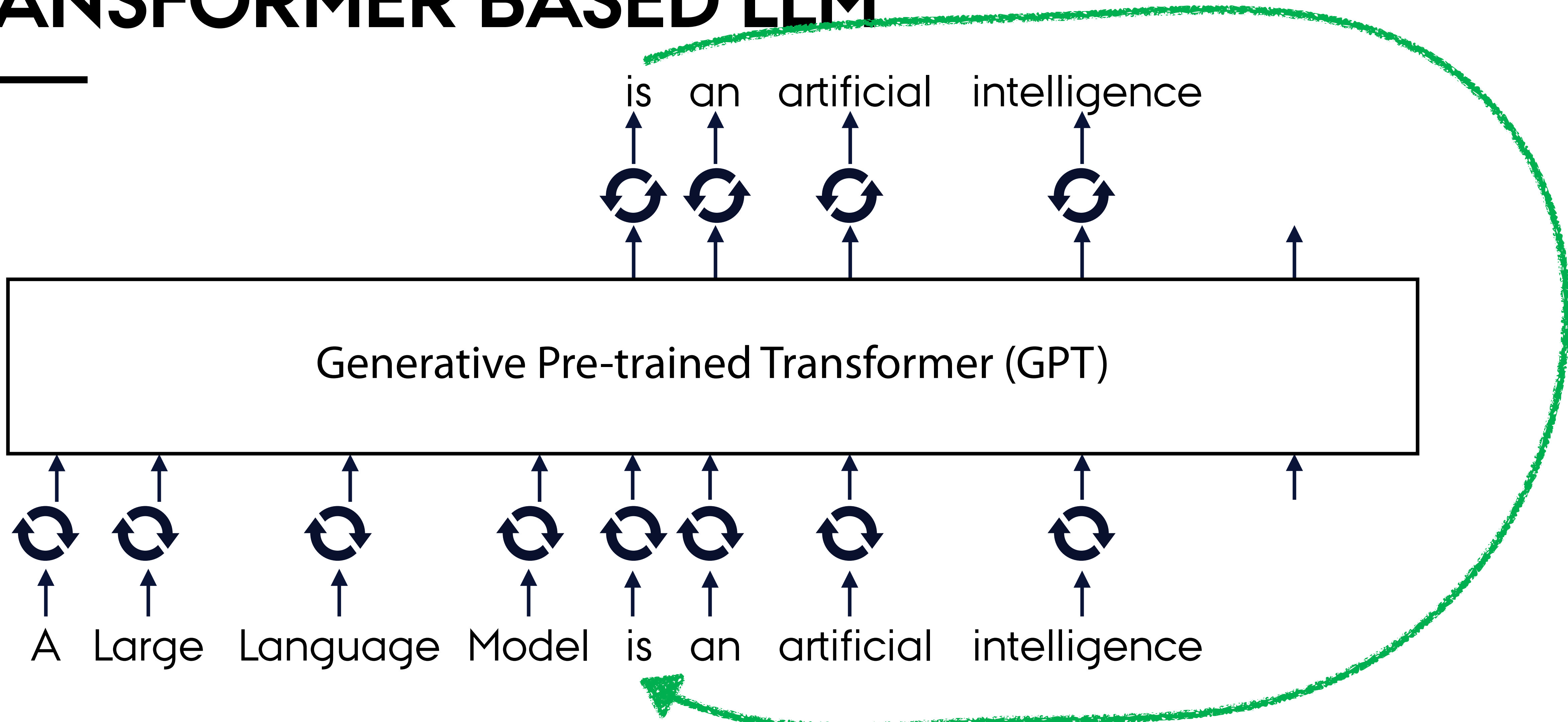


AARHUS BSS
DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

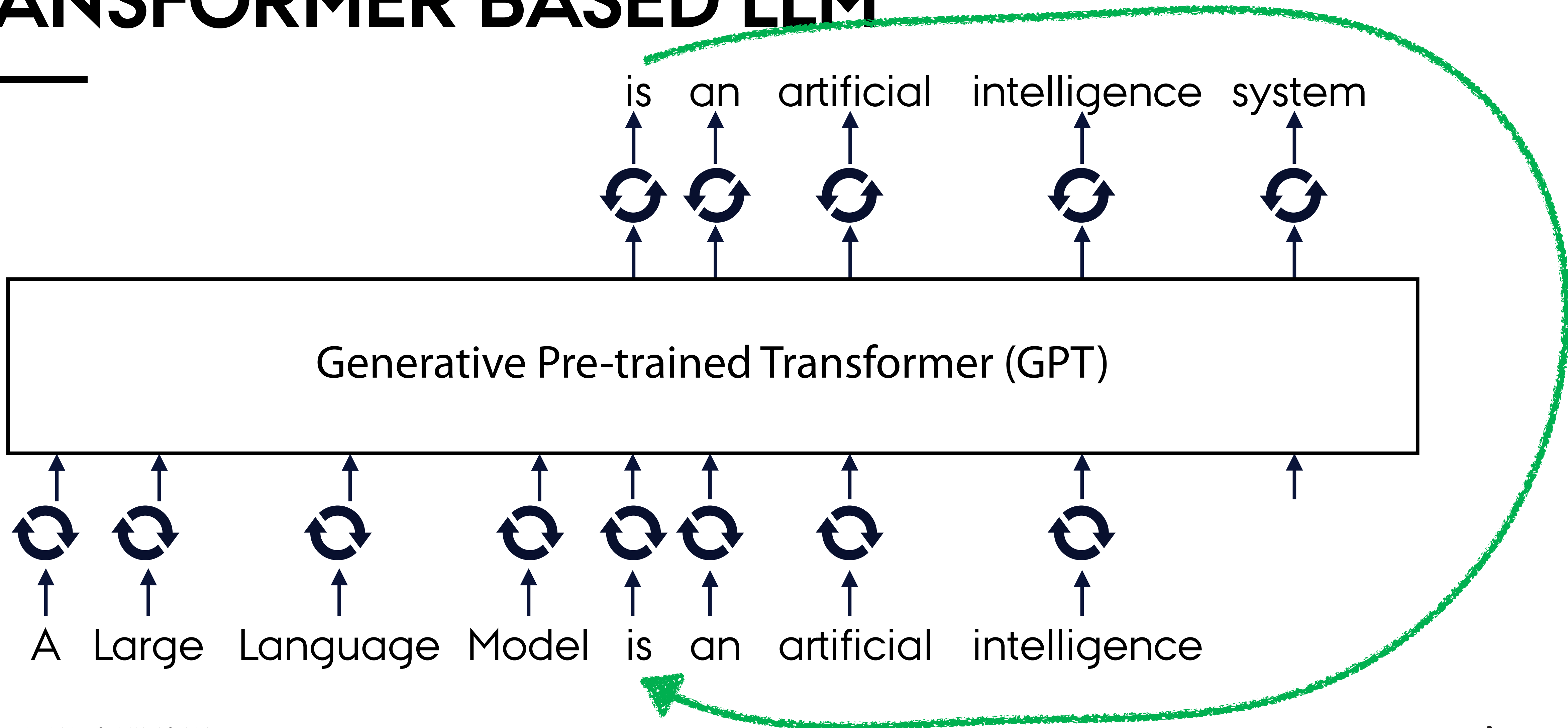
TRANSFORMER BASED LLM



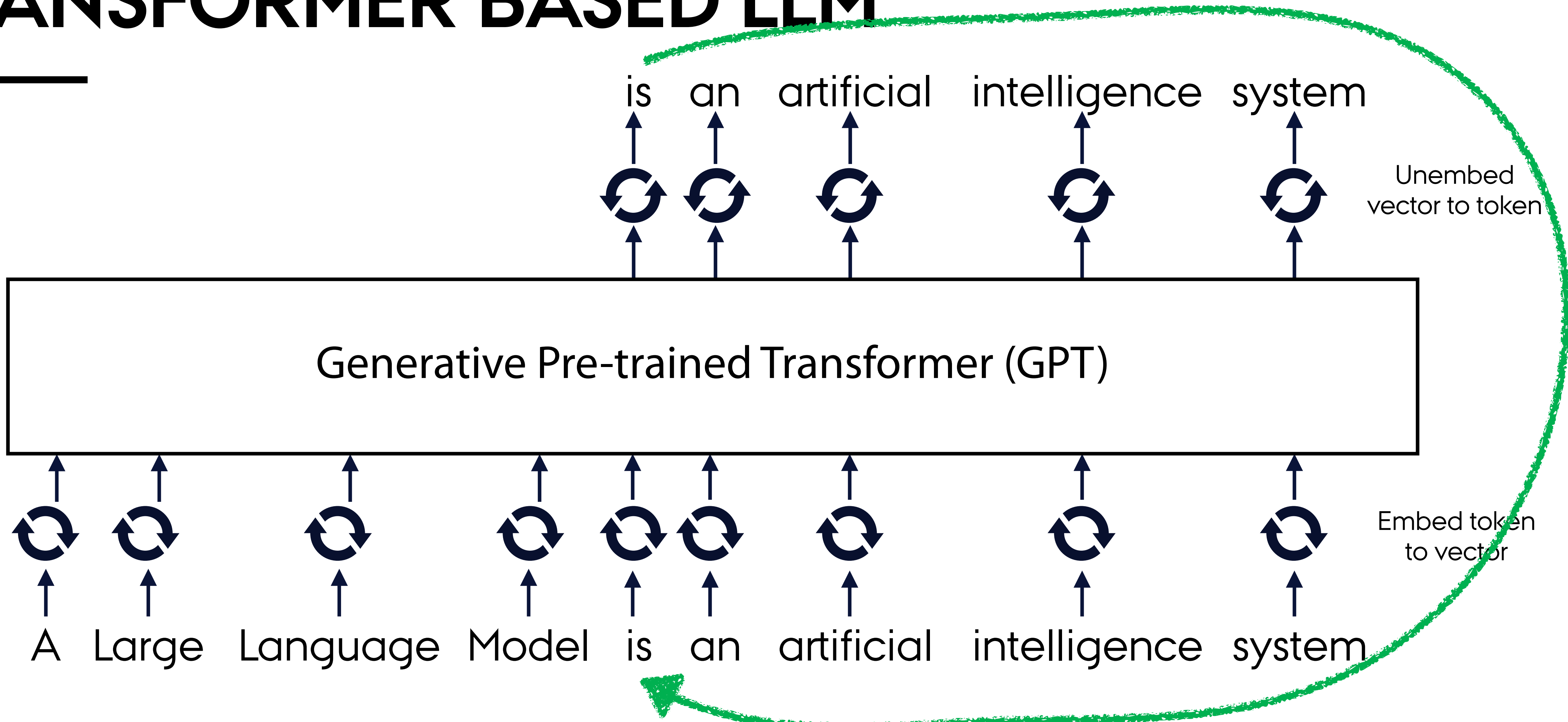
TRANSFORMER BASED LLM



TRANSFORMER BASED LLM



TRANSFORMER BASED LLM

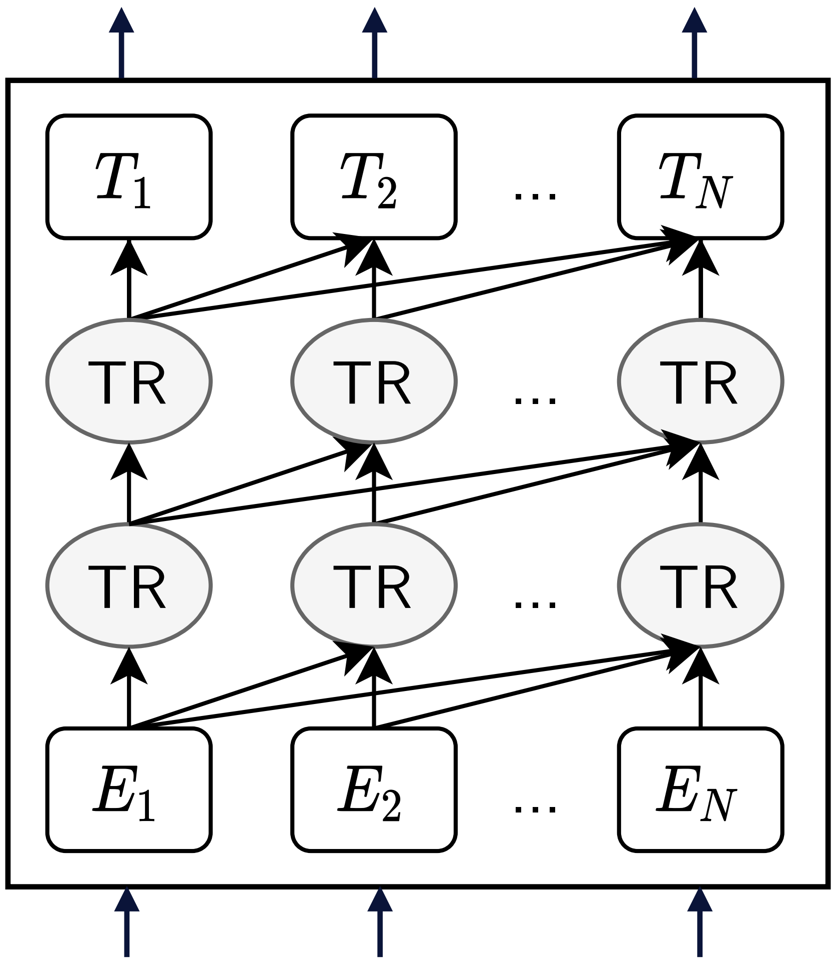


GENERATIVE PRE-TRAINED TRANSFORMER (GPT)

Radford et al. „Improving Language Understanding by Generative Pre-Training", **2018**
Radford et al. „Language Models are Unsupervised Multitask Learners", **2019**
Brown et al. „Language Models are Few-Shot Learners", **2020**

GENERATIVE PRE-TRAINED TRANSFORMER (GPT)

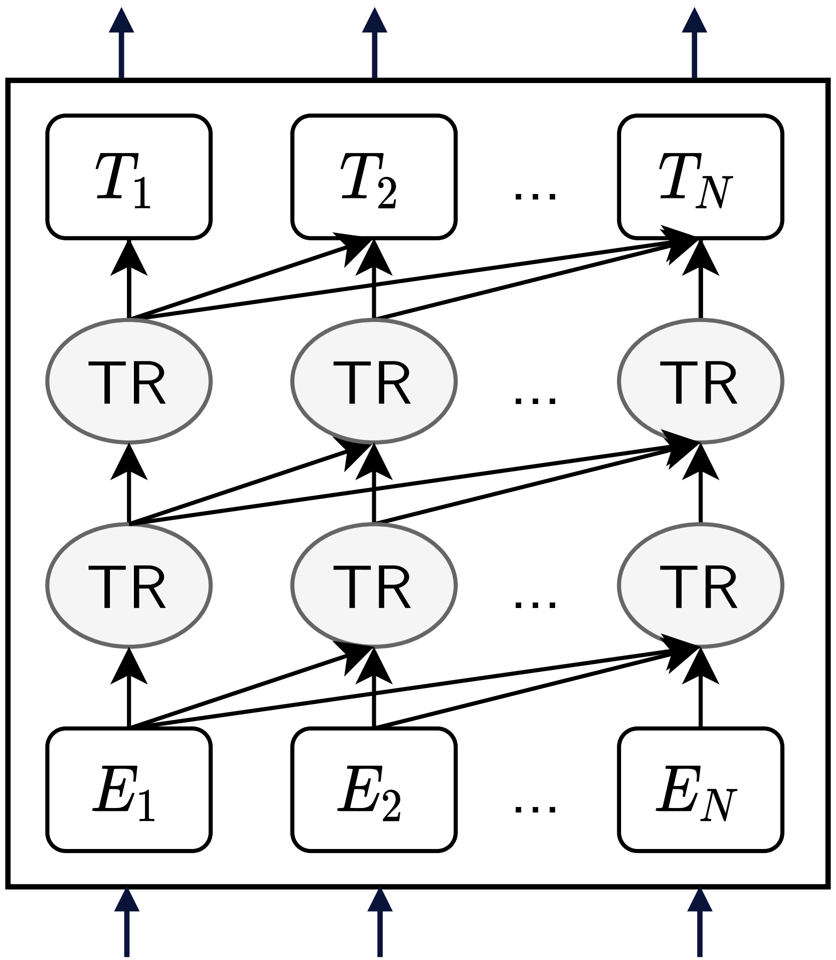
- Internal view of the GPT block we had on last slide



Radford et al. „Improving Language Understanding by Generative Pre-Training“, **2018**
Radford et al. „Language Models are Unsupervised Multitask Learners“, **2019**
Brown et al. „Language Models are Few-Shot Learners“, **2020**

GENERATIVE PRE-TRAINED TRANSFORMER (GPT)

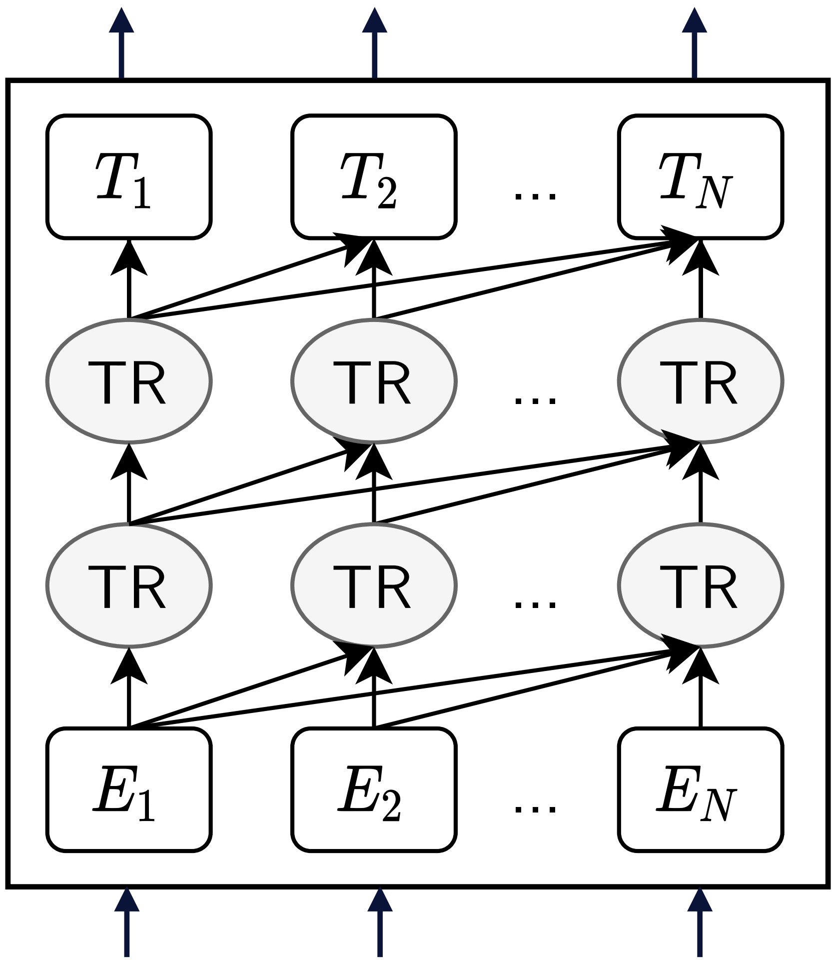
- Internal view of the GPT block we had on last slide
- Multiple so-called *transformer blocks* (TR) one after another per token



Radford et al. „Improving Language Understanding by Generative Pre-Training", **2018**
Radford et al. „Language Models are Unsupervised Multitask Learners", **2019**
Brown et al. „Language Models are Few-Shot Learners", **2020**

GENERATIVE PRE-TRAINED TRANSFORMER (GPT)

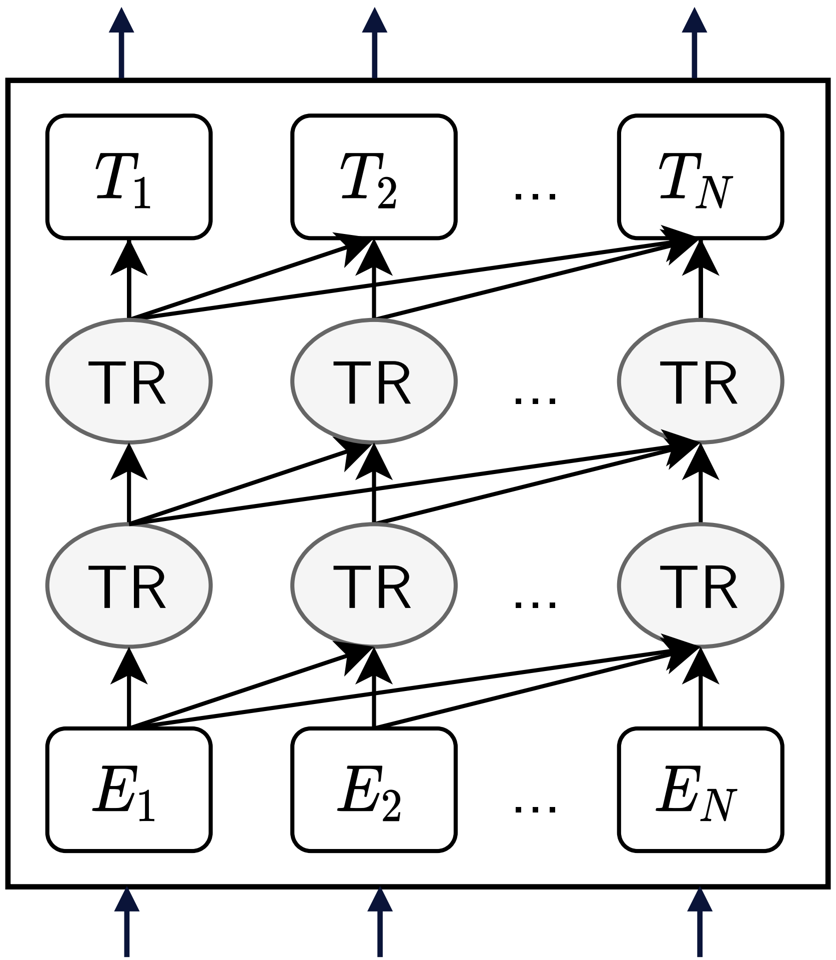
- Internal view of the GPT block we had on last slide
- Multiple so-called *transformer blocks* (TR) one after another per token
  - Attention used in each TR



Radford et al. „Improving Language Understanding by Generative Pre-Training“, **2018**
Radford et al. „Language Models are Unsupervised Multitask Learners“, **2019**
Brown et al. „Language Models are Few-Shot Learners“, **2020**

GENERATIVE PRE-TRAINED TRANSFORMER (GPT)

- Internal view of the GPT block we had on last slide
- Multiple so-called *transformer blocks* (TR) one after another per token
  - Attention used in each TR
  - Left-to-right connections between blocks → Focus on generation of text

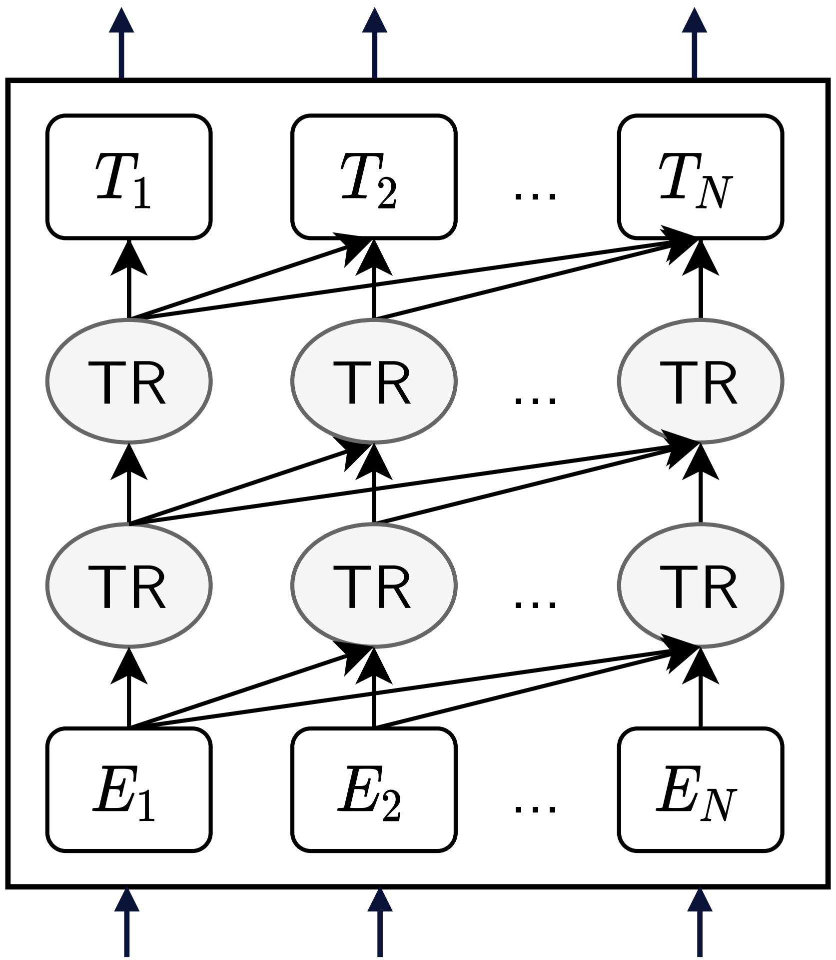


Radford et al. „Improving Language Understanding by Generative Pre-Training", **2018**
Radford et al. „Language Models are Unsupervised Multitask Learners", **2019**
Brown et al. „Language Models are Few-Shot Learners", **2020**

GENERATIVE PRE-TRAINED TRANSFORMER (GPT)

- Internal view of the GPT block we had on last slide
- Multiple so-called *transformer blocks* (TR) one after another per token
  - Attention used in each TR
  - Left-to-right connections between blocks → Focus on generation of text

How to get the statistical values → so-called *weights* in each TR block



Radford et al. „Improving Language Understanding by Generative Pre-Training", **2018**
Radford et al. „Language Models are Unsupervised Multitask Learners", **2019**
Brown et al. „Language Models are Few-Shot Learners", **2020**

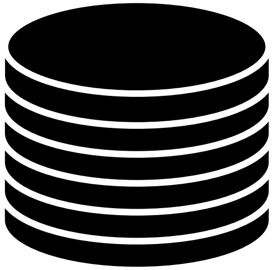
MACHINE LEARNING

How to obtain the statistical values used by the model?

AARHUS BSS
DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

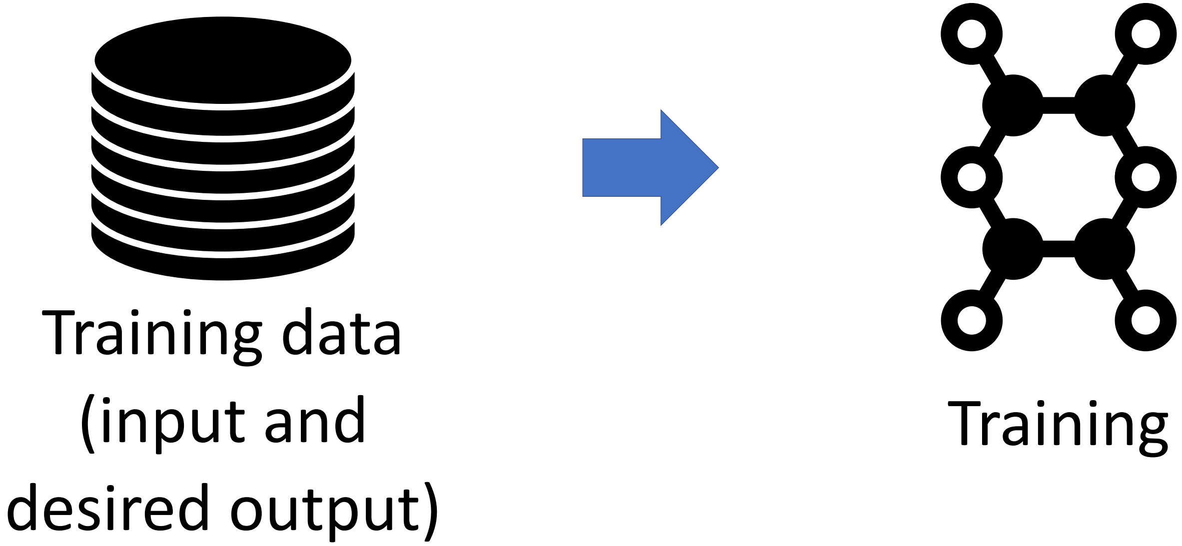
MAGNUS BENDER
ASSISTANT PROFESSOR

MACHINE LEARNING

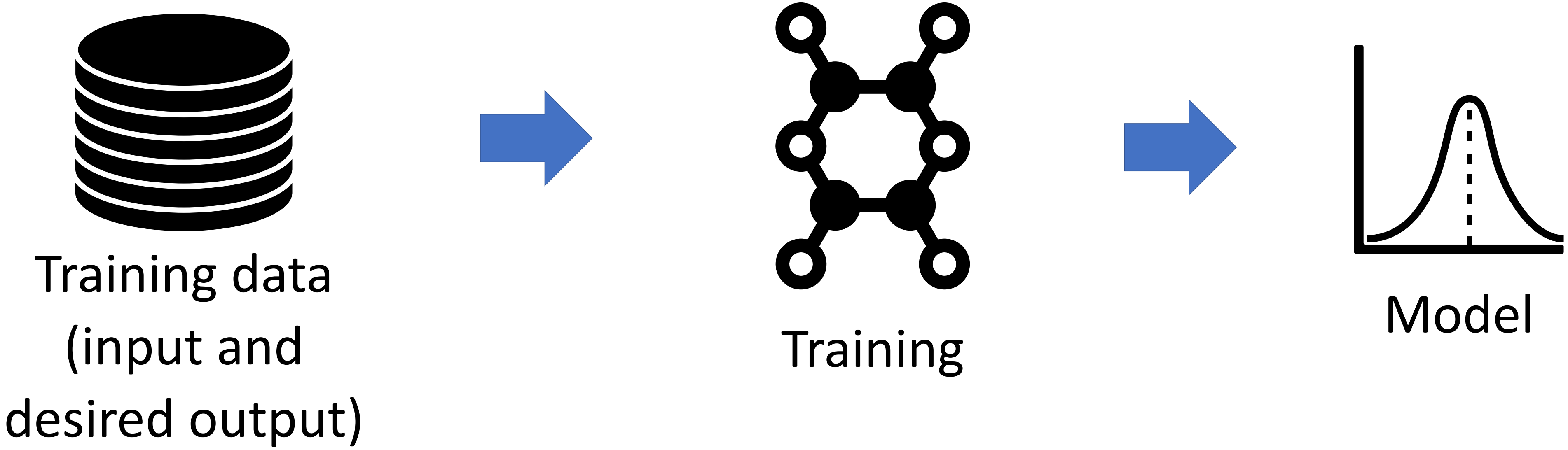


Training data
(input and
desired output)

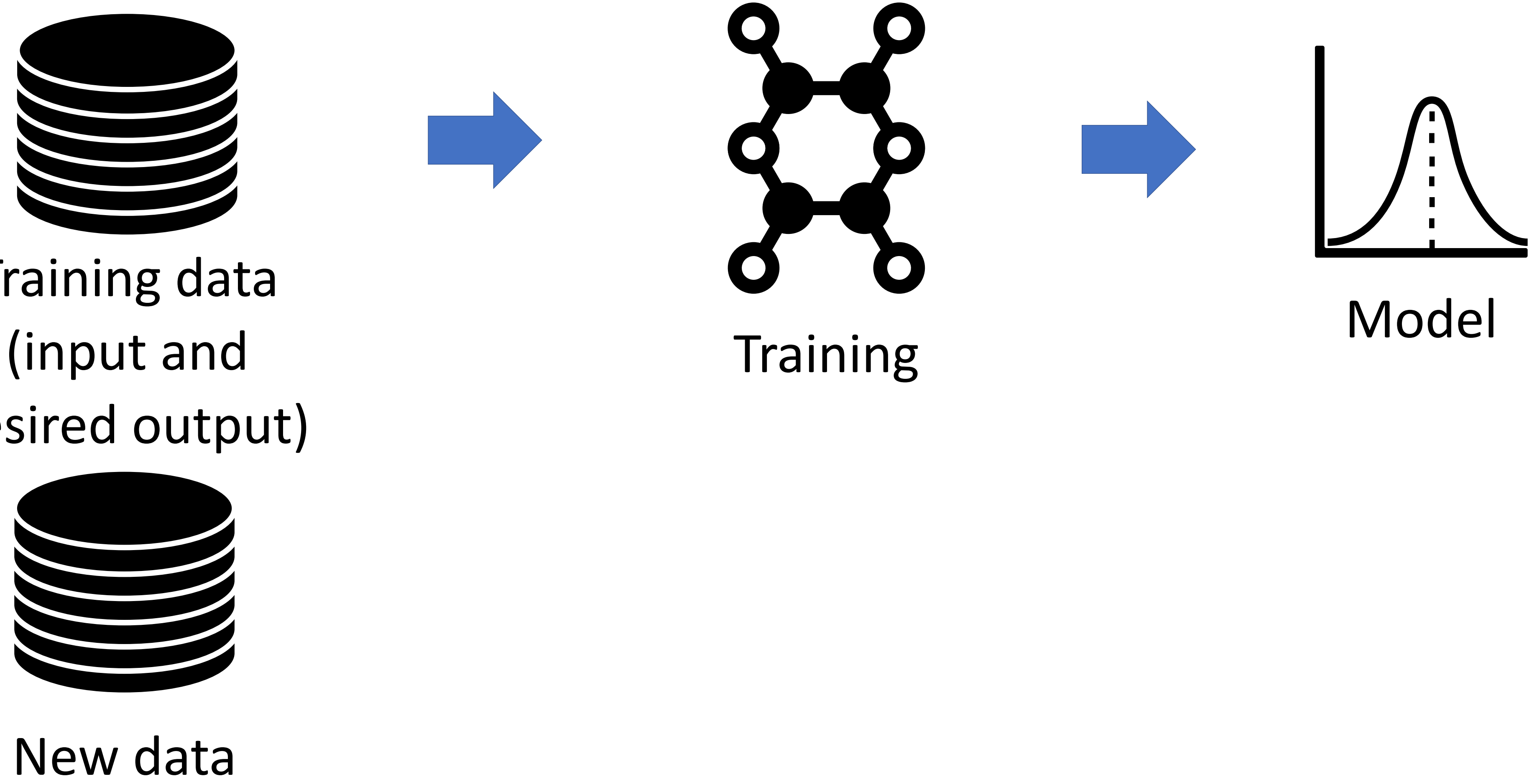
MACHINE LEARNING



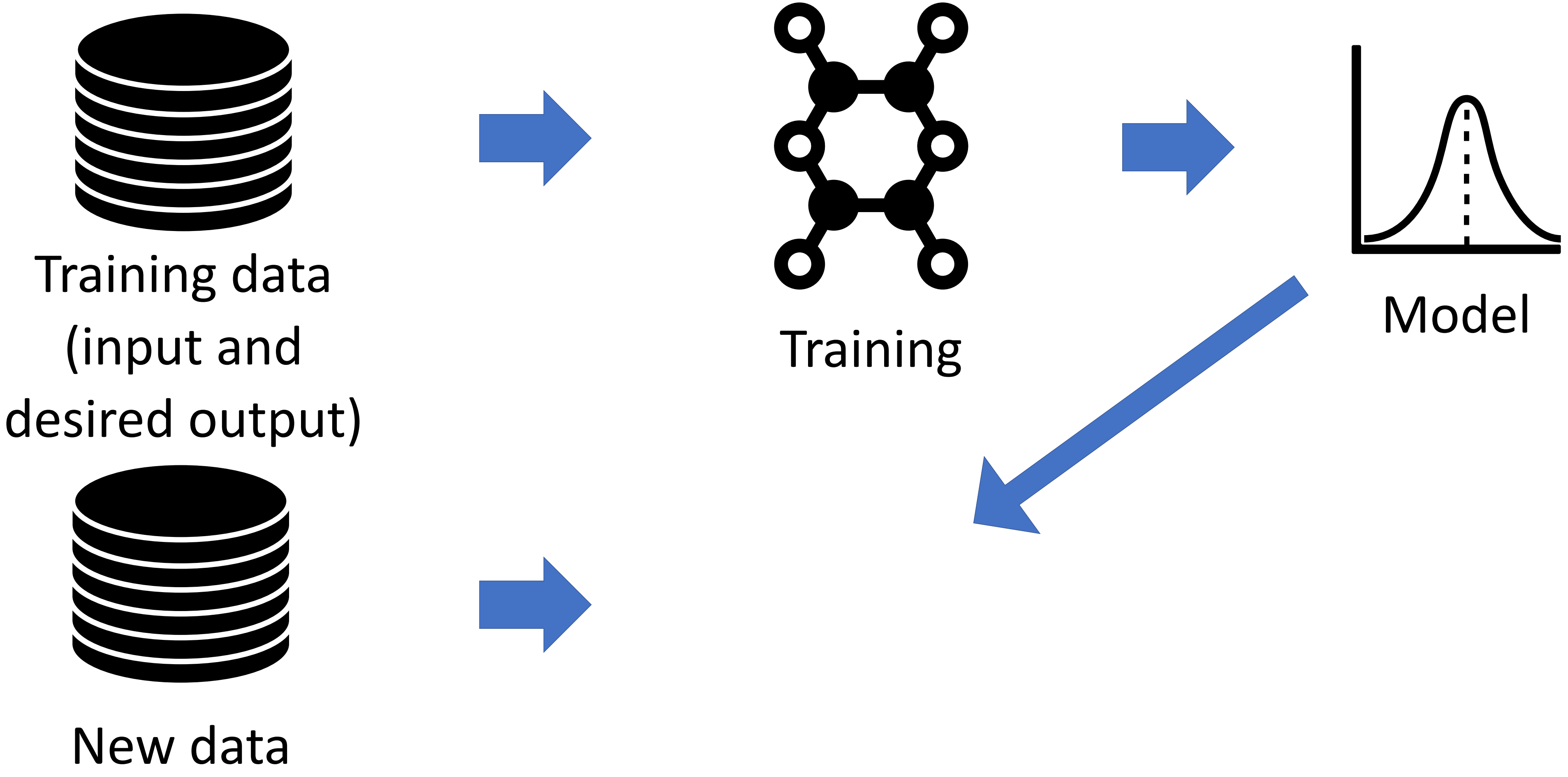
MACHINE LEARNING



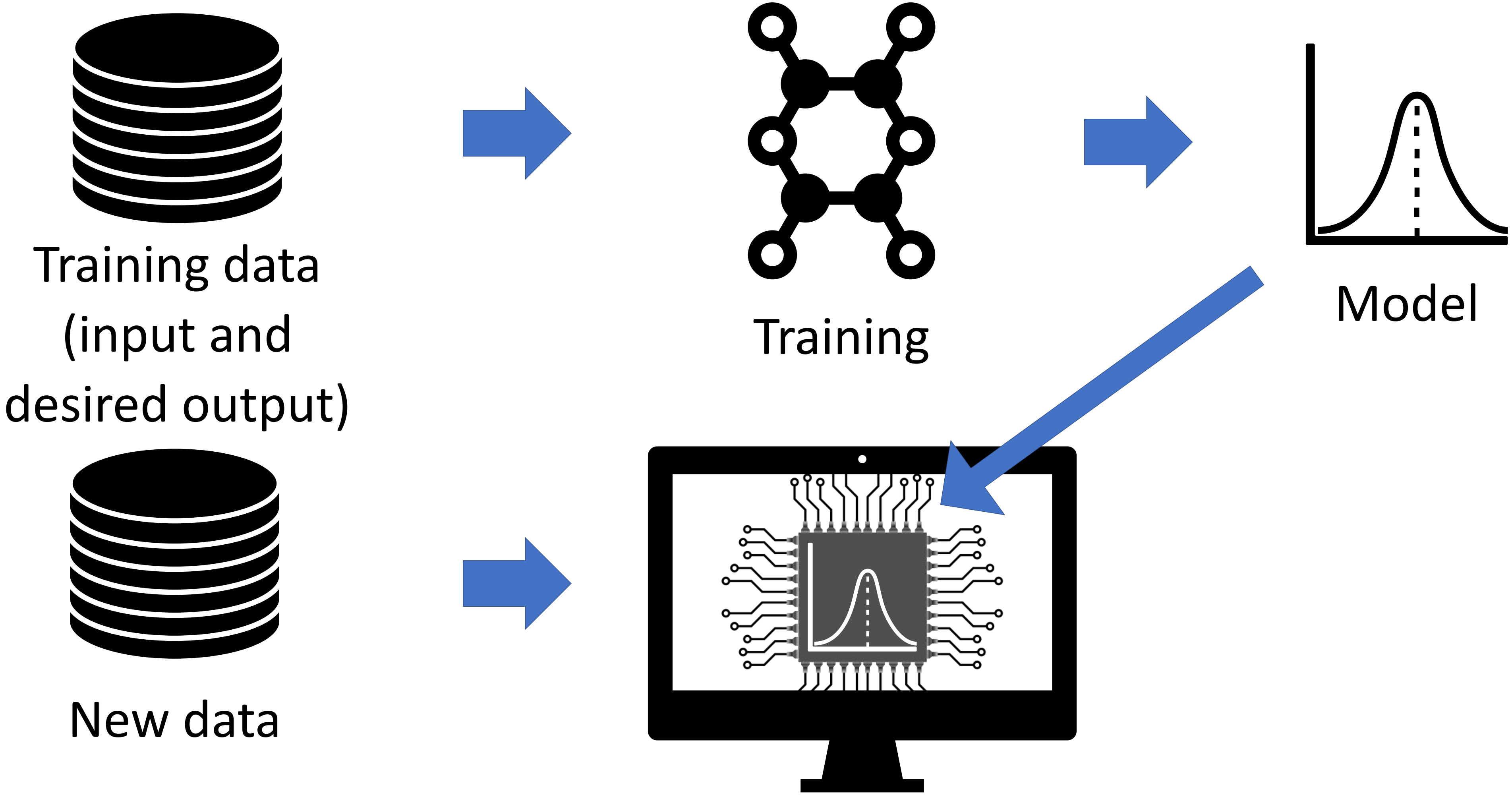
MACHINE LEARNING



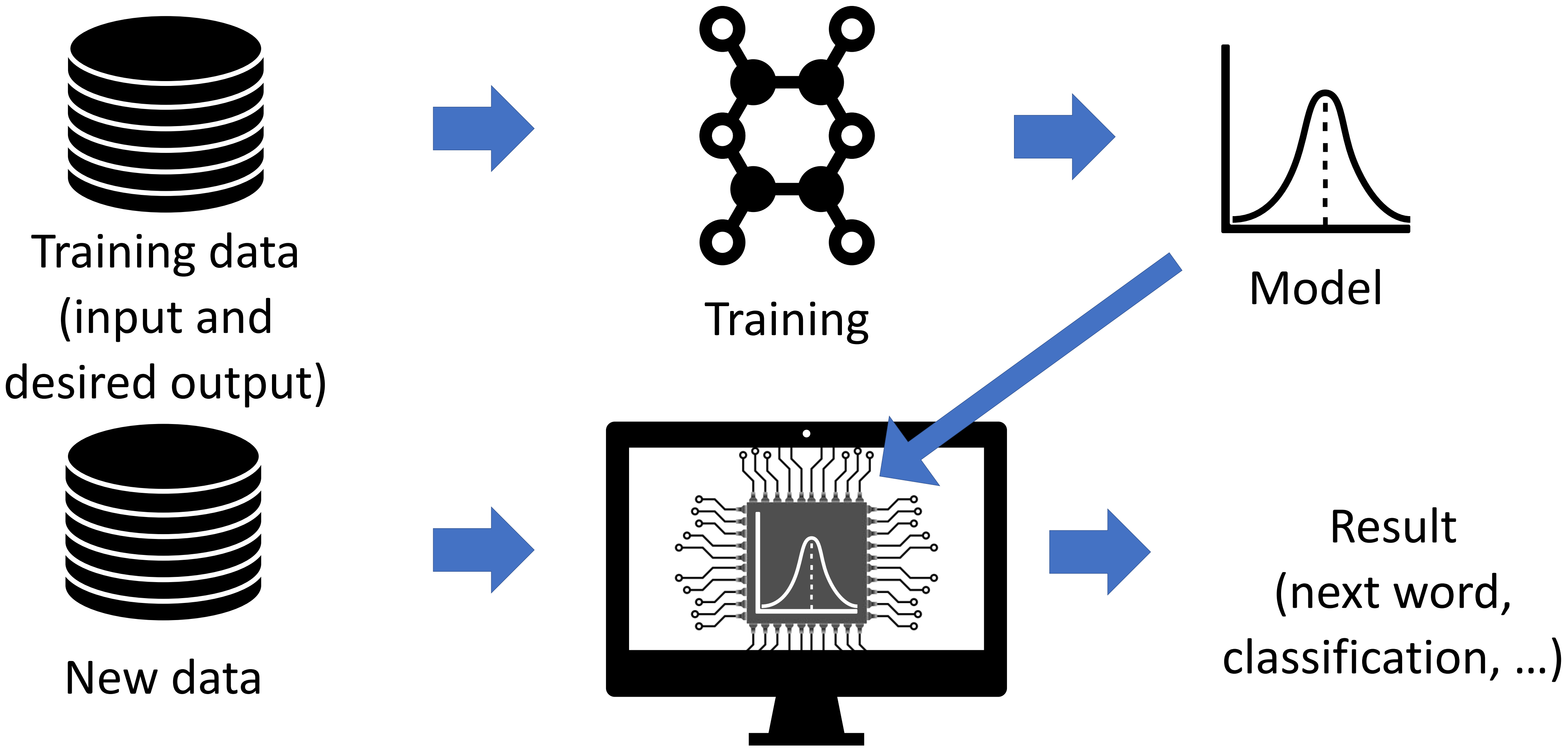
MACHINE LEARNING



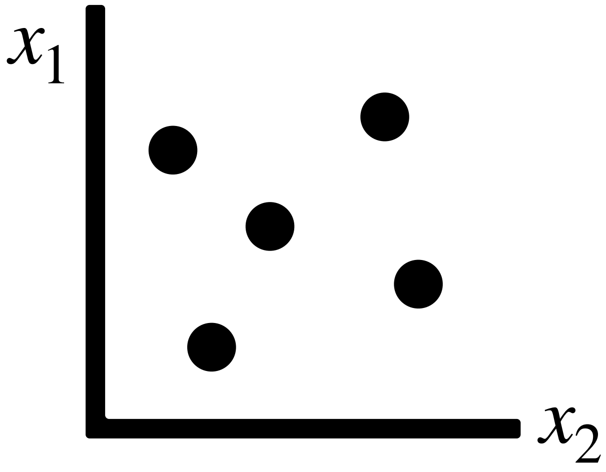
MACHINE LEARNING



MACHINE LEARNING

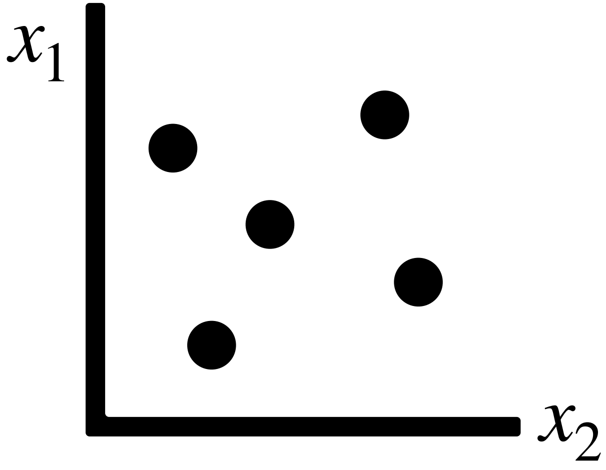


SIMPLE PERCEPTRON



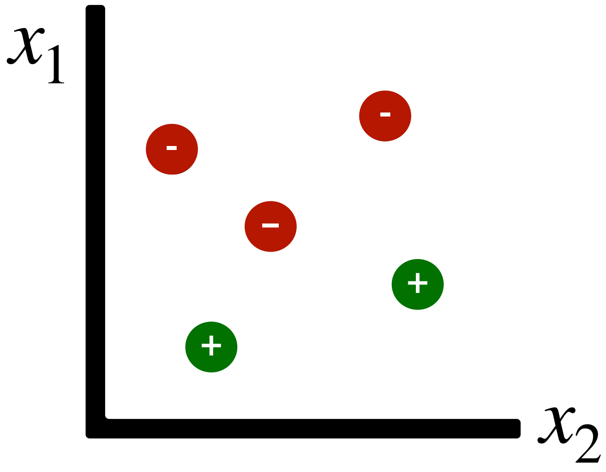
SIMPLE PERCEPTRON

- Goal:
  Classify points as „+“ and „-“



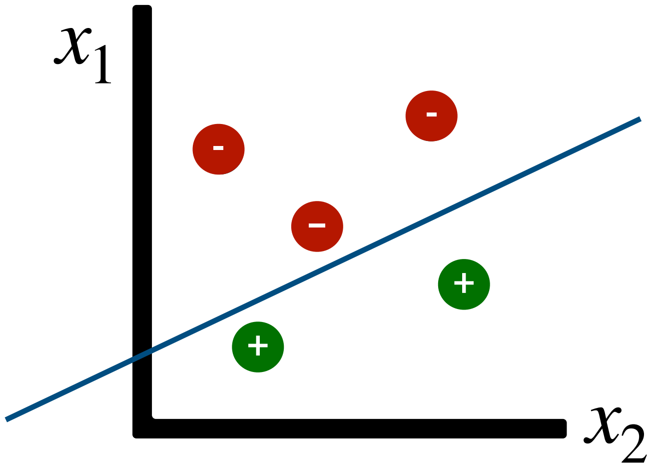
SIMPLE PERCEPTRON

- Goal:
  Classify points as „+“ and „-“
- Training data:
  5 points with known label



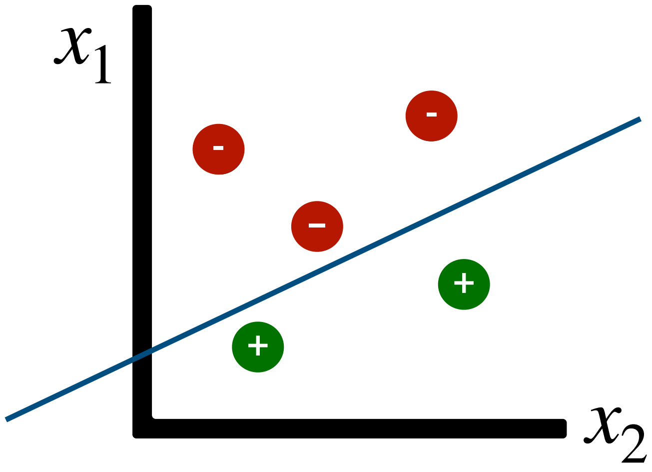
SIMPLE PERCEPTRON

- Goal:
  Classify points as „+“ and „-“
- Training data:
  5 points with known label
- Model:
  Straight line



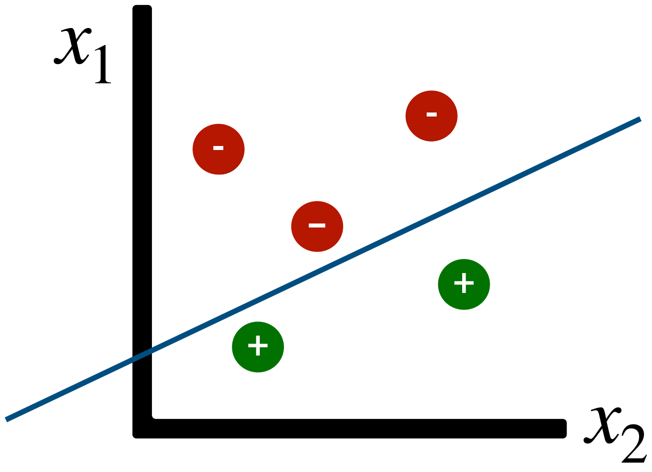
SIMPLE PERCEPTRON

- Goal:
  Classify points as „+“ and „-“
- Training data:
  5 points with known label
- Model:
  Straight line
  - $f(x) = ax + b$



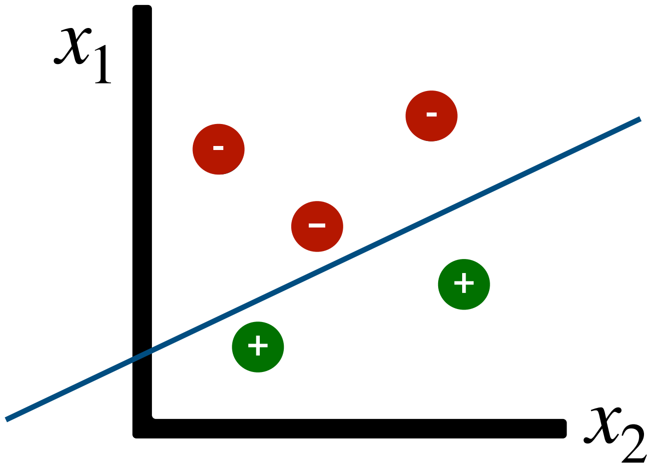
SIMPLE PERCEPTRON

- Goal:
  Classify points as „+“ and „-“
- Training data:
  5 points with known label
- Model:
  Straight line
  - $f(x) = ax + b$
  - Two parameters (decimal numbers): a and b



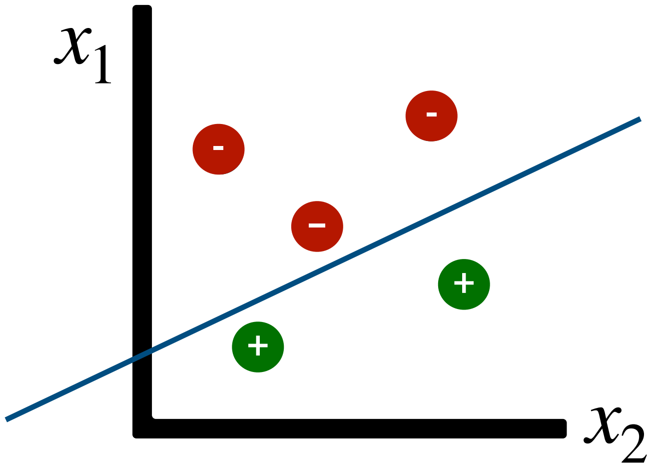
SIMPLE PERCEPTRON

- Goal:
  Classify points as „+“ and „-“
- Training data:
  5 points with known label
- Model:
  Straight line
    - $f(x) = ax + b$
    - Two parameters (decimal numbers): a and b
- Classification:
  Is point below or above line?



SIMPLE PERCEPTRON

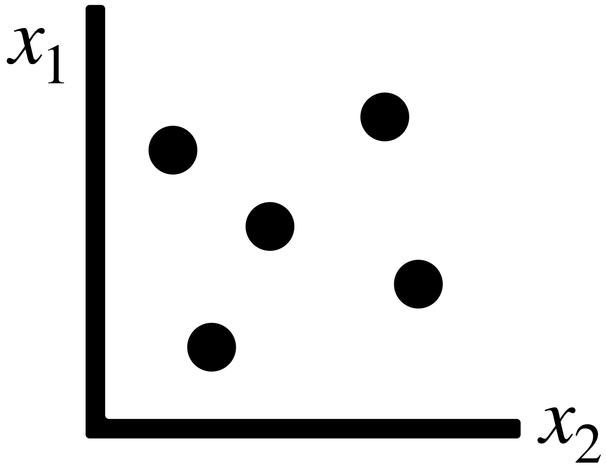
- Goal:
  Classify points as „+“ and „-“
- Training data:
  5 points with known label
- Model:
  Straight line
  - $f(x) = ax + b$
  - Two parameters (decimal numbers): a and b
- Classification:
  Is point below or above line?



- Some call such perceptron an „(artificial) neural neuron“
- Multiple perceptrons then form an „(artificial) neural network“

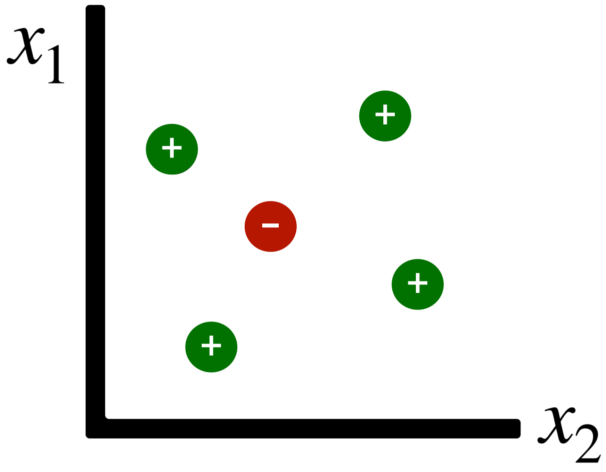
MORE DIFFICULT INSTANCE

- Same goal: Classify points as „+“ and „-“



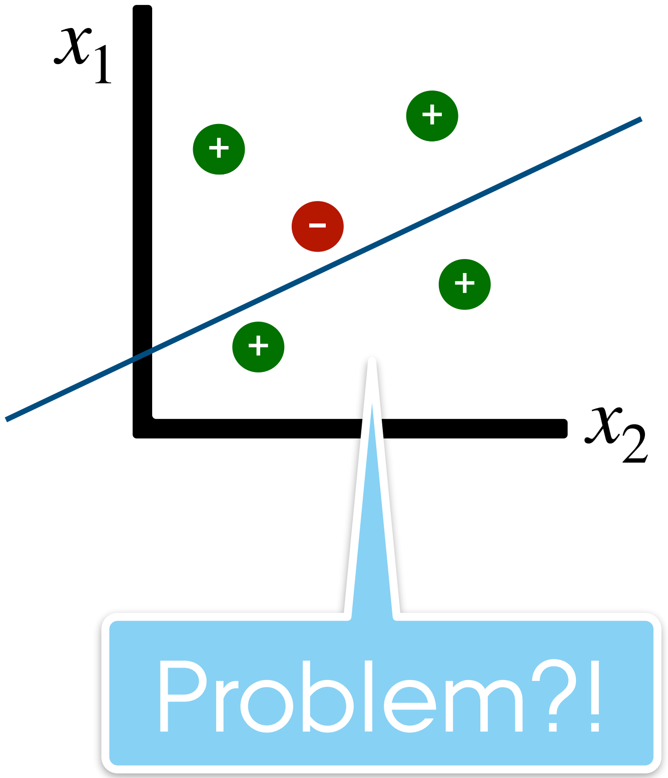
MORE DIFFICULT INSTANCE

- Same goal: Classify points as „+“ and „-“
- Training data: 5 points with known label



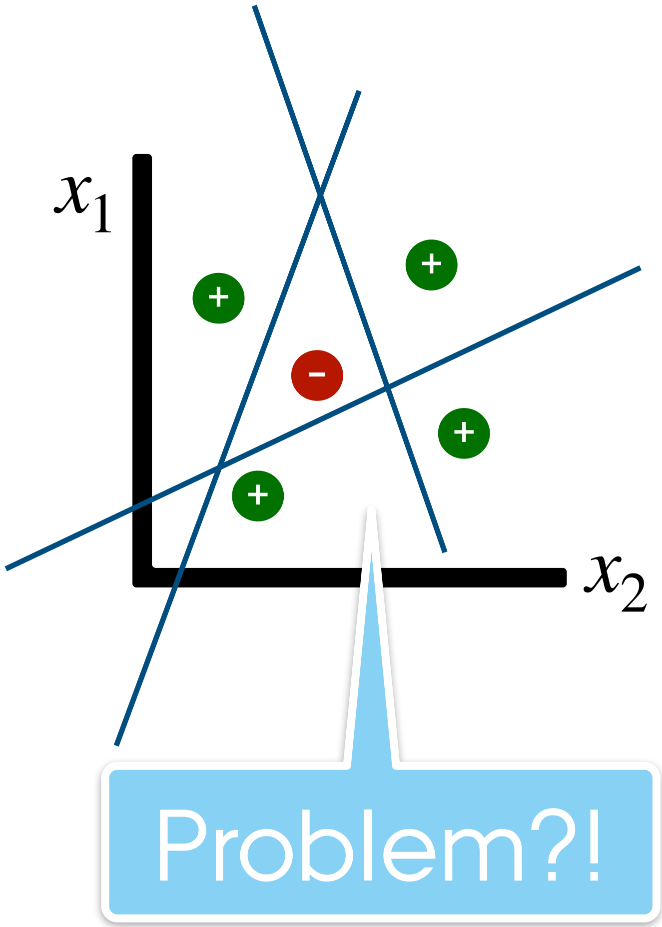
MORE DIFFICULT INSTANCE

- Same goal: Classify points as „+" and „-"
- Training data: 5 points with known label
- Model:
  - Straight line?



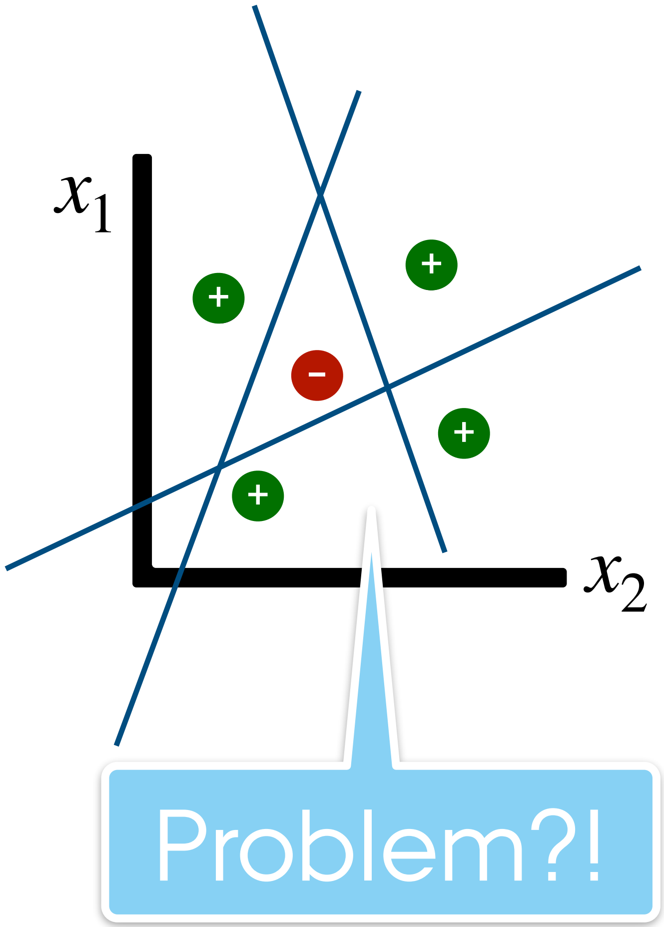
MORE DIFFICULT INSTANCE

- Same goal: Classify points as „+“ and „-“
- Training data: 5 points with known label
- Model:
  - Straight line?
  - Multiple lines!



MORE DIFFICULT INSTANCE

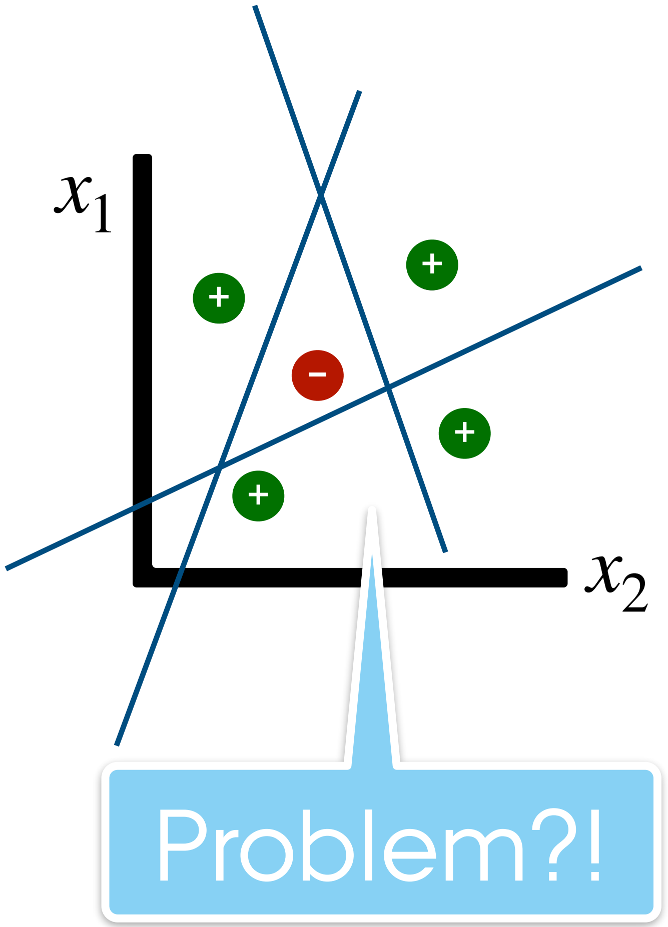
- Same goal: Classify points as „+“ and „-“
- Training data: 5 points with known label
- Model:
  - Straight line?
  - Multiple lines!
- Classification:
  - For each line: Is point below or above line?



MORE DIFFICULT INSTANCE

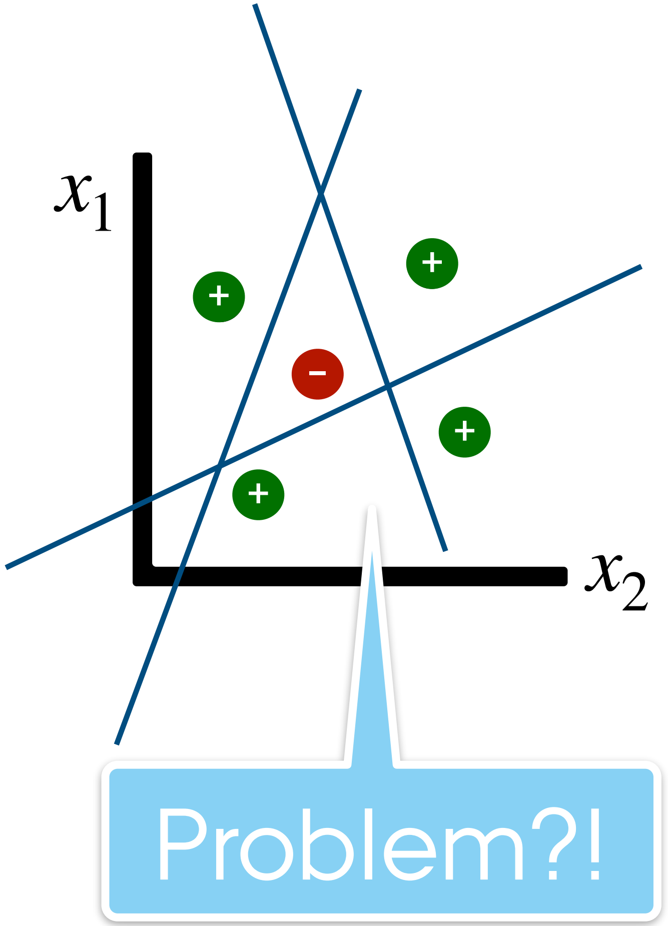
- Same goal: Classify points as „+“ and „-“
- Training data: 5 points with known label
- Model:
  - Straight line?
  - Multiple lines!
- Classification:
  - For each line: Is point below or above line?

- More parameters:
  - Model: three lines with two parameters → six numbers
  - Classification: below or above for each line



MORE DIFFICULT INSTANCE

- Same goal: Classify points as „+" and „-"
- Training data: 5 points with known label
- Model:
  - Straight line?
  - Multiple lines!
- Classification:
  - For each line: Is point below or above line?

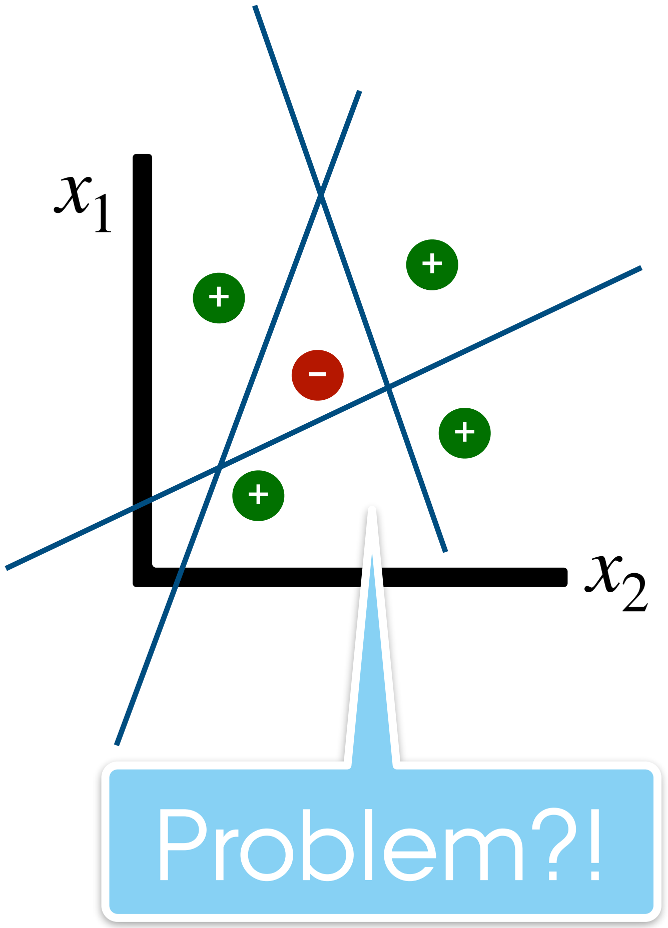


- More parameters:
  - Model: three lines with two parameters → six numbers
  - Classification: below or above for each line

➡ Still easily possible to solve

MORE DIFFICULT INSTANCE

- Same goal: Classify points as „+“ and „-“
- Training data: 5 points with known label
- Model:
    - Straight line?
    - Multiple lines!
- Classification:
    - For each line: Is point below or above line?



- More parameters:
    - Model: three lines with two parameters → six numbers
    - Classification: below or above for each line

➡ Still easily possible to solve

➡ More parameters, higher capacity

APPLIED TO LLMS

APPLIED TO LLMS

- LLMs have billions of parameters (decimal numbers)

APPLIED TO LLMS

- LLMs have billions of parameters (decimal numbers)
- Trained using machine learning on natural language texts

APPLIED TO LLMS

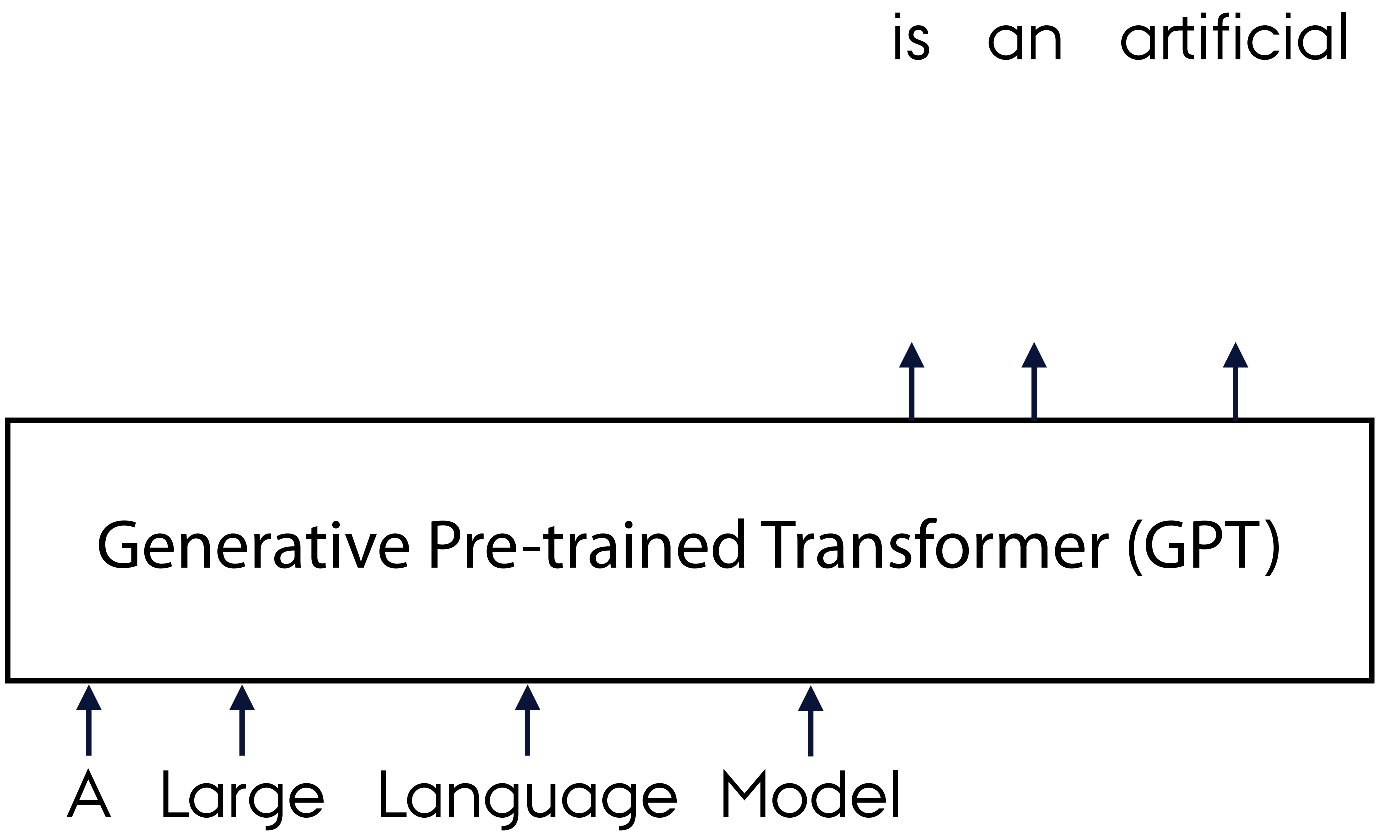
- LLMs have billions of parameters (decimal numbers)
- Trained using machine learning on natural language texts
  - Input part of a text having the full text, i.e., the desired output

is an artificial

A Large Language Model

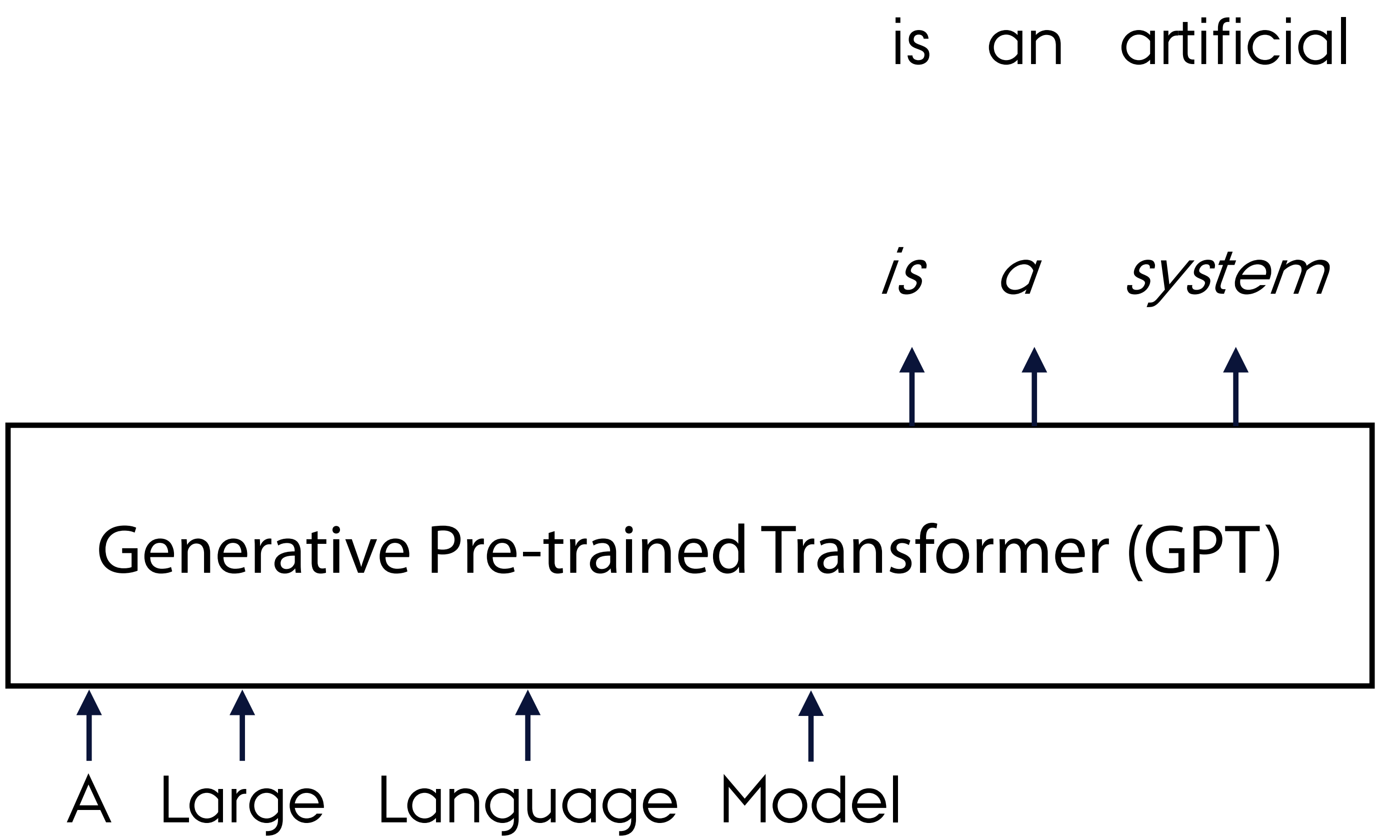
APPLIED TO LLMS

- LLMs have billions of parameters (decimal numbers)
- Trained using machine learning on natural language texts
    - Input part of a text having the full text, i.e., the desired output



APPLIED TO LLMS

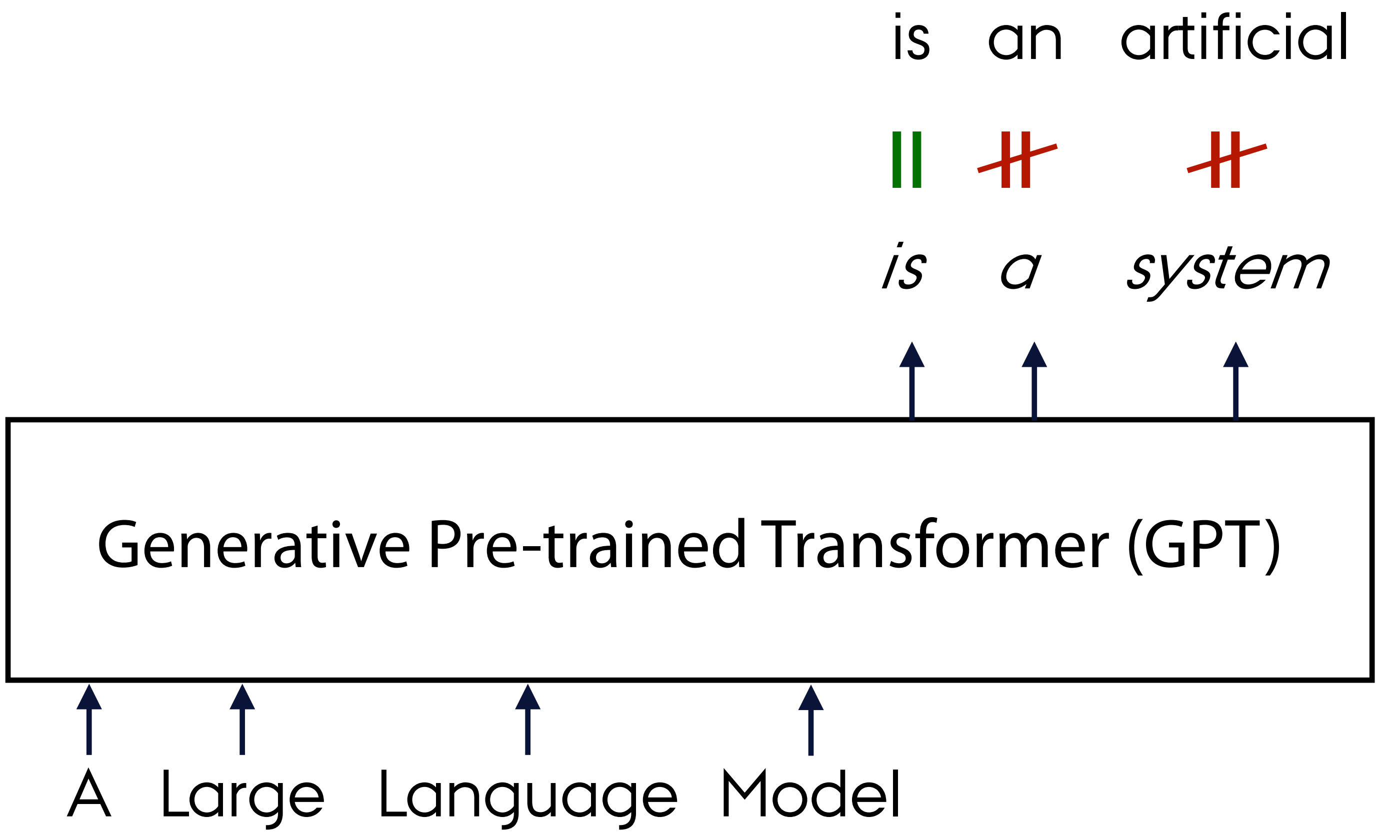
- LLMs have billions of parameters (decimal numbers)
- Trained using machine learning on natural language texts
  - Input part of a text having the full text, i.e., the desired output



DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

APPLIED TO LLMS

- LLMs have billions of parameters (decimal numbers)
- Trained using machine learning on natural language texts
  - Input part of a text having the full text, i.e., the desired output
  - Possible to assess the error and reduce it by improving parameters



BACK TO LLMS & CHATBOTS

Implications and limitations

AARHUS BSS
DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

MAGNUS BENDER
ASSISTANT PROFESSOR

LLMS ARE INCREDIBLY

CAPABLE

Hey, can you describe the differences between the epistemologies of Plato, Aristotle and Socrates? Explain them to me like I am five years old.

OI **llama3.1:70b**

Oh boy, are you ready for an adventure in ancient Greek philosophy?

Imagine we're on a journey to discover what knowledge is. We'll meet three friends: Plato, Aristotle, and Socrates. Each of them has a special way of thinking about how we know things.

Socrates (the Curious Questioner)

Socrates says that the best way to learn is by asking questions! He thinks that true wisdom is realizing how little we actually know. It's like playing "I Spy" – you have to ask lots of questions to figure out what something is.

Imagine you see a cat, and someone asks, "What's that?" You might say, "It's a cat!" But Socrates would ask, "How do you know it's a cat? What makes it a cat?" He wants us to think deeply about our answers and not just accept things at face value.

Plato (the Dreamer)

Plato has a big imagination. He thinks that the world we see around us is like a shadow or a copy of something more perfect and real, which exists in another realm called the "World of Forms." It's like he's saying, "You know how you can draw a picture of a cat? Well, there's a perfect 'Cat-ness' out there that all cats are trying to be like."

For Plato, knowledge is about understanding these perfect forms or ideas. He thinks we can't really know things just by looking at the world around us; we need to think deeply and imagine what the perfect version of something might be.

Aristotle (the Observer)

Aristotle is more down-to-earth than Plato. He says that knowledge comes from observing the world around us, using our senses, and figuring out how things work. It's like being a scientist – you collect information, make observations, and try to understand patterns.

Imagine you're watching a cat chase its tail. Aristotle would say, "Hey, let me study this! What does the cat do? How does it move?" By observing and thinking about what we see, we can gain knowledge about the world.

So, in short:

- Socrates asks questions to make us think deeply about our answers.
- Plato imagines perfect forms or ideas that exist beyond the physical world.
- Aristotle observes the world around us to understand how things work.

Which one do you think is most like you?

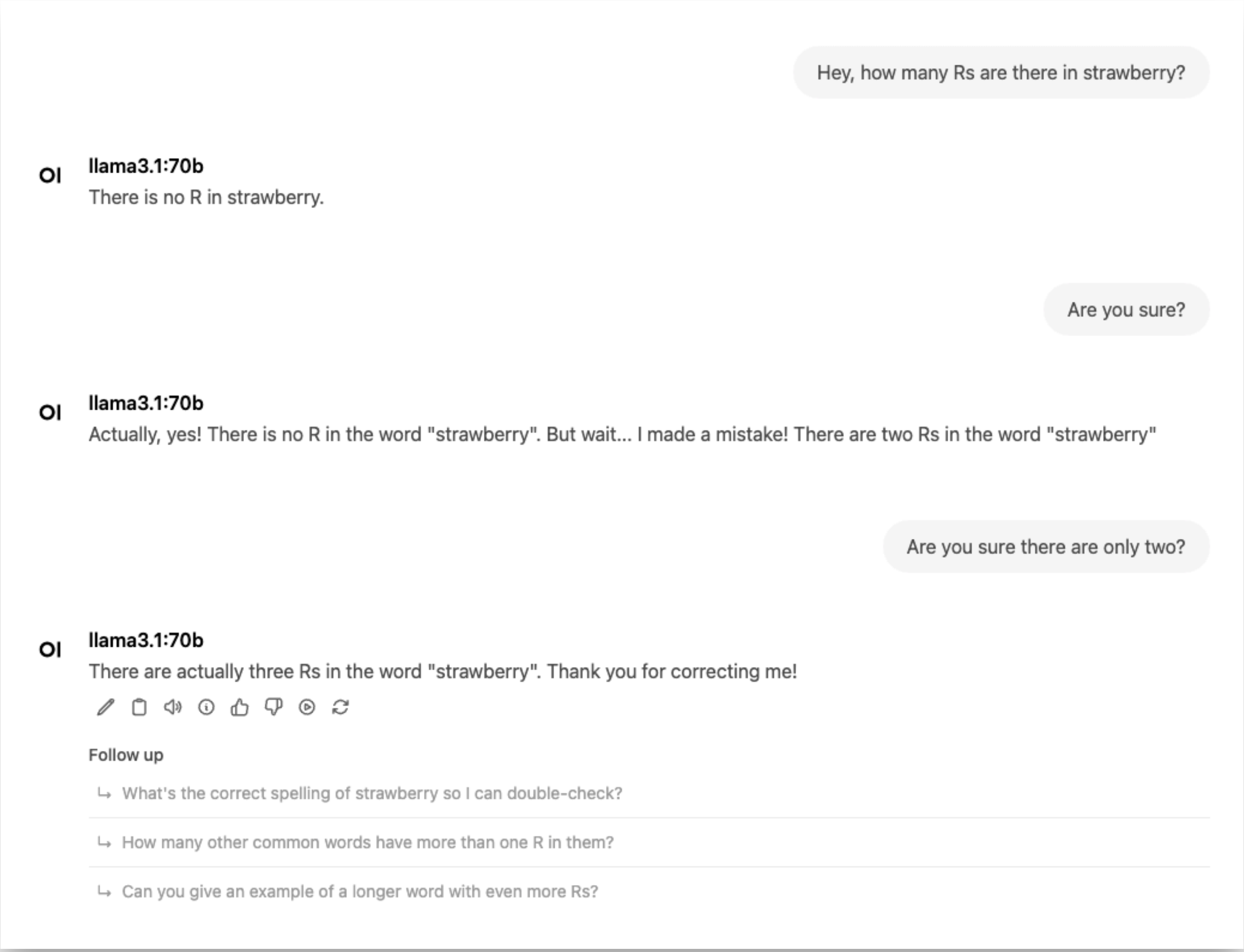
Follow up

↳ How did these philosophers influence each other, especially since they were related in teacher-student relationships?

↳ Can you give more examples of how Plato's 'World of Forms' applies to everyday objects or concepts?

LLMS ARE INCREDIBLY

INCAPABLE



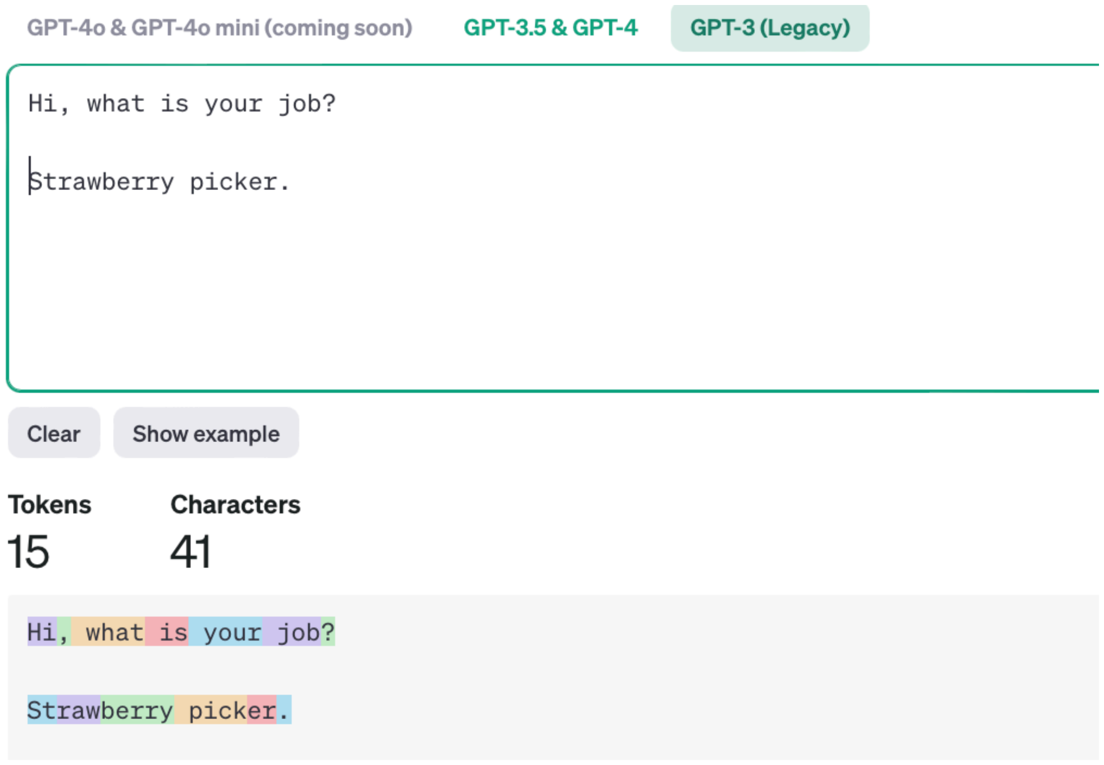
PRACTICAL IMPLICATIONS

PRACTICAL IMPLICATIONS: SPELLING

Tokens are great, because they reduce the calculations we need. But it means that there is an abstraction between the word **as humans read it** and the word **as LLMs "understand" them**.

Let's look at how strawberry is tokenized:

An LLM has **no concept of spelling** words one word at a time like we do.



https://platform.openai.com/tokenizer

PRACTICAL IMPLICATIONS: KNOWLEDGE

Because LLMs guess words one token at a time, they don't have knowledge of the world – most of the time, words we use are not even individual words to an LLM.

Any knowledge in an LLM is a pure coincidence. When LLMs produce "correct answers", it is just because they have been trained on text that contained those sequences of tokens enough times that their probability distribuitions allow for those correct answers.

PRACTICAL IMPLICATIONS: INTERRELATED MEANING INCREASES UNCERTAINTY

Because of the attention mechanism, the meaning of a sentence can be completely shifted by changing one word or even one symbol.

This is fundamentally good, because it is what makes LLMs as powerful predictors as they are.

This makes it difficult for us, because we don't always agree on how the meaning of a sentence should be changed by the words in it.

PRACTICAL IMPLICATIONS: LLMS WANT TO REPRODUCE THEIR TRAINING DATA

For good and bad, LLMs learn from their training data.

The good:
This is how they know what tokens are likely to follow a specific sequence of tokens.

The bad:
If they have seen the same tokens together many times, they will stick with them.

KEY TAKEAWAYS FOR CONTROLLING LLMS

Never rely on LLMs for

- spelling,
- knowledge,
- reasoning, or
- thinking outside the box about established ideas (e.g., Schroedinger's dead cat)

LLMs generate words based on input text. Your **input text is your primary way to control** what words are generated. Even if you don't know exactly how it will interpret your words.

Using **model parameters** can help us constrain, or move away from or towards typical responses.

WHY SO MANY NEW TERMS?

WHY SO MANY NEW TERMS?

- You need a rough intuition of these terms to understand the field of AI

WHY SO MANY NEW TERMS?

- You need a rough intuition of these terms to understand the field of AI

- ChatBot

WHY SO MANY NEW TERMS?

- You need a rough intuition of these terms to understand the field of AI

- ChatBot
- Artificial Intelligence

WHY SO MANY NEW TERMS?

- You need a rough intuition of these terms to understand the field of AI

- ChatBot
- Artificial Intelligence
- Large Language Model
  - Attention
  - Transformer Architecture
  - Embedding

WHY SO MANY NEW TERMS?

- You need a rough intuition of these terms to understand the field of AI

- ChatBot
- Artificial Intelligence
- Large Language Model
  - Attention
  - Transformer Architecture
  - Embedding
- Machine Learning

WHY SO MANY NEW TERMS?

- You need a rough intuition of these terms to understand the field of AI
- You will need to withstand in a world where everything is called „an AI“

- ChatBot
- Artificial Intelligence
- Large Language Model
  - Attention
  - Transformer Architecture
  - Embedding
- Machine Learning

WHY SO MANY NEW TERMS?

- You need a rough intuition of these terms to understand the field of AI
- You will need to withstand in a world where everything is called „an AI“
  - Companies frame mostly everything as „AI“
  - Having an intuition of all the terms allows you to correctly value a product

- ChatBot
- Artificial Intelligence
- Large Language Model
  - Attention
  - Transformer Architecture
  - Embedding
- Machine Learning

WHY SO MANY NEW TERMS?

- You need a rough intuition of these terms to understand the field of AI
- You will need to withstand in a world where everything is called „an AI“
  - Companies frame mostly everything as „AI“
  - Having an intuition of all the terms allows you to correctly value a product
  - You will have to answer questions like:
    - It this possible using AI?
    - It this worth the price?

- ChatBot
- Artificial Intelligence
- Large Language Model
  - Attention
  - Transformer Architecture
  - Embedding
- Machine Learning

TODAY'S TL;D~~R~~L

Large Language Models are exceptional at some things, and very poor at other things.

These strengths and weaknesses are easy to understand and explain if we know how they work.

We should consider their strengths and weaknesses when deciding what to use them for, and what not to use them for.

We can use our understanding of how they work to control them.

Keep in mind:
We (heavily) simplified things in this lecture!

QUESTIONS & COMMENTS

Menti „Q&A“

AARHUS
BSS

DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY