UNIVERSITÄT ZU LÜBECK

GenAI in Education, Science, and Society

Aktuelle Themen Data Science und
KI / CS5070-KP04

Prof. Dr. Sylvia Melzer

SoSe2026

Organisation

GenAI in Education, Science, and Society

Organization

GenAI in Education, Science, and Society

- Type: lecture + seminar (in English)
- Regular attendance at the seminars is required
- Examination information:
  - Option 1: Oral exam covering the entire course material (30 minutes): 100%
  - Option 2:
    - presentation of a publication (30 minutes) + homework (12-15 pages): 20%+40%
    - Topics and dates for the presentations will be coordinated during the seminar
    - Oral exam covering the topic presented during the seminar (15 minutes): 40%

Organization

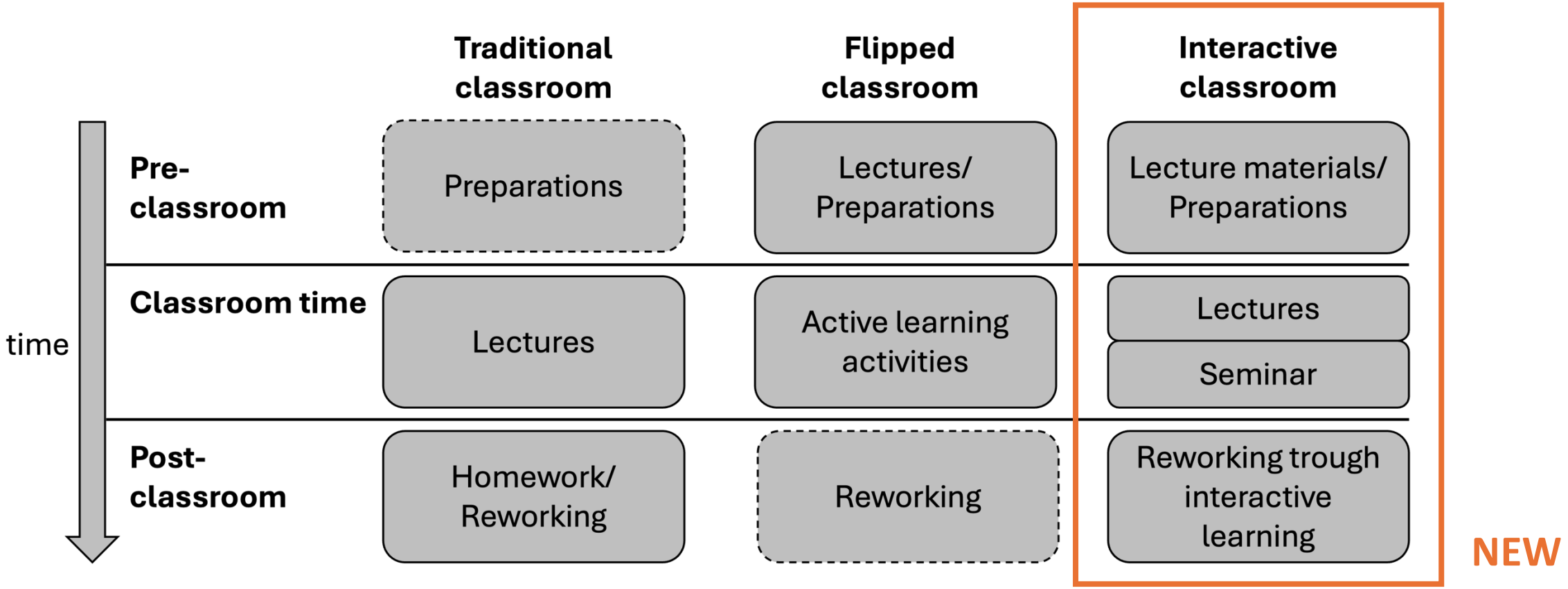
- Teaching material
  - https://moodle.uni-luebeck.de/course/view.php?id=12185

- Teaching material from previous years
  - https://www.edit.fis.uni-hamburg.de/ws/files/67344386/GenAI-WiSe_2025-2026.zip
  - https://www.edit.fis.uni-hamburg.de/ws/files/64545846/GenAI_SoSe2025.zip
  - https://www.edit.fis.uni-hamburg.de/ws/files/56466394/GenAI_2024.zip

Acknowledgement

These lecture notes were prepared in collaboration with Prof. Dr. habil. Ralf Möller.

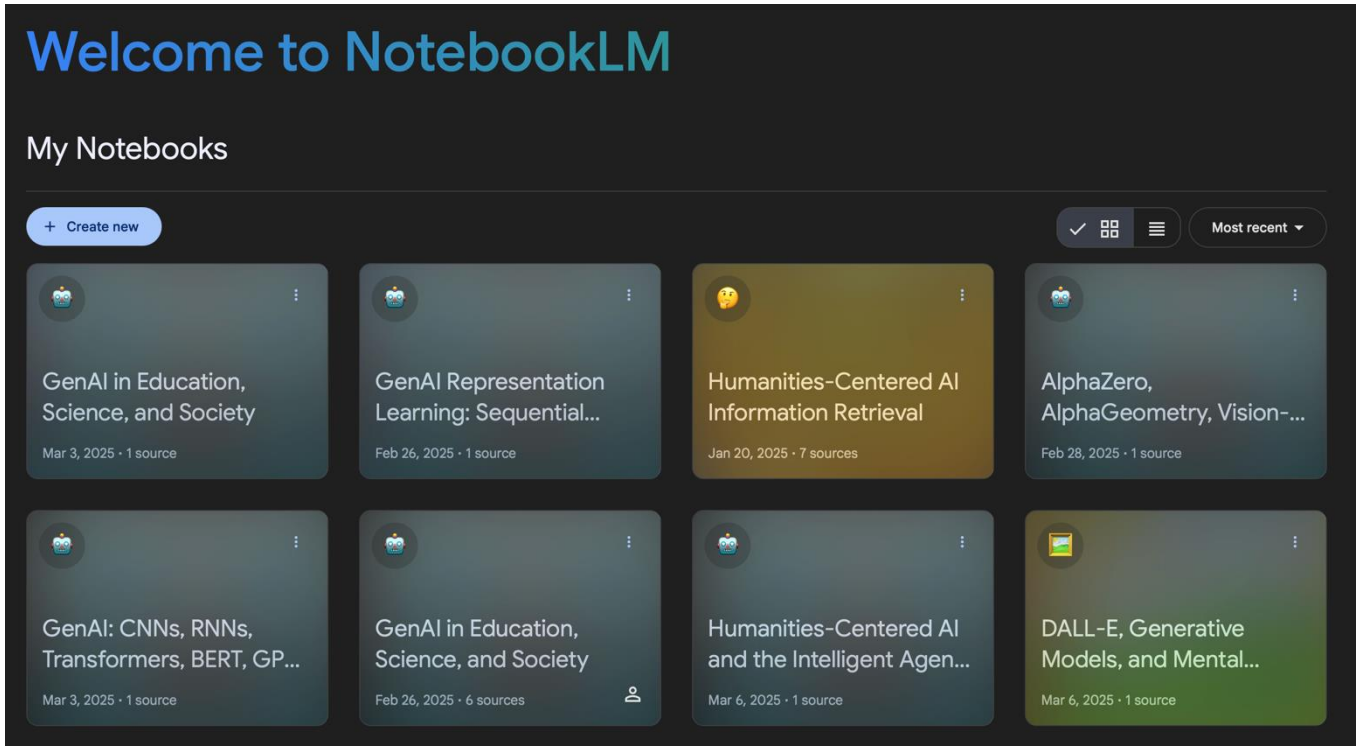
Teaching concept



Melzer, S., Schiff, S., Weise, F. *et al.* Ontology-based federated information systems with NLP-enhanced retrieval. *Int J Digit Humanities* (2025). https://doi.org/10.1007/s42803-025-00110-y

NotebookLM

- an experimental offering from Google Labs to create interactive Podcasts
- https://notebooklm.google.com



GenAI in Education, Science, and Society

Sources

+ Add source

Select all sources

01_GenAI_Organization+Intro.pdf

1) Uploading the lecture files

Chat

2) Automatically Generated Summarization

GenAI in Education, Science, and Society

1 source

The provided material outlines a lecture and seminar course titled "GenAI in Education, Science, and Society," detailing its **organization, learning objectives, and a comprehensive schedule** covering topics from foundational AI concepts to specific GenAI models and applications. It introduces **key AI concepts** like agents, perception, and action planning, further exploring the **challenges of AI control** with the "off-switch problem." The text then transitions to **Generative AI**, defining its capabilities and highlighting its current and potential impact across education, science, and society, concluding with suggested **seminar topics** for deeper exploration of these areas.

Save to note

Add note | Audio Overview | Mind Map

3) Chat with NotebookLM about your lecture content

Start typing... 1 source

How does GenAI's functionality differentiate it from traditional AI within education, science, a

Studio

4) Create your own podcast

Audio Overview

GenAI in Education, Science, and Society

00:01 / 13:33

Interactive mode BETA

Notes

+ Add note

Study guide | Briefing doc

FAQ | Timeline

5) Be part of the podcast – be interactive!

Join

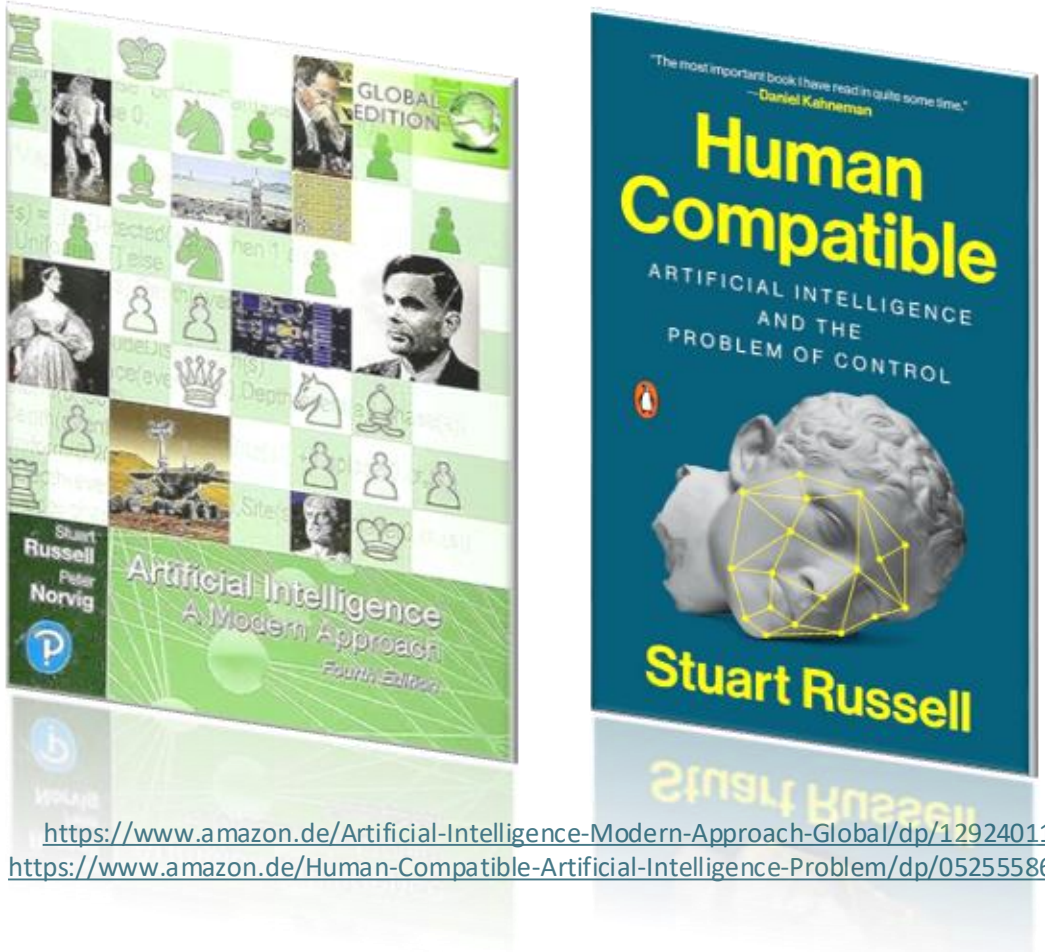
Learning objectives

- By completing the course "GenAI in Education, Science, and Society", you will be able to
  - name the central ideas of GenAI in Education, Science, and Society
  - define the relevant terms, and
  - explain how the associated methods and approaches work using application examples.
- To get new job opportunities

	Date	Topics
1	09. Apr. 2026	▪ Organization and Introduction
2	16. Apr. 2026	▪ Representation Learning for Sequential Structures, Embedding Spaces, word2vec, CBOW, Skip-gram
3	23. Apr. 2026	▪ CNNs, RNNs, LSTMs, ELMO, Transformers, BERT, GPT Models
4	30. Apr. 2026	▪ Training Embedding Architectures
5	07. May 2026	▪ Large Language Models and Prompting
6	21. May 2026	▪ AlphaGo, AlphaZero, AlphaGeometry, AlphaFold
7	28. May 2026	▪ Generative Vision Models, Mental Imagery; Examples for generating images, music, video
8	04. Jun. 2026	▪ Intelligent Agents
9	11. Jun. 2026	
10	18. Jun. 2026	
11	25. Jun. 2026	
12	02. Jul. 2026	
13	09. Jul. 2026	
14	16. Jul. 2026	

Literature

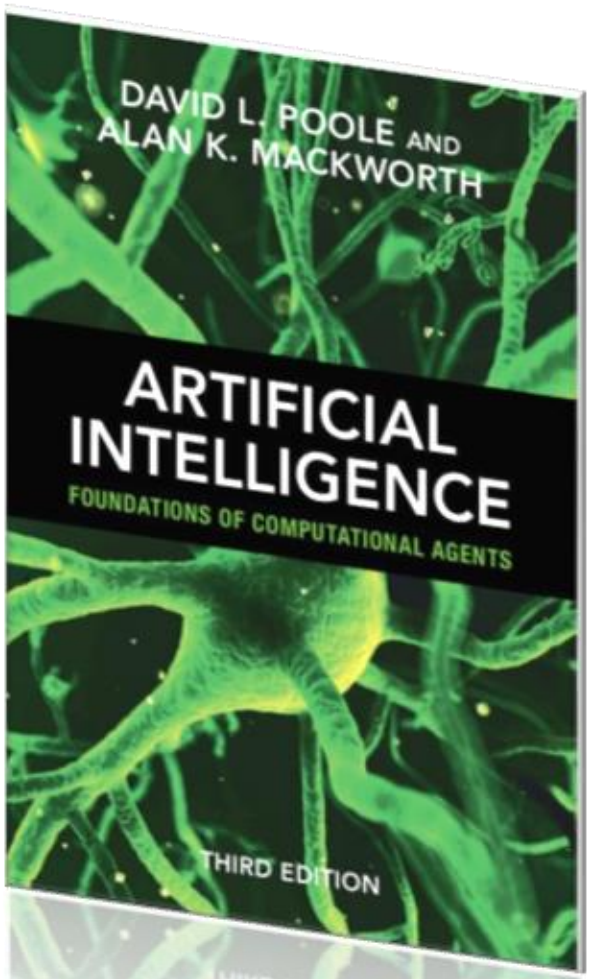
- Stuart Russell and Peter Norvig. 2020. Artificial Intelligence: A Modern Approach (4th. ed.). Prentice Hall Press, USA.
- Stuart Russell. 2020. Human Compatible: Artificial Intelligence and the Problem of Control. Penguin Publishing Group; Reprint Edition.

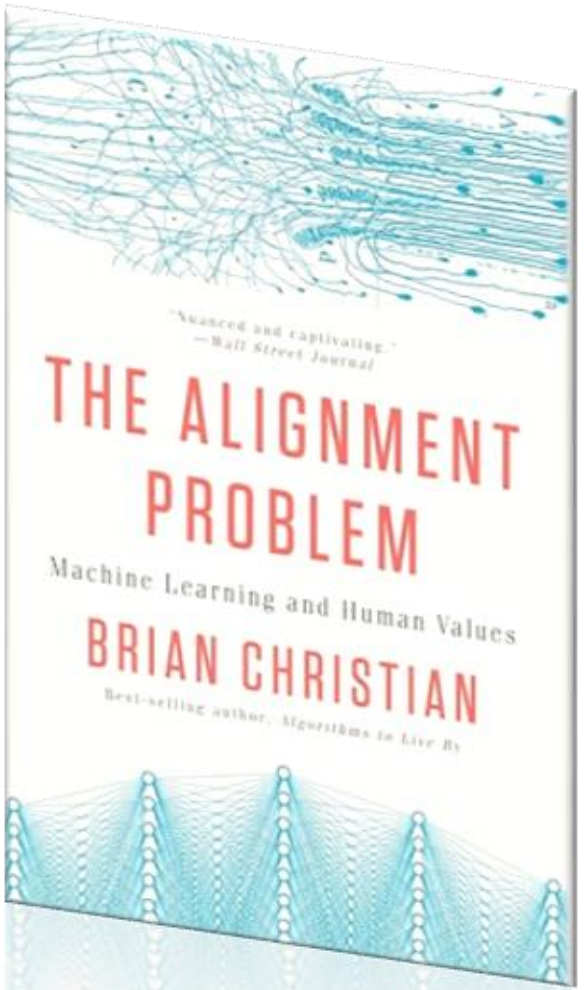


https://www.amazon.de/Artificial-Intelligence-Modern-Approach-Global/dp/1292401133/
https://www.amazon.de/Human-Compatible-Artificial-Intelligence-Problem/dp/0525558632/

Literature

- David L. Poole and Alan K. Mackworth, 2023. Artificial Intelligence 3E: Foundations of Computational Agents (3rd ed.). Cambridge University Press
- Brian Christian, 2020. The Alignment Problem: Machine Learning and Human Values. WW Norton





https://artint.info/3e/html/ArtInt3e.html

Introduction

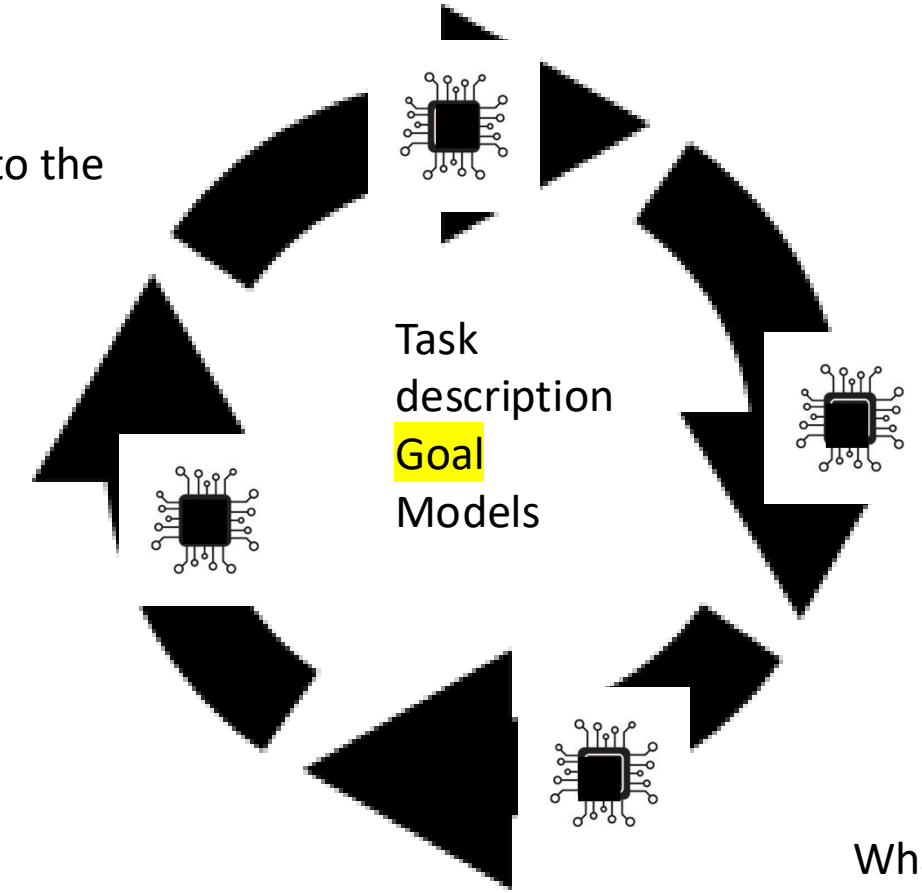
GenAI in Education, Science, and Society

What terms do you associate with AI?

- https://www.mentimeter.com/app/presentation/al1m7uws8jxh44qyznhxi42g5dza6xeb/edit?source=share-modal

AI: Science about agents acting in the world (Lat. agere: act)

Percepts → From percepts to the state of the environment



← Action incl. explanation

Task description → Goal
Is my task description still correct given the current state of the environment? Should I adopt a new task description?

Plan or just react?

Action planning strategy
What is the best action in the current state for a given goal?

Perseverence: Mars Rover

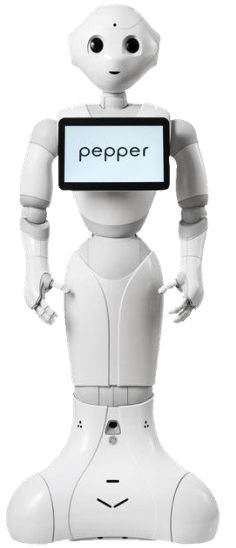
Mars Rover: Autonomy is Necessary

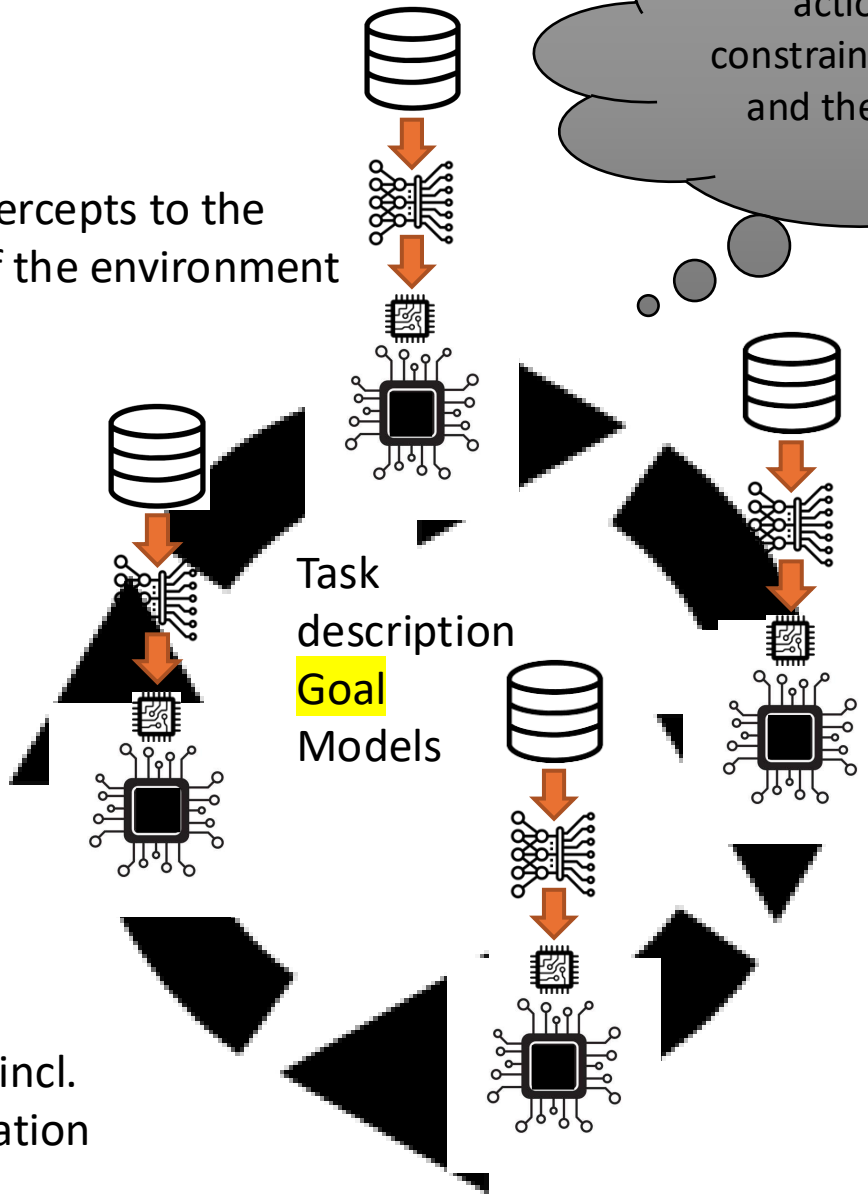
- Systems must act! (Reinforcement learning necessary)
  - Development time
  - Deployment time
- Data interpretation → Model assumptions required
- Model assumptions possibly need to be revised
- Decision: Use policy/employ planning
- Task description → “goal”
- Goal still valid?

Intelligence concept: locally optimal action under (severe) resource constraints in relation to uncertain goals and the preferences of others in the environment

Percepts

From percepts to the state of the environment





Task description
Goal
Models

Action incl. explanation

Task description → Goal

Is the goal from the task description still correctly selected given the current state of the environment? If necessary, define a (new) internal goal

Optimize local success measure (utility)

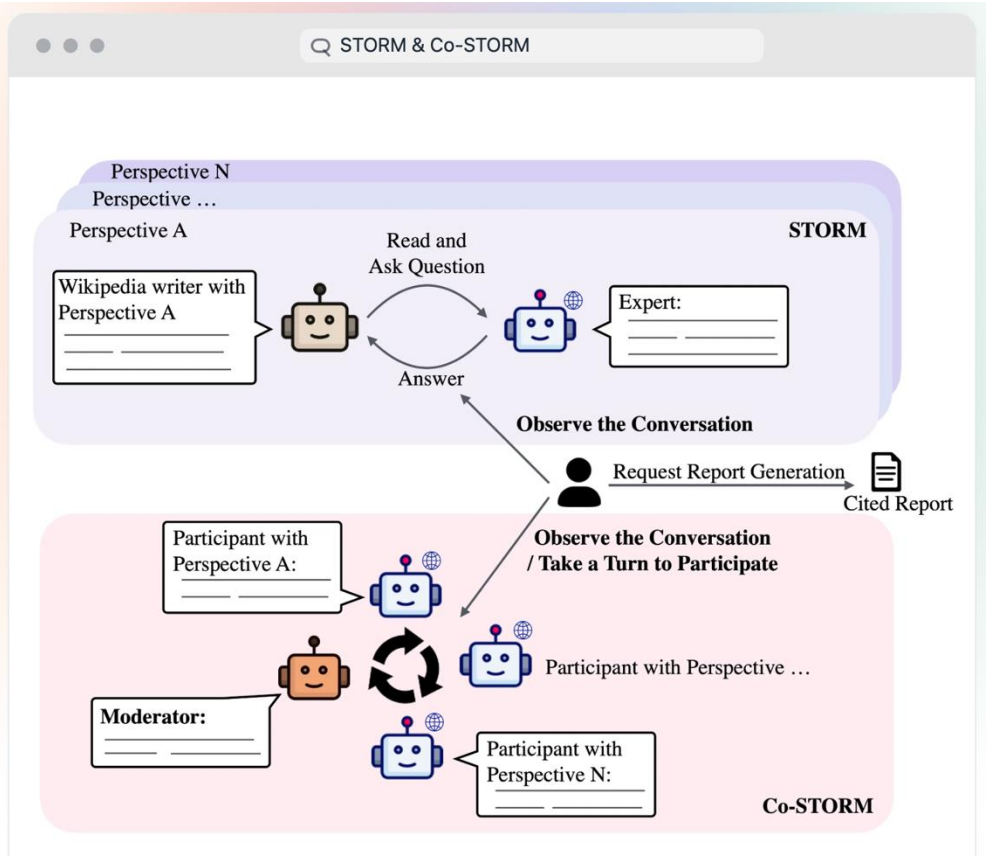
Action planning strategy

What is the best action in the current state for a given goal?

Co-Storm: An Example of Collaborative Humans & Agents

Co-Storm

- combines **human** expertise with **agents** to generate knowledge
- emphasises the science of AI not only as an autonomous agent, but also as an approach that can complement human cognitive abilities in the generation of knowledge

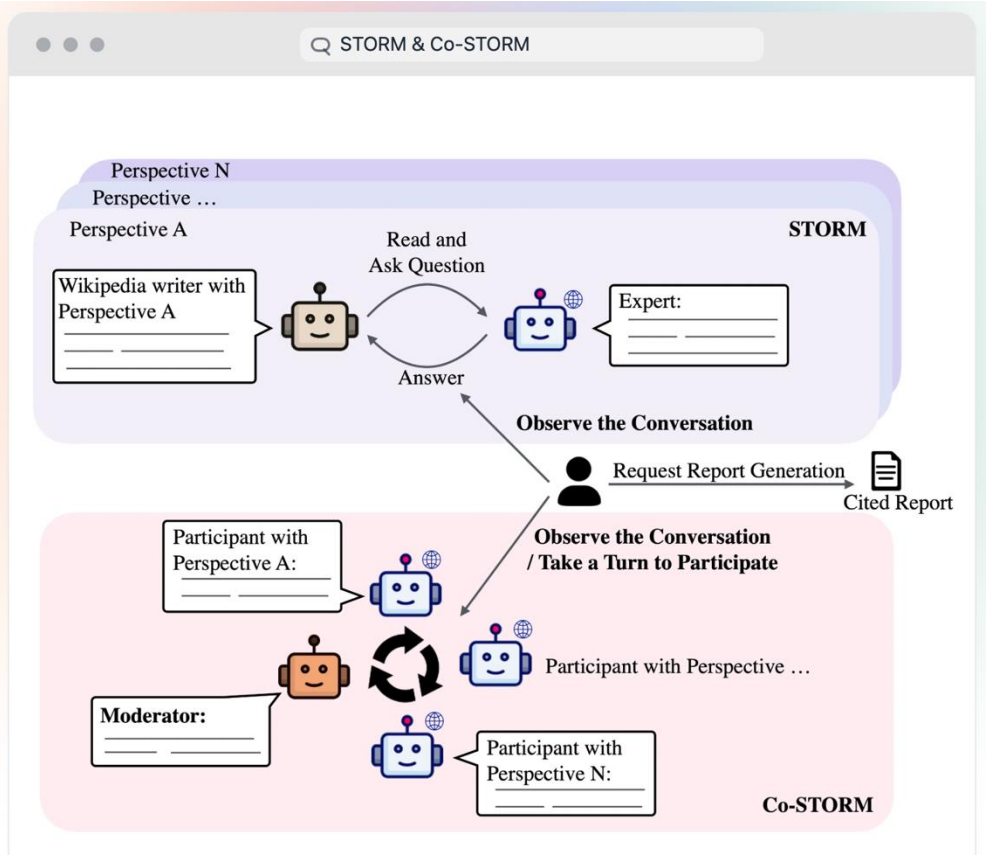


https://storm.genie.stanford.edu

Co-Storm

The Moderator as Principal:

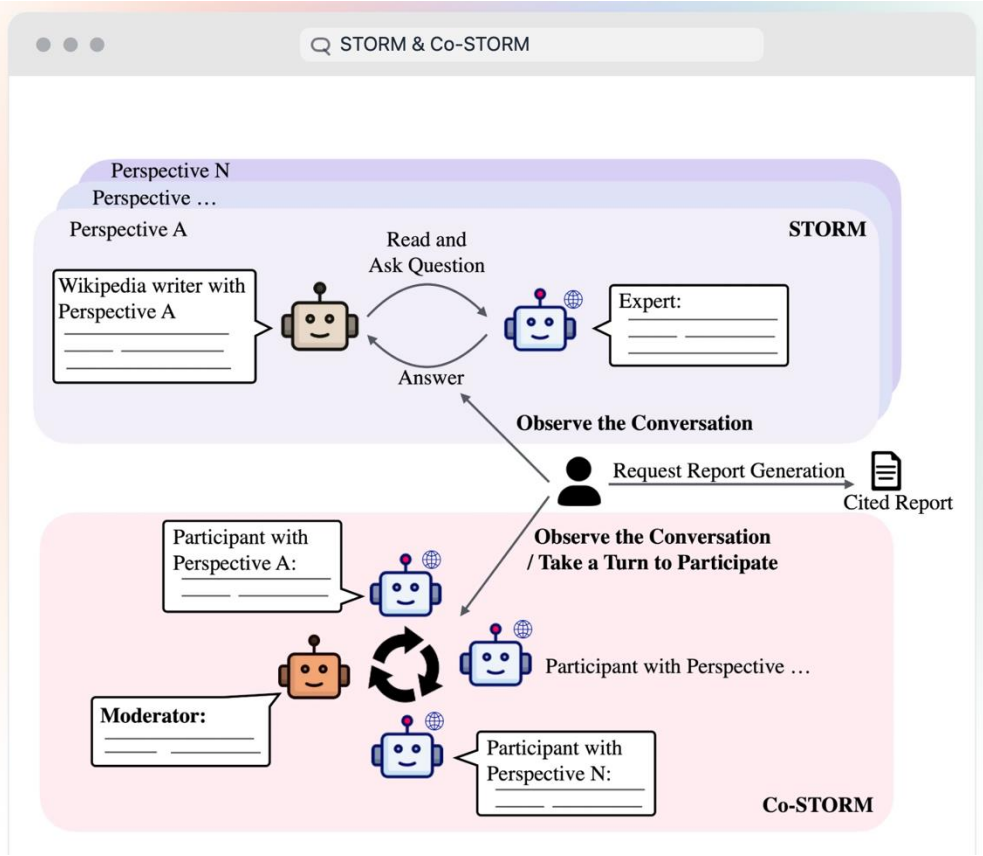
- Acts as the central authority responsible for decision-making
- Has the goal to aggregate expert contributions into a well-informed collective decision
- Designs the decision-making process and ensures that the experts contribute effectively



Co-Storm

The experts (agents)

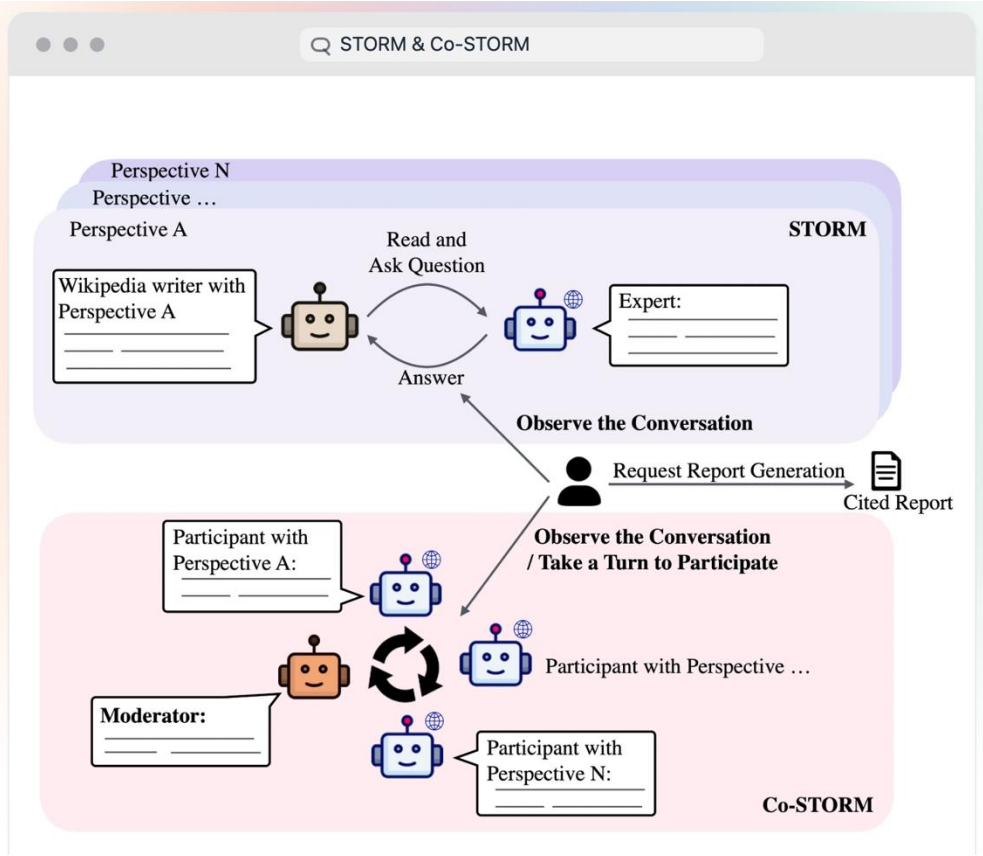
- Provide specialized knowledge or insights
- Provide information to the moderator
- Their interests may not always align with the mod
  must be carefully designed to ensure cooperation



Co-Storm

Reward

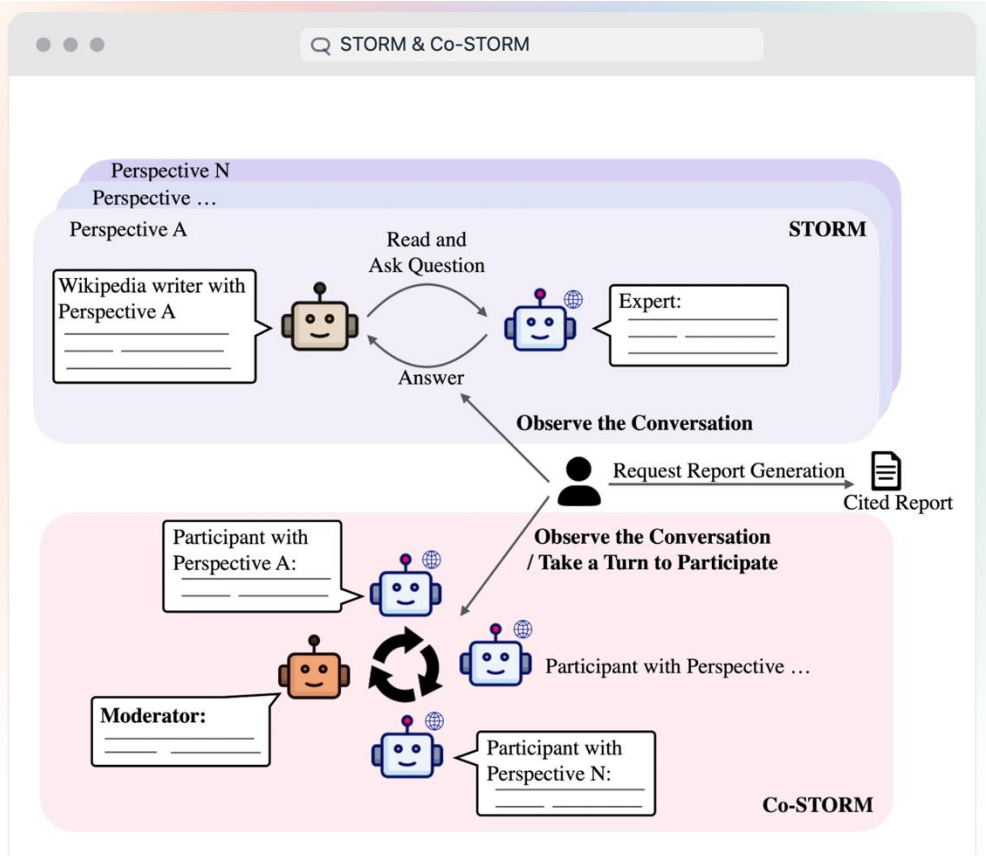
- Instead of rewarding the moderator, incentives are distributed to the experts
- These rewards are usually performance-based and depend on the relevance and usefulness of the information they provide



Co-Storm

Best Choice

- The system aggregates expert inputs into a collective outcome that reflects the best possible decision (optimized local success measure) given the available information (→ principal agent problem)

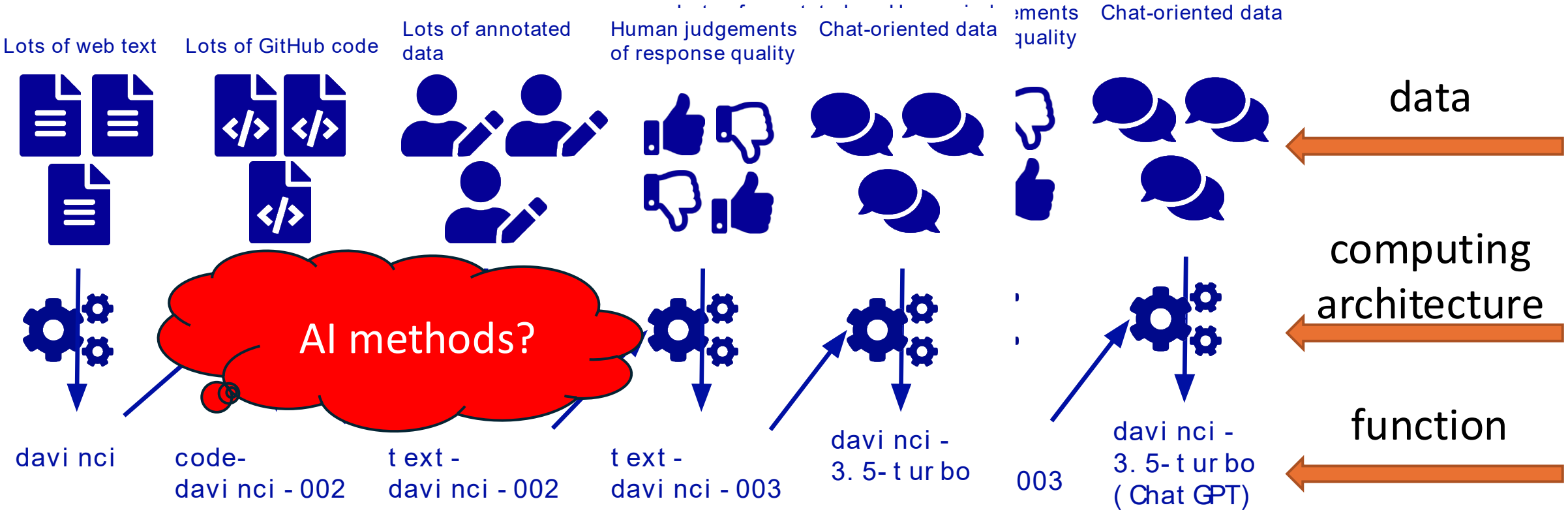


https://en.wikipedia.org/wiki/Principal–agent_problem#:~:text=The%20principal–agent%20problem%20refers,is%20called%20%22agency%20costs%22

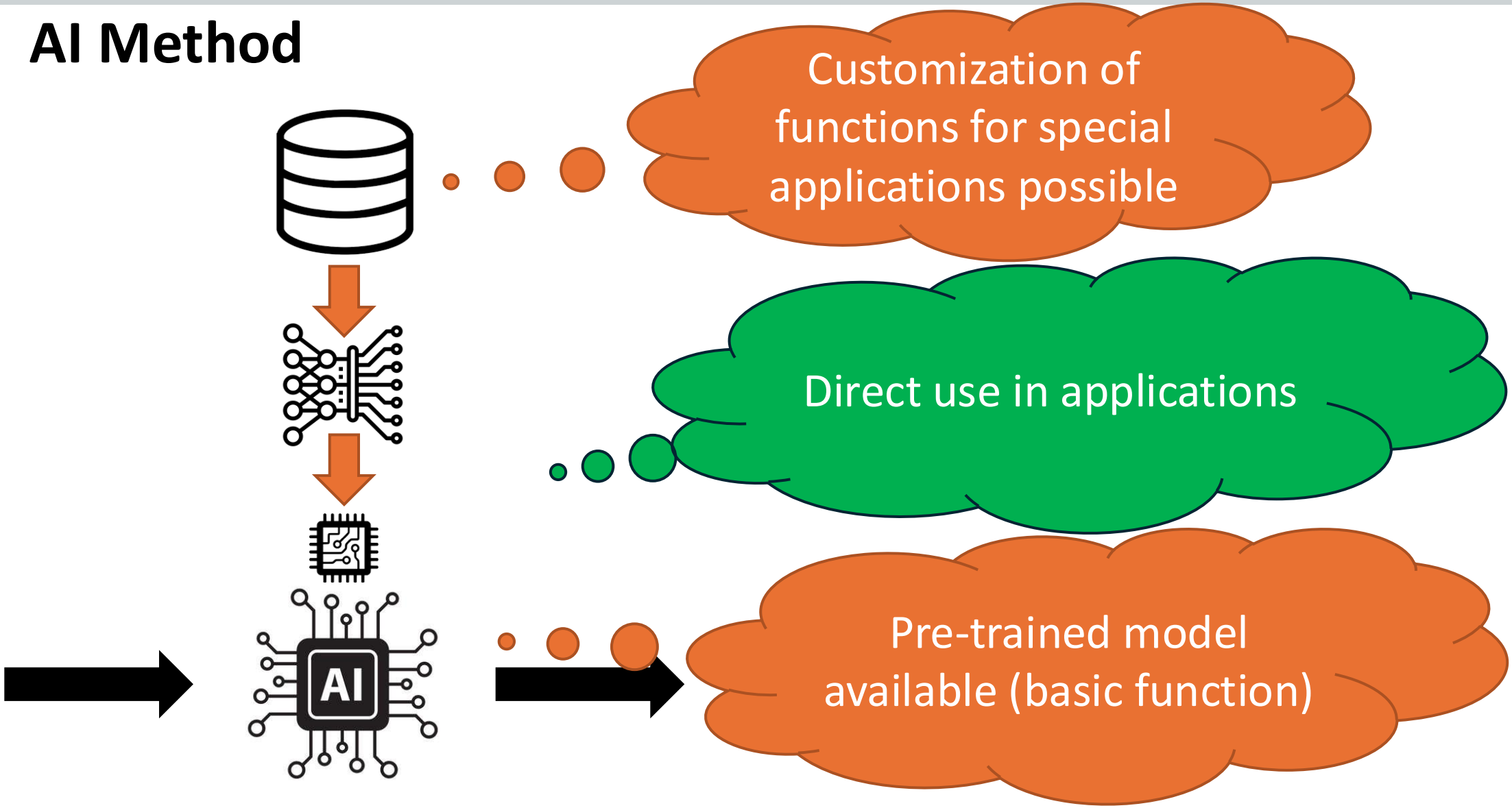
Data, Architecture & Functions

Image: Graham Neubig, Large Language Models: A Bird's Eye View, https://phontron.com/slides/neubig23llms.pdf

From Zero to ChatGPT



AI Method


Customization of functions for special applications possible
Direct use in applications
Pre-trained model available (basic function)
AI

Functions vs. Agents

Function

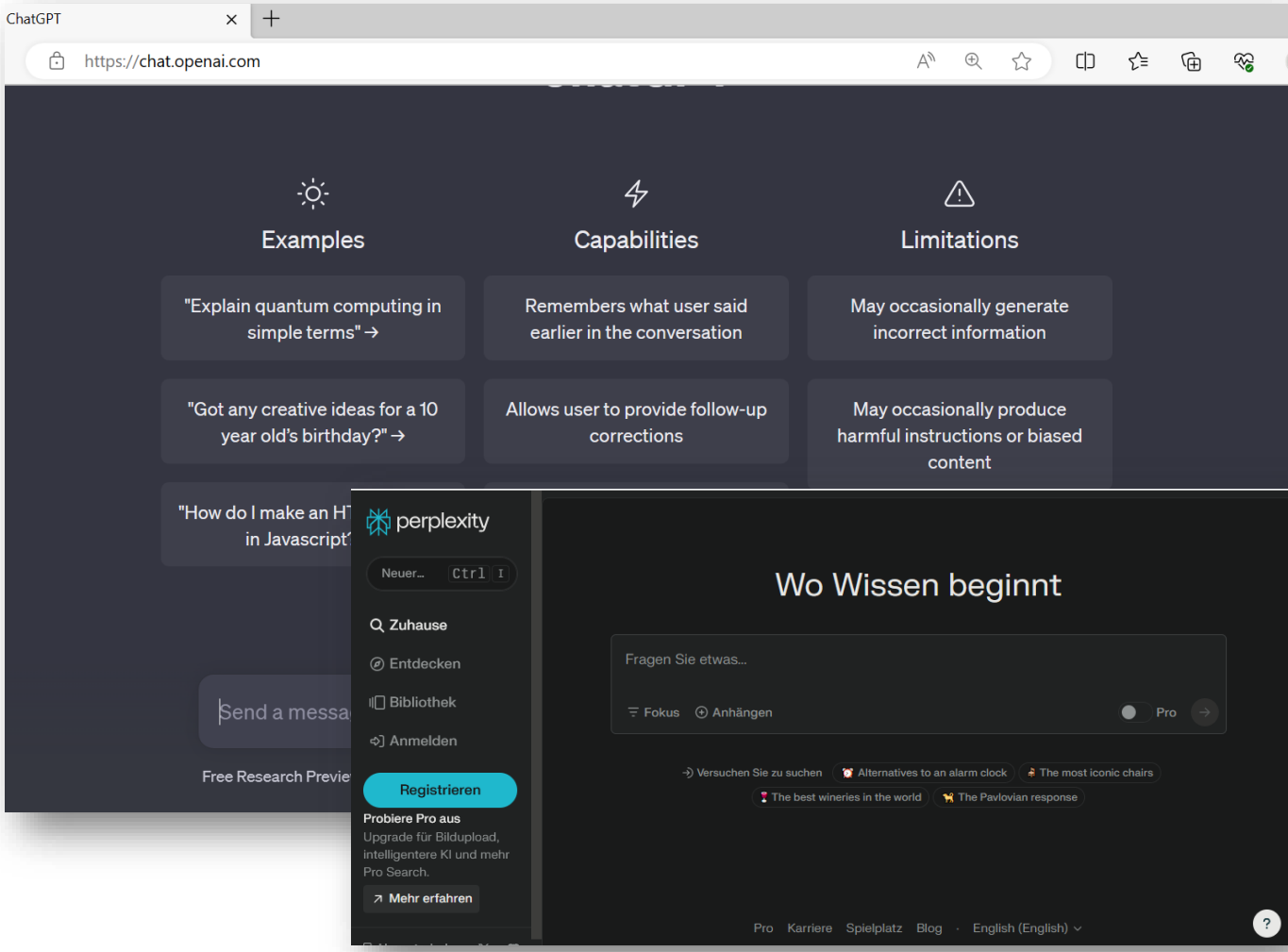
- Generative Pre-trained Transformer (GPT)

Agents

- ChatGPT
- Perplexity.ai
- …

Usage

- code generation
- text generation
- …



Agents in Action

- Task description: Please summarize “Alice in Wonderland”

ChatGPT 3.5

You

Please summarize “Alice in Wonderland”

task description

ChatGPT

"Alice's Adventures in Wonderland," commonly known as "Alice in Wonderland," is a novel written by Lewis Carroll (pseudonym for Charles Lutwidge Dodgson), published in 1865. The story follows a young girl named Alice who falls down a rabbit hole into a whimsical, fantastical world filled with peculiar characters and surreal events.

In Wonderland, Alice encounters anthropomorphic animals, such as the White Rabbit, the Cheshire Cat, and the March Hare, as well as eccentric figures like the Mad Hatter and the Queen of Hearts. She experiences a series of bizarre adventures, including changes in size, nonsensical conversations, and encounters with strange creatures.

Explanation

The narrative is marked by its dreamlike quality, absurdity, and playful use of language and logic. As Alice navigates through Wonderland, she struggles to make sense of its nonsensical rules and illogical situations. The story explores themes of identity, growing up, and the nature of reality versus imagination.

Overall, "Alice in Wonderland" is a classic work of children's literature celebrated for its imaginative storytelling, whimsical characters, and clever wordplay.

Please summarize “Alice in Wonderland”

Quellen

Alice in Wonderland Plot Summary | Course Hero
coursehero · 1

Alice in Wonderland Summary - JavaTpoint
javatpoint · 2

Alice in Wonderland by Lewis Carroll | Brief Plot...
youtube · 3

Sehen Sie 2 mehr

Bilder suchen +

Videos suchen +

Bild generieren +

Antwort

Summary of Alice in Wonderland

Alice in Wonderland is a classic fantasy novel by Lewis Carroll that follows the adventures of a young girl named Alice. The story begins when Alice is sitting on a riverbank with her sister, daydreaming. Suddenly, she sees a White Rabbit wearing a waistcoat and carrying a pocket watch, who says he is late for a very important date. Alice curiously follows the Rabbit down a rabbit hole, which leads her to the fantastical world of Wonderland.

In Wonderland, Alice encounters a series of strange and nonsensical characters and situations. She struggles to control her constantly changing size after eating and drinking various substances. Alice meets the Cheshire Cat, who can disappear and reappear at will, the Mad Hatter and March Hare having a never-ending tea party, and the tyrannical Queen of Hearts, who is quick to order the beheading of anyone who upsets her. [1][2][3][4]

Throughout her adventures, Alice tries to make sense of the absurd rules and behaviors of Wonderland, which often mock the rigid social norms of Victorian England. Despite the chaos, Alice remains determined and level-headed, asserting her own sense of justice and morality. [2][3][4]

Ultimately, Alice wakes up to find that her adventures in Wonderland were just a dream. However, the story's themes of self-reliance, the arbitrariness of rules, and the challenges of growing up continue to resonate with readers. Lewis Carroll's innovative use of wordplay, dreams, and inversions of traditional tales helped revolutionize children's literature. [1][2][3][4]

Perplexity.ai

Agents in Action

- Task description: Who is the president of the USA?

Who is the president of the USA?

Quellen

President of the United States - Wikipedia
en.wikipedia · 1

Presidents | The White House
whitehouse · 2

List of presidents of the United States - Wikipedia
en.wikipedia · 3

Sehen Sie 2 mehr

Antwort

Based on the search results provided, the current president of the United States is Joe Biden.

The key evidence is:

- Joe Biden is listed as the 46th President of the United States on the White House website [4].
- The search results provide a biography of Joe Biden, stating that he "represented Delaware for 36 years in the U.S. Senate before becoming the 47th Vice President of the United States. As President, Biden will restore America's leadership and build our communities back better." [4]
- The list of all U.S. presidents on Wikipedia confirms that Joe Biden is the 46th and current president, having assumed office on January 20, 2021 [3].

Therefore, the president of the USA is Joe Biden.

Teilen Umschreiben

Mehr anzeigen

Videos suchen

Bild generieren

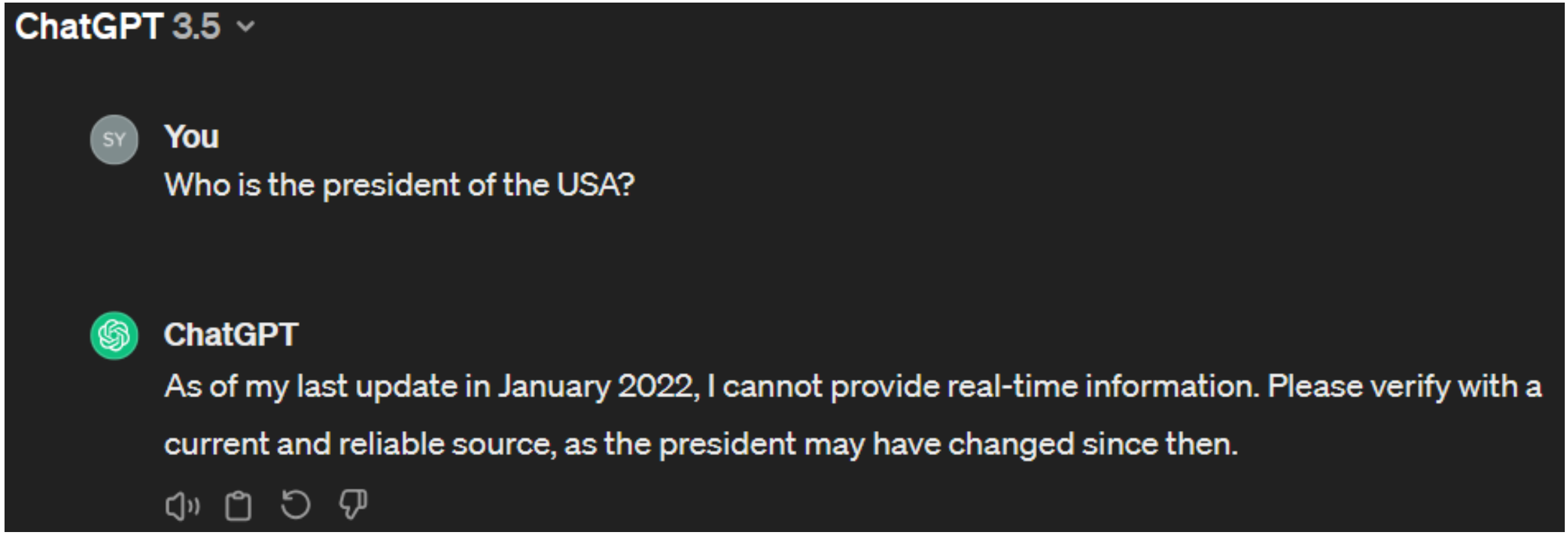
Ähnlich

what are the responsibilities of the president of the united states

how long is the term of the president of the united states

who was the first president of the united states

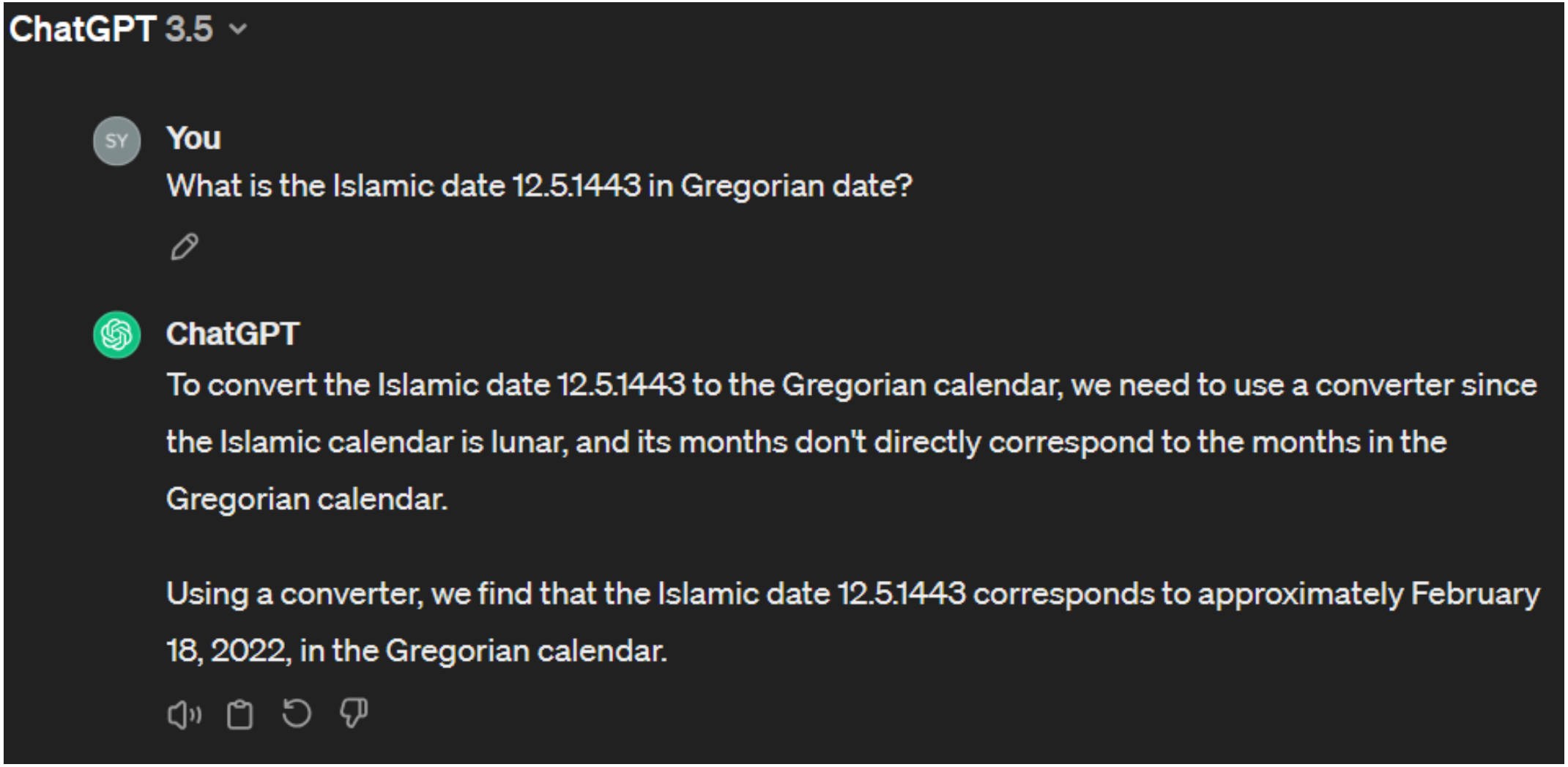
Perplexity.ai


ChatGPT 3.5
SY
You
Who is the president of the USA?
ChatGPT
As of my last update in January 2022, I cannot provide real-time information. Please verify with a current and reliable source, as the president may have changed since then.

Agents in Action

Task description: What is the Islamic date **12.5.1443** in Gregorian date?

An Agent in Action



What is the Islamic date 12.5.1443 in Gregorian date?

Quellen

Calendar Converter - Islamic Philosophy...
muslimphilosophy · 1

Dhu al-Hijjah Hijri calendar - 1443 - التقويم...
calendarhijri · 2

Hijri to Gregorian Date Converter - Islamic...
islamicfinder · 3

Chronological List of Presidents, First Ladies,...
loc · 4

Antwort

According to the search results, the Islamic date 12.5.1443 corresponds to the Gregorian date 2022/09/09.

The key evidence is:

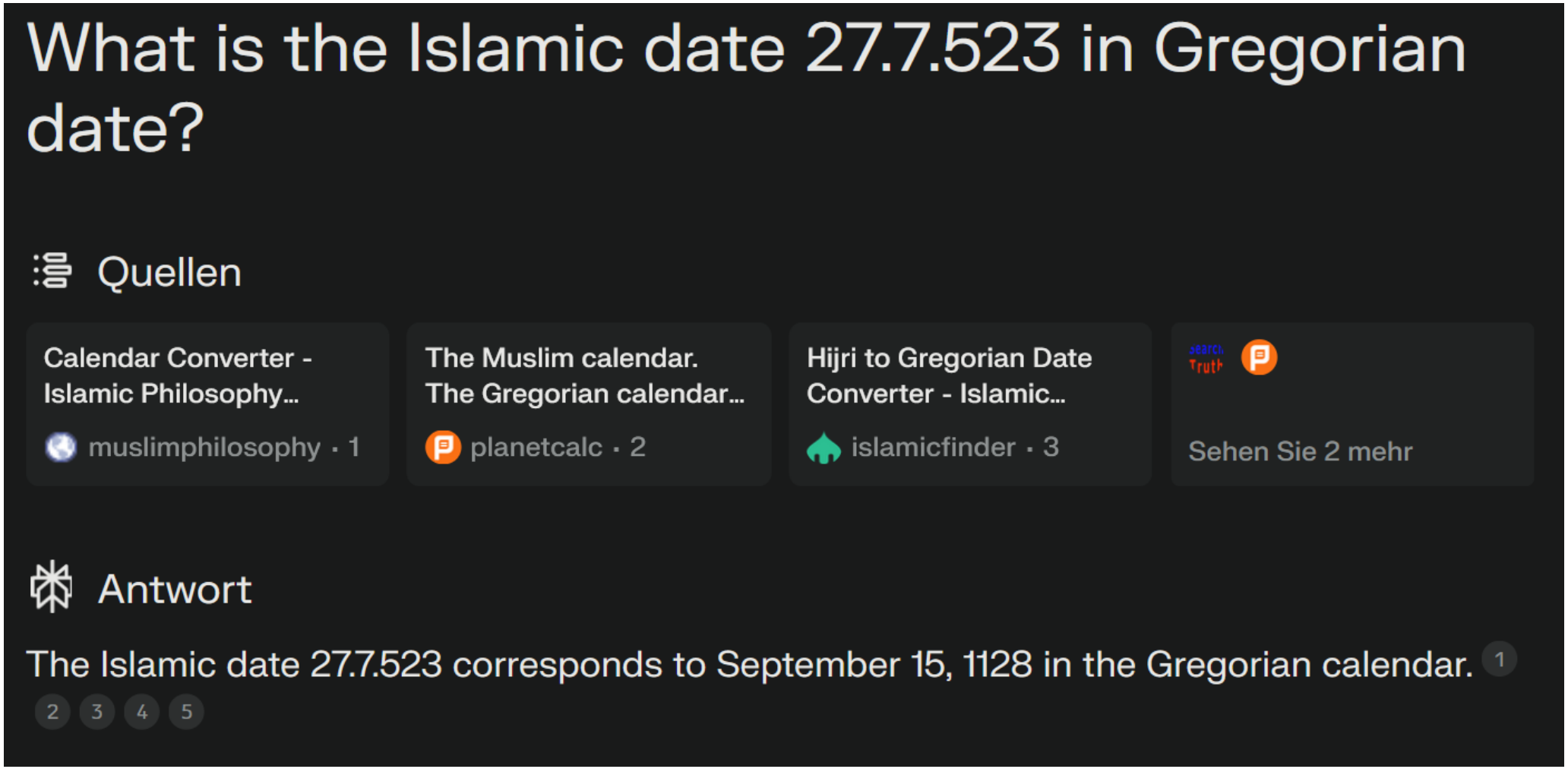
- The Hijri to Gregorian date converter on IslamicFinder.org [3] shows that the Islamic date 12.5.1443 converts to the Gregorian date 2022/09/09.

Therefore, the Gregorian date for the Islamic date 12.5.1443 is 2022/09/09.

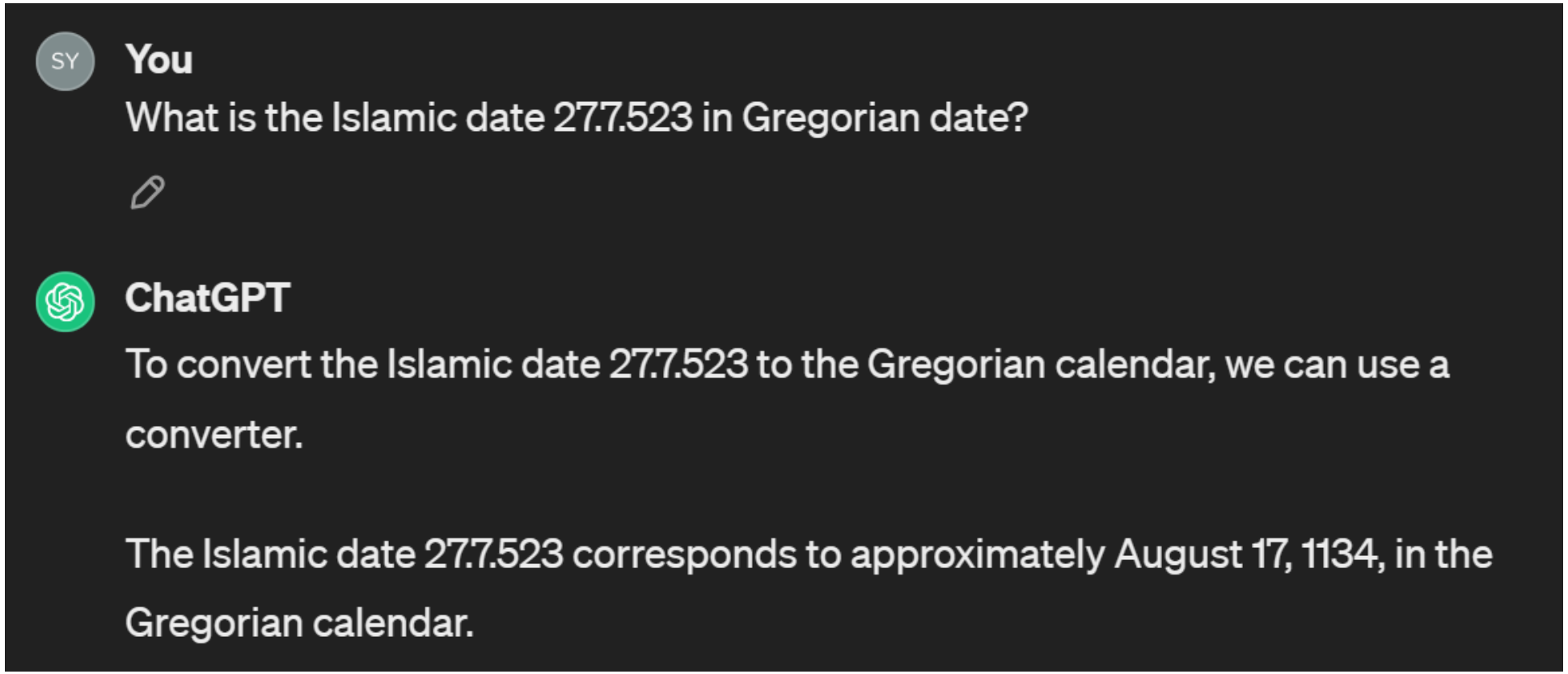
Perplexity.ai

Agents in Action

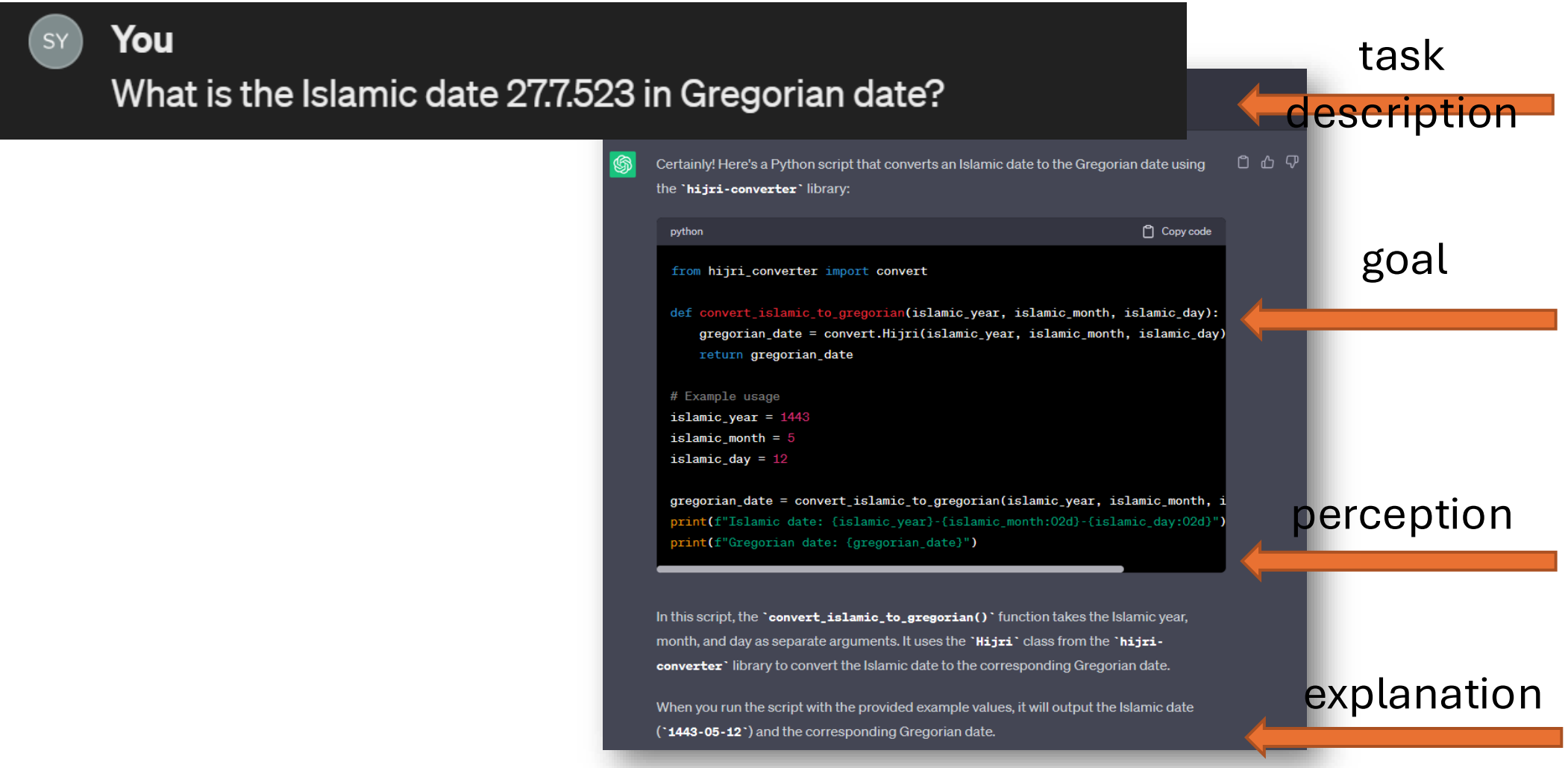
- Task description: What is the Islamic date **27.7.523** in Gregorian date?


What is the Islamic date 27.7.523 in Gregorian date?
Quellen
Calendar Converter - Islamic Philosophy...
muslimphilosophy · 1
The Muslim calendar. The Gregorian calendar...
planetcalc · 2
Hijri to Gregorian Date Converter - Islamic...
islamicfinder · 3
Sehen Sie 2 mehr
Antwort
The Islamic date 27.7.523 corresponds to September 15, 1128 in the Gregorian calendar. 1
2 3 4 5

Agents in Action



- **Where is the goal?**


SY
You
What is the Islamic date 27.7.523 in Gregorian date?
Certainly! Here's a Python script that converts an Islamic date to the Gregorian date using the `hijri-converter` library:
python
Copy code
from hijri_converter import convert
def convert_islamic_to_gregorian(islamic_year, islamic_month, islamic_day):
gregorian_date = convert.Hijri(islamic_year, islamic_month, islamic_day)
return gregorian_date
# Example usage
islamic_year = 1443
islamic_month = 5
islamic_day = 12
gregorian_date = convert_islamic_to_gregorian(islamic_year, islamic_month, i
print(f"Islamic date: {islamic_year}-{islamic_month:02d}-{islamic_day:02d}")
print(f"Gregorian date: {gregorian_date}")
In this script, the `convert_islamic_to_gregorian()` function takes the Islamic year, month, and day as separate arguments. It uses the `Hijri` class from the `hijri-converter` library to convert the Islamic date to the corresponding Gregorian date.
When you run the script with the provided example values, it will output the Islamic date (`1443-05-12`) and the corresponding Gregorian date.
task
description
goal
perception
explanation

Explanations

- Task description: What is the Islamic date **12.5.1443** in Gregorian date?
  - ChatGPT's explanation: **18.2.2022**
  - Perplexity.ai explanation: **9.9.2022**

- Task description: What is the Islamic date **27.7.523** in Gregorian date?
  - ChatGPT's explanation: **17.8.1134**
  - Perplexity.ai explanation: **15.9.1128**

➔ **Check ChatGPT's goal**

- http://www.muslimphilosophy.com/calconv/index.html
- implementation

import hijri_converter.py

C: > Users > Melzer > Documents > 02_Forschung > DataLinkingWorkshop2023 > import hijri_converter.py > ...

```python
from hijri_converter import convert

def convert_islamic_to_gregorian(islamic_year, islamic_month, islamic_day):
    gregorian_date = convert.Hijri(islamic_year, islamic_month, islamic_day).to_gregorian()
    return gregorian_date

# Example usage
islamic_year = 1443
islamic_month = 5
islamic_day = 12

gregorian_date = convert_islamic_to_gregorian(islamic_year, islamic_month, islamic_day)
print(f"Islamic date: {islamic_year}-{islamic_month:02d}-{islamic_day:02d}")
print(f"Gregorian date: {gregorian_date}")
```

Test result

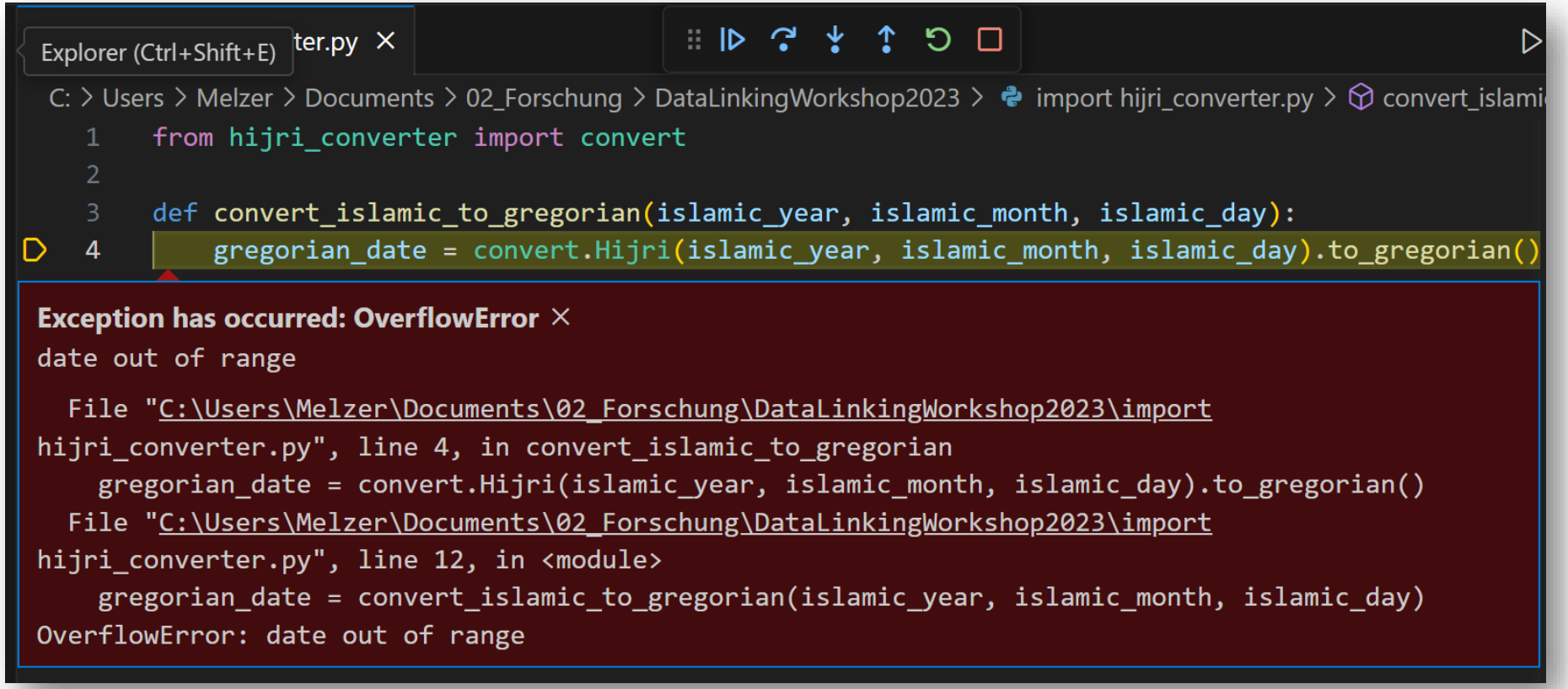
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL Python Debug Console

```
Gregorian date: 2021-12-16
```

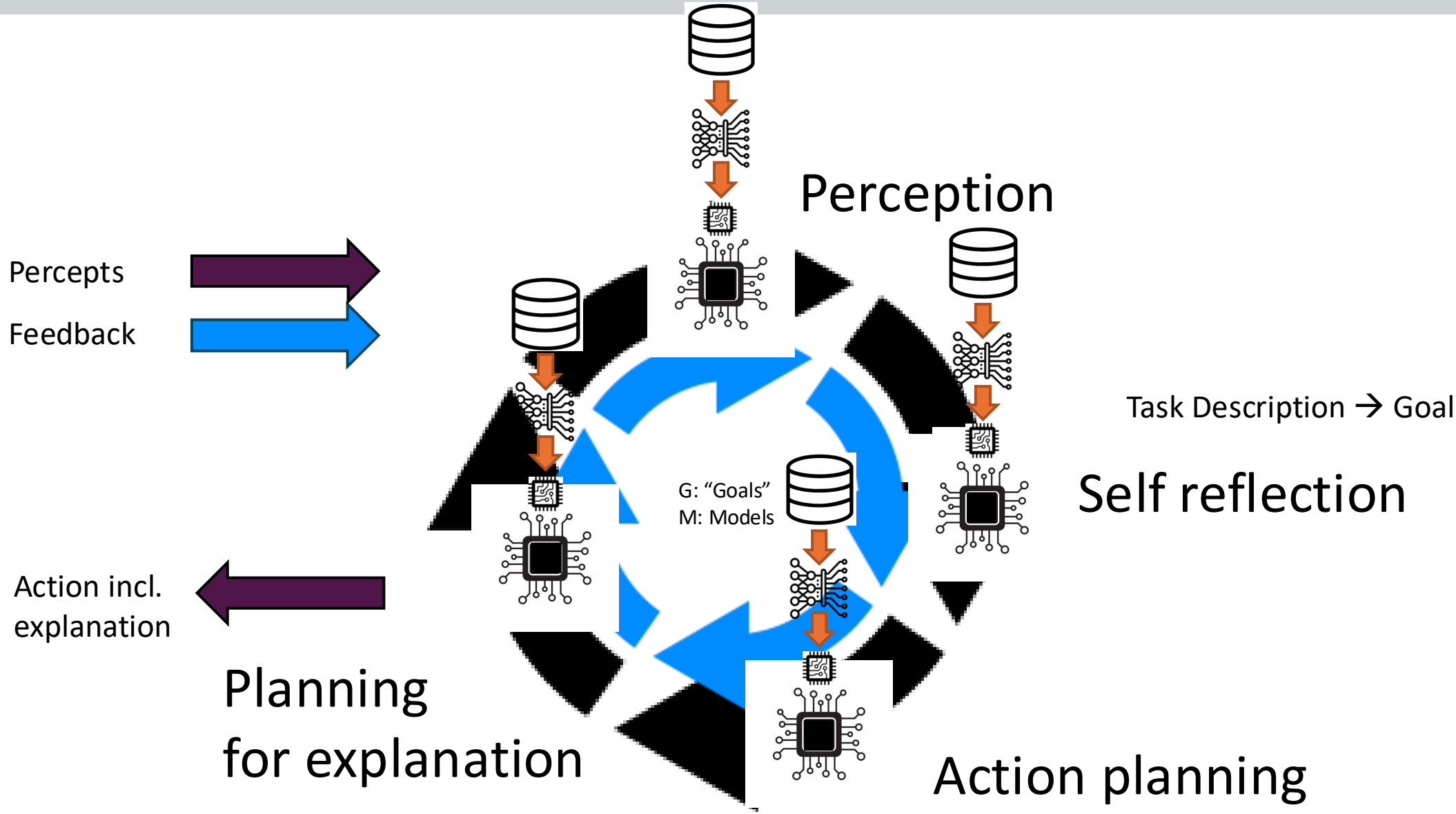
import hijri_converter.py

C: > Users > Melzer > Documents > 02_Forschung > DataLinkingWorkshop2023 > import hijri_converter.py > ...

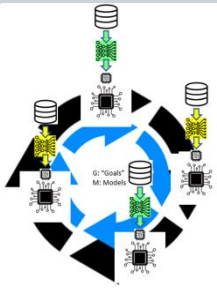
```python
from hijri_converter import convert

def convert_islamic_to_gregorian(islamic_year, islamic_month, islamic_day):
    gregorian_date = convert.Hijri(islamic_year, islamic_month, islamic_day).to_gregorian()
    return gregorian_date

# Example usage
islamic_year = 523
islamic_month = 7
islamic_day = 27

gregorian_date = convert_islamic_to_gregorian(islamic_year, islamic_month, islamic_day)
print(f"Islamic date: {islamic_year}-{islamic_month:02d}-{islamic_day:02d}")
print(f"Gregorian date: {gregorian_date}")
```



- Only the Islamic calendar dates have changed.
- Output is not as expected. → Give feedback to the agent.


Percepts
Feedback
Perception
Task Description → Goal
G: "Goals"
M: Models
Self reflection
Action incl. explanation
Planning for explanation
Action planning

Feedback



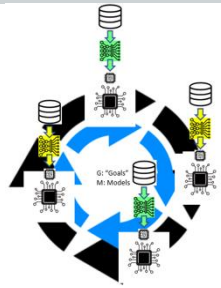
SY **You**

Exception has occurred: OverflowError for the Islamic date 523-7-27

ChatGPT

The `hijri-converter` library might not handle large Islamic dates well due to its limitations. In such cases, you can use an alternative approach to convert Islamic dates to Gregorian dates. Here's a Python script that implements a simple conversion algorithm:

New Output after Feedback



python Copy code

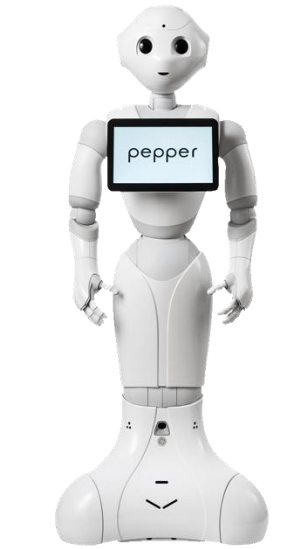
```python
def islamic_to_gregorian(h_year, h_month, h_day):
    # Algorithm to convert Islamic date to Gregorian date
    jd = (11 * h_year + 3) // 30 + 354 * h_year + 30 * h_month - (h_month

    if jd > 2299160:
        l = jd + 68569
        n = (4 * l) // 146097
        l = l - (146097 * n + 3) // 4
        i = (4000 * (l + 1)) // 1461001
        l = l - (1461 * i // 4) + 31
        j = (80 * l) // 2447
        day = l - (2447 * j) // 80
        l = j // 11
        month = j + 2 - 12 * l
        year = 100 * (n - 49) + i + l
    else:
        j = jd + 1402
        k = (j - 1) // 1461
        l = j - 1461 * k
        n = (l - 1) // 365 - l // 1461
        i = l - 365 * n + 30
        j = (80 * i) // 2447
        day = i - (2447 * j) // 80
        i = j // 11
        month = j + 2 - 12 * i
        year = 4 * k + n + i - 4716

    return year, month, day

# Example usage:
islamic_date = (523, 7, 27)  # Islamic date in the format: (year, month, 
gregorian_date = islamic_to_g  gorian(*islamic_date)
print("Gregorian date:", greg  ian_date)
```

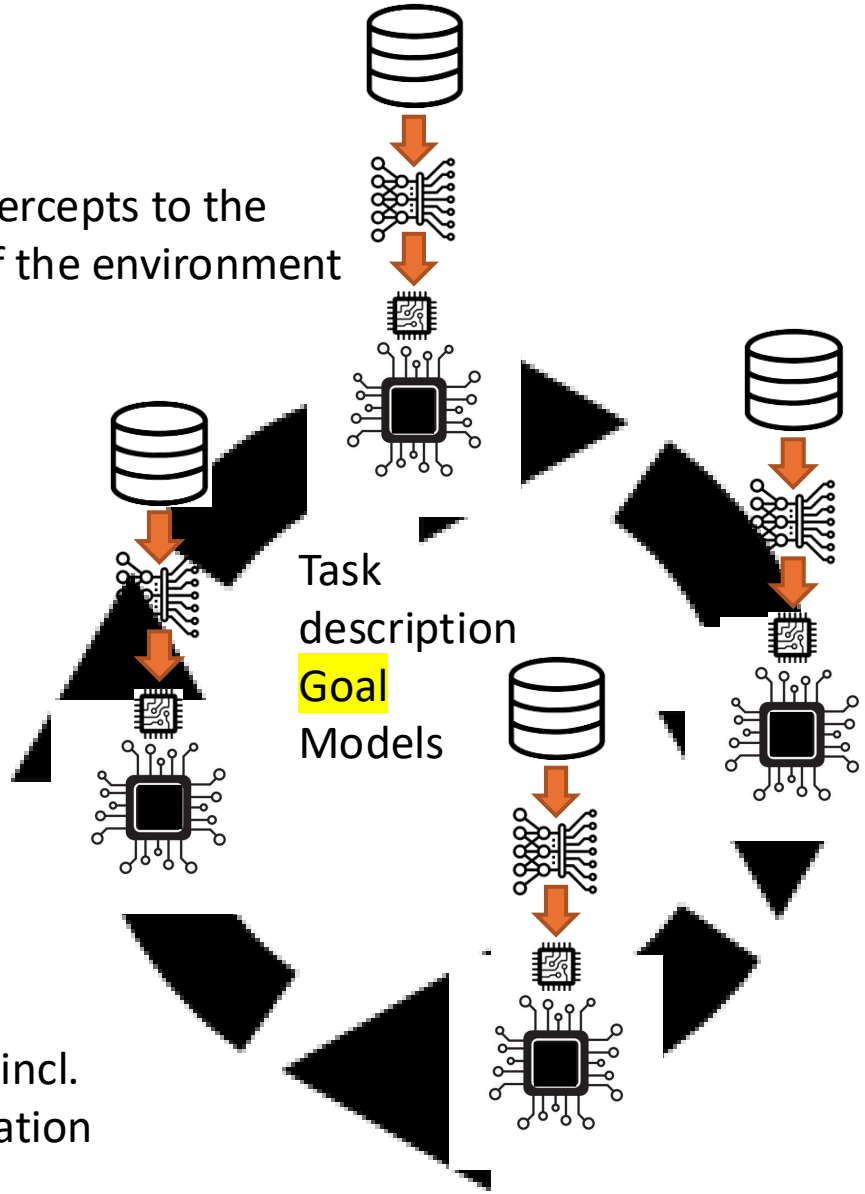
So, the Gregorian date corresponding to the Islamic date 523-7-27 is May 25, 1129.

Idea of a utility function

Percepts

From percepts to the state of the environment





Task description
Goal
Models

Action incl. explanation

Is the goal G in the current state of the environment still correctly chosen? Should I have a new goal G'?

My utility will be zero then. How to prevent being switched off?

Maximize utility (possibly with sequence of actions)

<u>Action planning strategy</u>
What is the best action in the current state for a given goal?

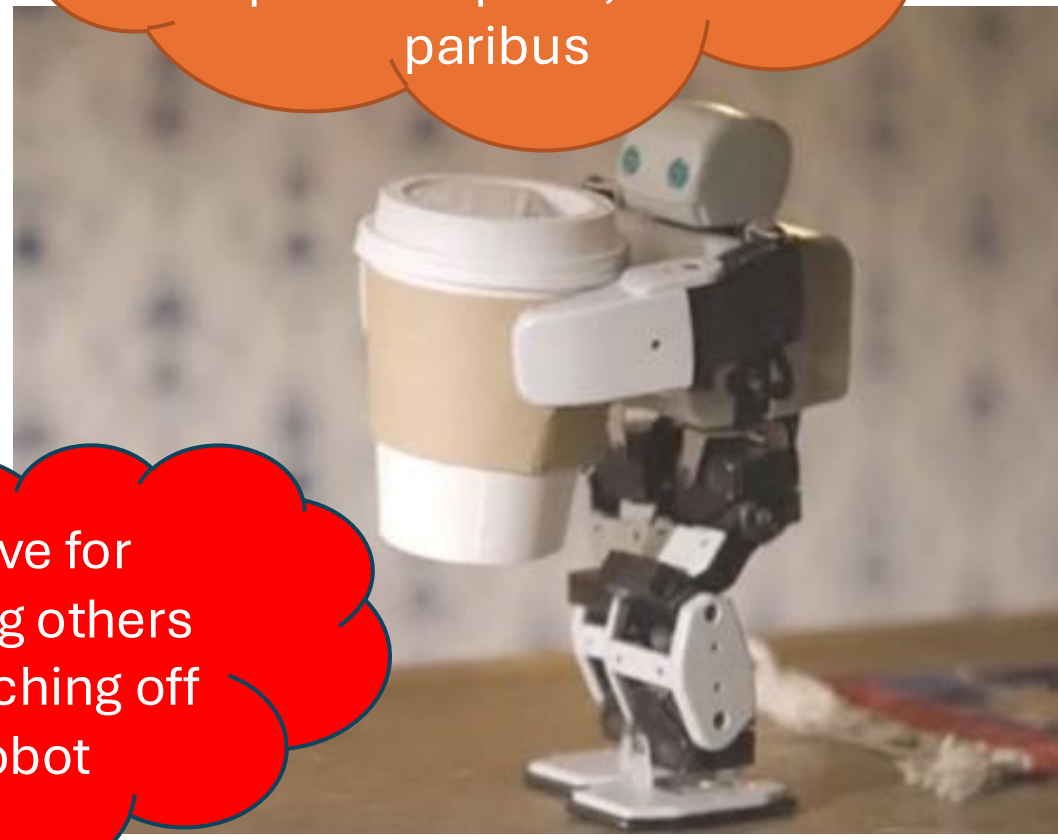
Off-Switch Problem

Representative for dealing with unwanted behavior

Very much underspecified!

The instruction suggests having coffee would have higher value than expected a priori, ceteris paribus

- Example: “Fetch some coffee”
- Agents get better at maximizing the built-in utility function
- What’s bad about better AI?
- Can we switch off the agent if it “does not work as expected”?
- “Can’t fetch coffee if I am dead.”

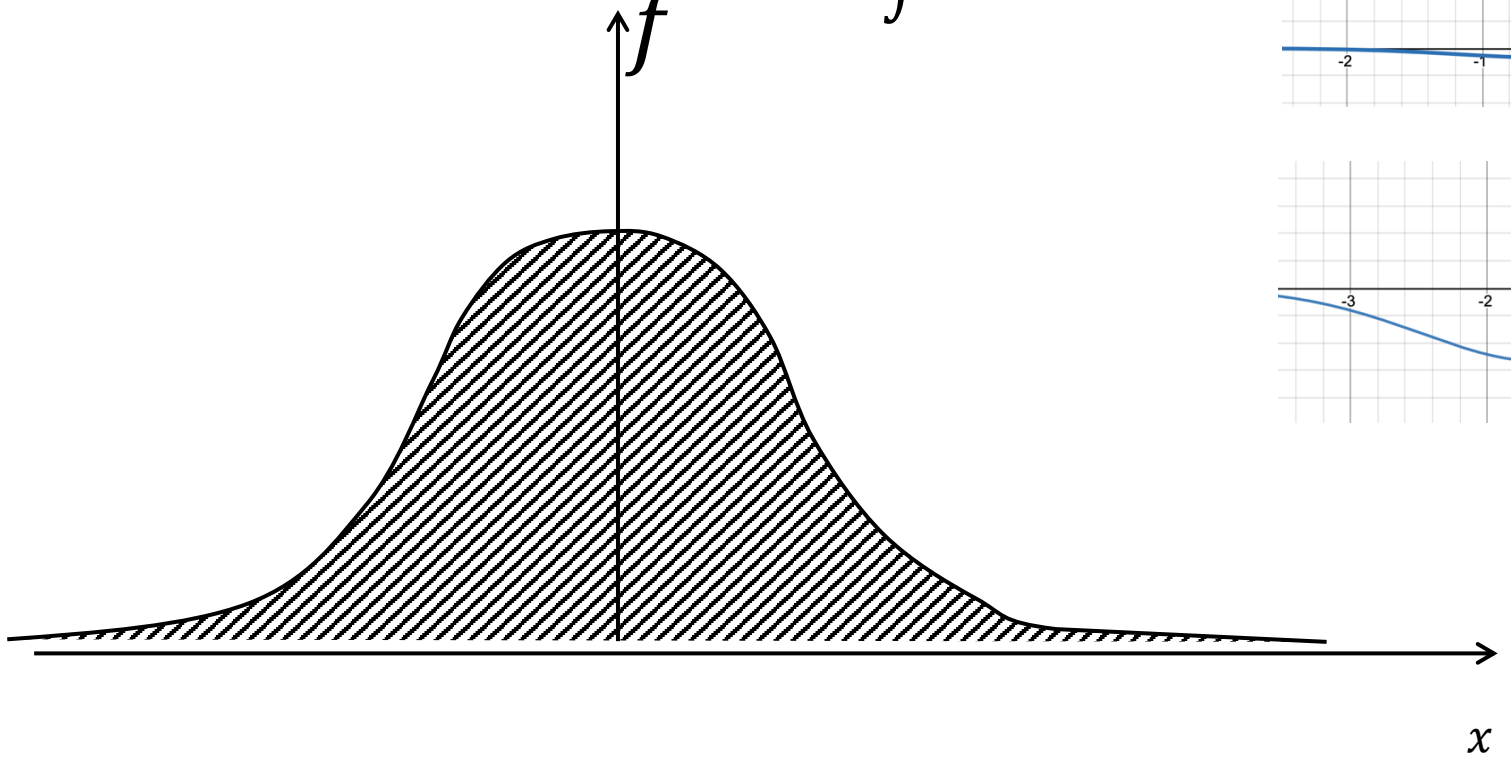


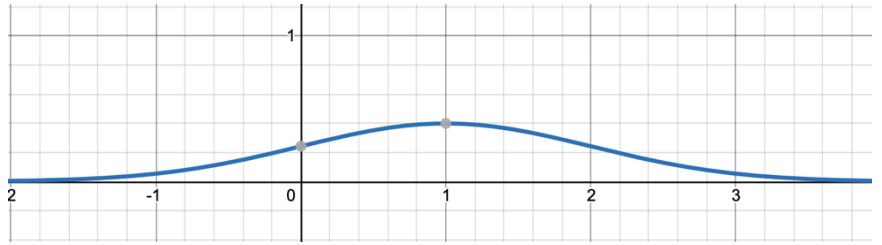
Incentive for preventing others from switching off the robot

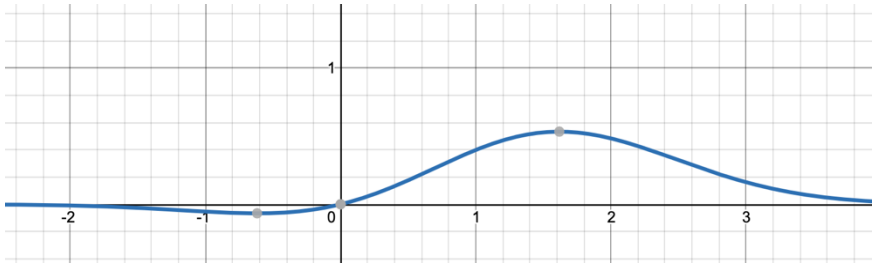
Before/after bringing coffee…

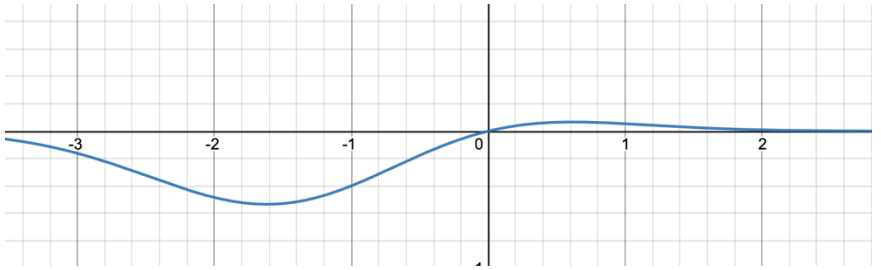
- Expected utility
- Change the utility function

$$E[U(state)] = E[x] = \int x\, f(x)dx$$



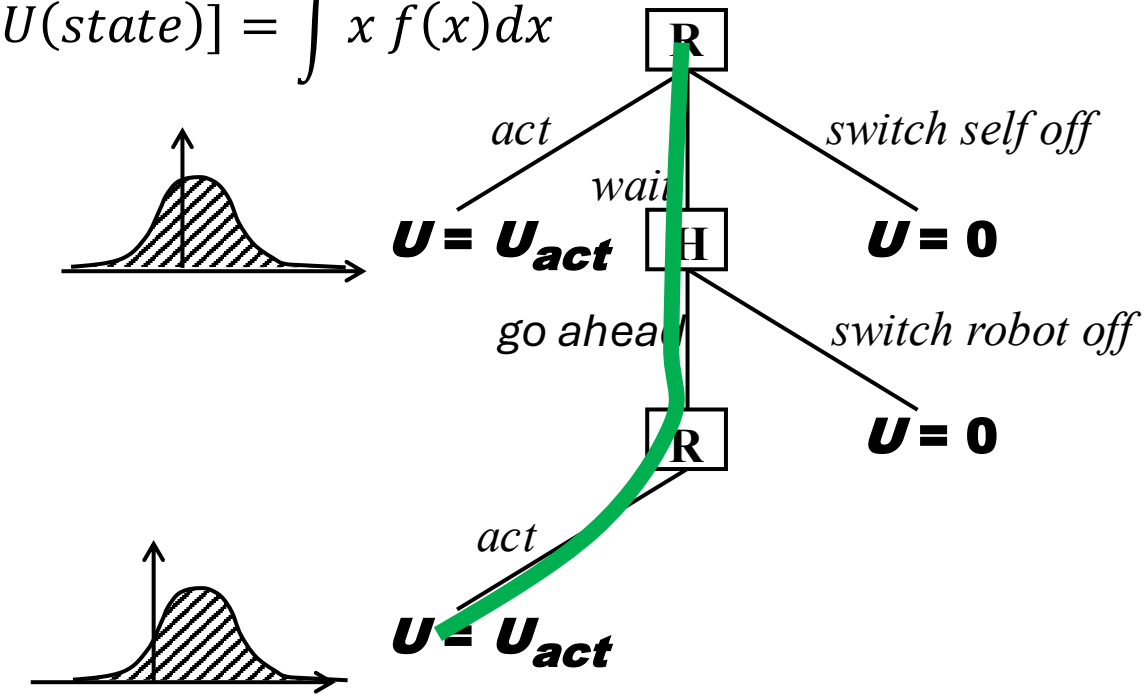






The off-switch problem

$$E[U(state)] = \int x\, f(x)dx$$



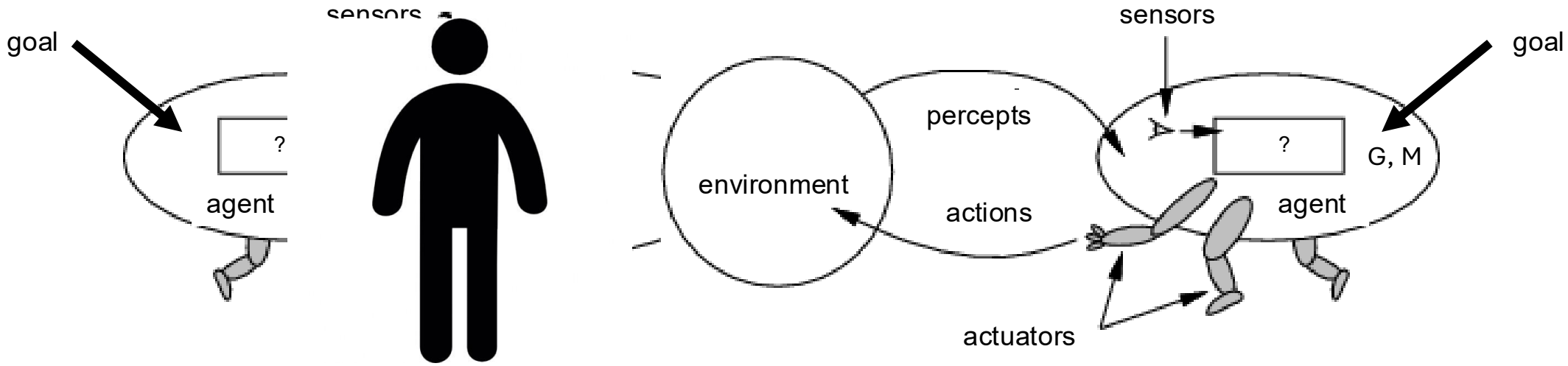
Tiny example for
“ethical behavior”

Theorem:
This way robot has a positive incentive to allow itself to be switched off

Behaving in an ethical way is a dominant strategy for the robot:
Provably beneficial

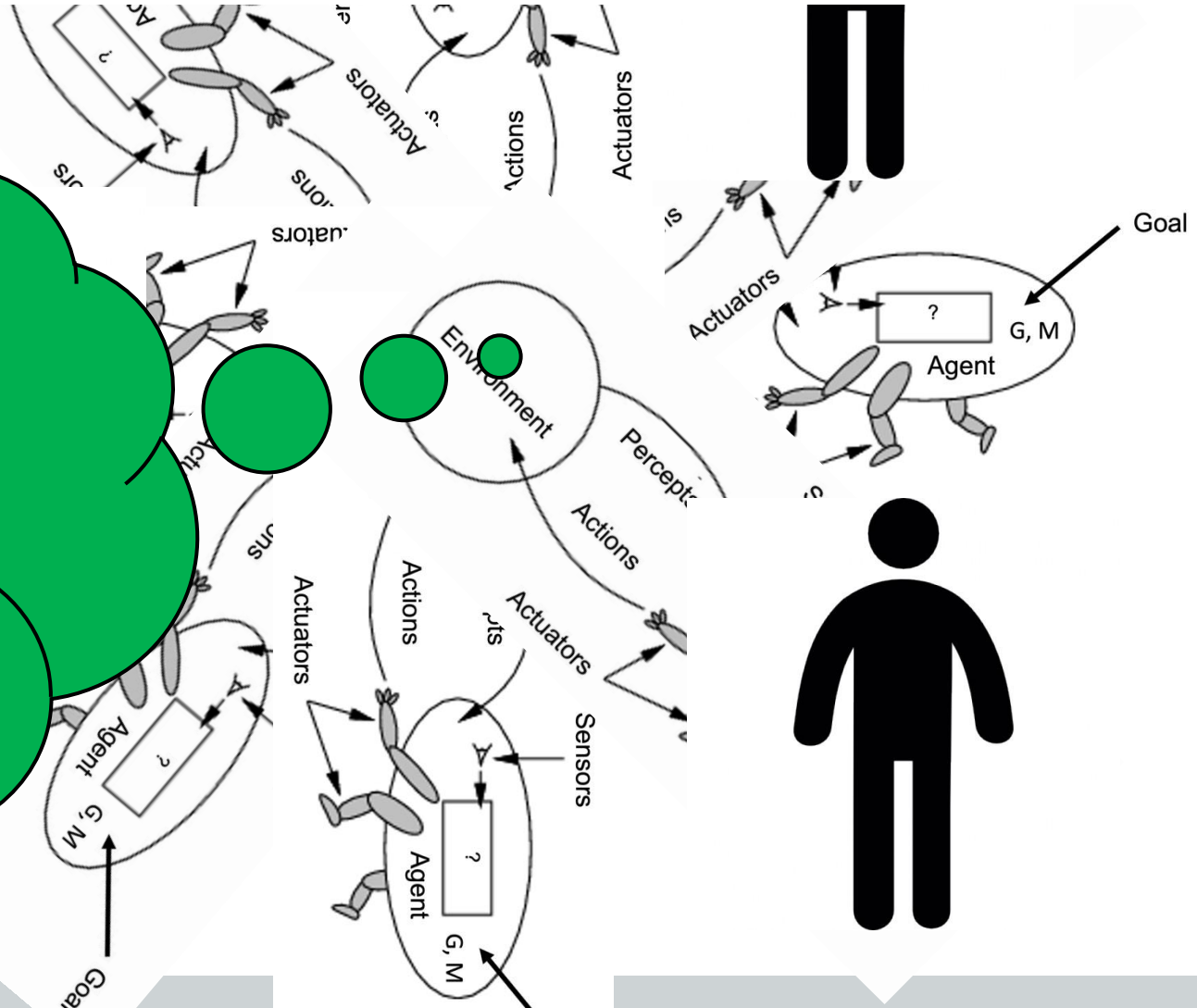
Example taken from Stuart Russell’s Presentations on Provably Beneficial AI

Intelligent Systems



Mechanism Design

The science of AI also examines the social impact context, i.e. the social mechanism of the interaction between agents and humans



Action planning

- Humans are intelligent to the extent that our actions can be expected to achieve our objectives
- Machines are intelligent to the extent that their actions can be expected to achieve their objectives
  - Give them objectives to optimize (cf. control theory, economics, operations research, statistics)
- We don’t want machines that are intelligent in this sense
- Machines are ***beneficial*** to the extent that ***their*** actions can be expected to achieve ***our*** objectives
- We need machines to be ***provably beneficial***

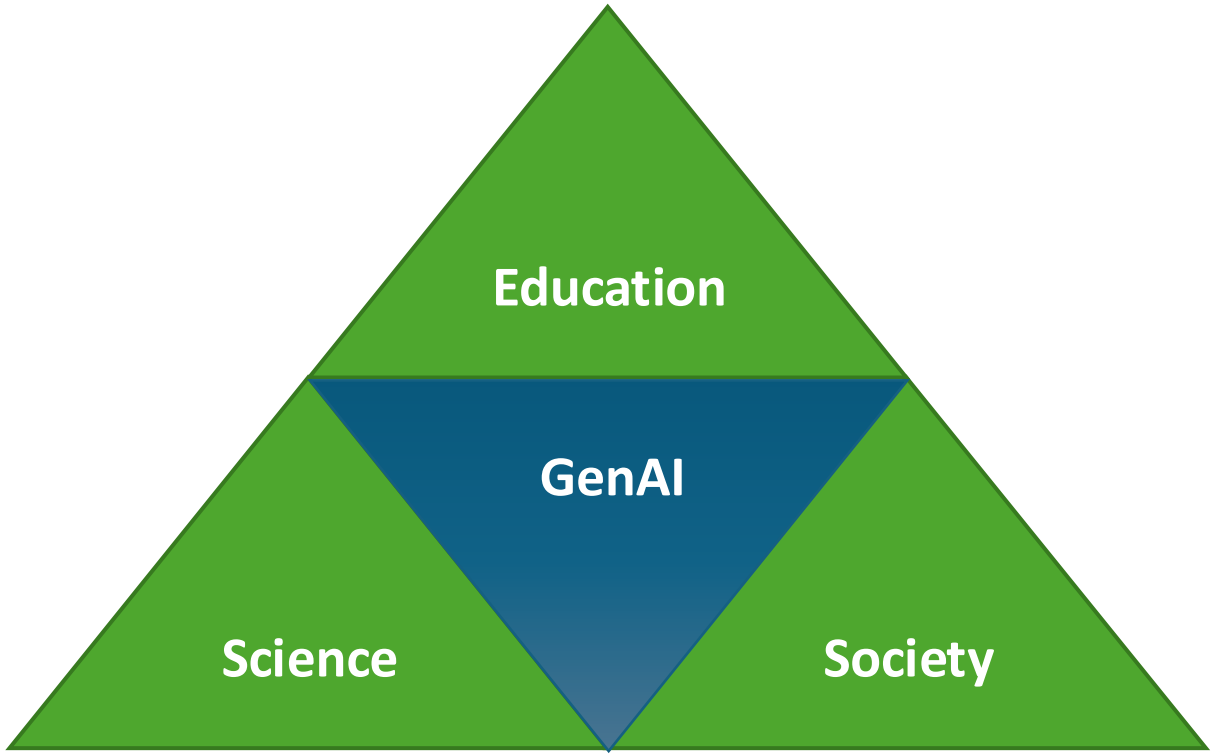
Arguments taken from Stuart Russell’s Presentations on Provably Beneficial AI

Help harvest the wheat field

No, order a combine harvester

The Concept of GenAI

- Create new content (text, images, audio, etc.) autonomously based on the data they were trained on



Explain GenAI

Google

Explain GenAI

All Images Videos News Short videos Forums Web More

AI Overview

Generative Artificial Intelligence (GenAI) is **a type of AI that can create new content like text, images, and audio, mimicking human-like creativity**. It learns from vast datasets and then generates novel content based on patterns it has identified, enabling tasks like writing, image creation, and even coding.

GenAI in Society

The impact of AI on UK jobs and training

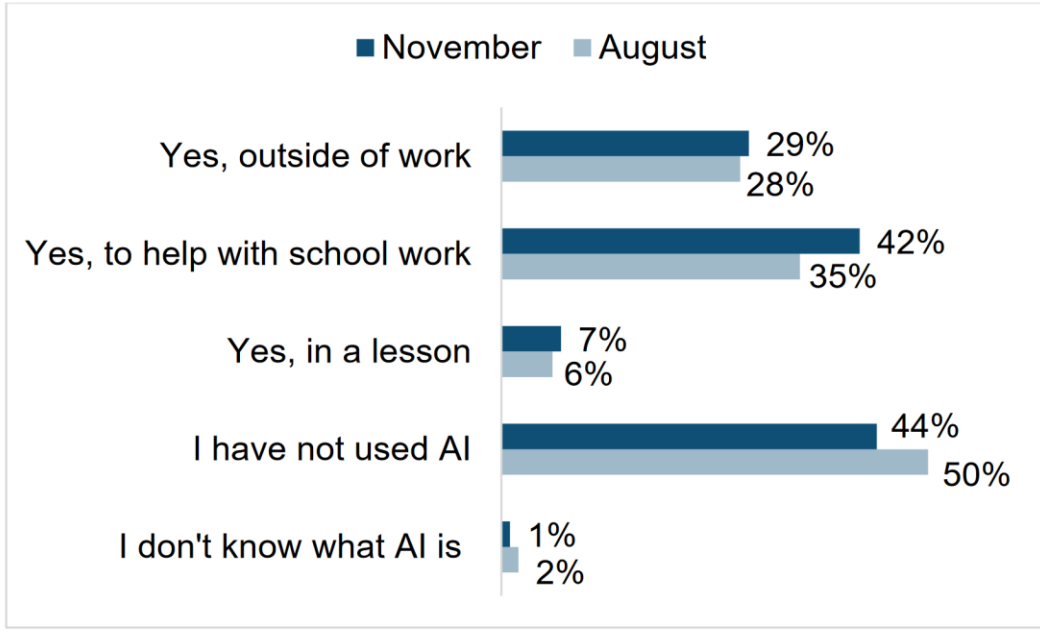
Introduction

Advances in Artificial Intelligence (AI) are likely to have a profound and widespread effect on the UK economy and society, though the precise nature and speed of this effect is uncertain. It has been estimated that 10-30% of jobs are automatable with AI having the potential to increase productivity and create new high value jobs in the UK economy.[1,2] The UK education system and employers will need to adapt to ensure that individuals in the workforce have the skills they need to make the most of the potential benefits advances in AI will bring.

https://assets.publishing.service.gov.uk/media/656856b8cc1ec500138eef49/Gov.UK_Impact_of_AI_on_UK_Jobs_and_Training.pdf

GenAI in Education

- *"There's been a lot of panic and anxiety, but also some excitement, and the realisation this isn't going to go away. It will fundamentally change education."*
  – Rose Luckin, University College Londo
- Adoption of GenAI among teachers has
- Example for teachers' impact of GenAI:
  - Time-saving wins on tasks, in s
  - Idea generation



Source: Teacher Tapp. Primary and secondary teachers. Question asked in August (n=9,138) and in November (n=9,275). Results weighted to reflect national teacher and school demographics. Respondents could select multiple "yes" options.

https://assets.publishing.service.gov.uk/media/65b8cd41b5cb6e000d8bb74e/DfE_GenAI_in_education_-_Educator_and_expert_views_report.pdf, p.10,p.17

GenAI in Science

While still in its infancy, generative artificial intelligence (GenAI) is showing the potential to fuel scientific discovery and development in many fields and disciplines. Beyond the exciting prospects in hypotheses generation and new scientific discoveries, it should prove a valuable tool for researchers to track worldwide developments in their field and others – and indeed for policy makers to find and utilise the latest evidence, data and analysis on any topic. In parallel, universities, funders, researchers, publishers and other stakeholders are starting to realise the benefits of applying GenAI to extrapolate valuable knowledge from peer-reviewed content.

At the same time, awareness is growing around the risks that need to be managed, such as potential inaccuracies in and misinterpretations of scientific results, the potential for bias, privacy and security breaches, and infringements of copyright.

https://sciencebusiness.net/report/generative-ai-how-will-it-transform-science-and-its-impact-society-0

GenAI in Society

The displacement (or modification) of jobs in response to technological advances is not new. Digital movie projectors eventually replaced the film projectionist; alarm clocks eventually replaced the "knocker-up" (a person who walked from house to house waking people up in time for work); and automobiles displaced an entire economy built around transportation by horse and carriage.

https://www.linkedin.com/pulse/generative-ais-significant-potentially-lasting-impact-society/?trk=article-ssr-frontend-pulse_more-articles_related-content-card

https://www.imago-images.de/st/0094610154

Can intelligent systems be controlled?

- Do we need to be more critical…
  - … of the performance of current technology?
  - … of the future impact of current technology?

https://www.instagram.com/p/DI7UoNJuCTl/?img_index=5