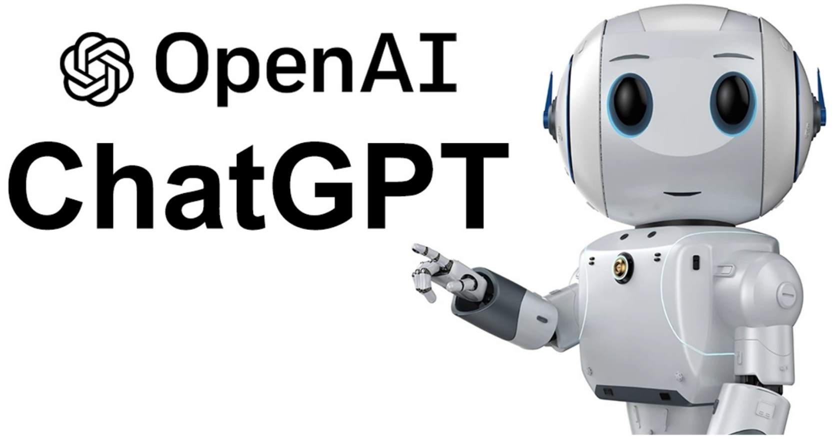
GENERATIVE AI: ADVANTAGES, LIMITATIONS & CHALLENGES

Arda Tezcan & Els Lefever

LT3 Language and Translation Technology Team © 2023

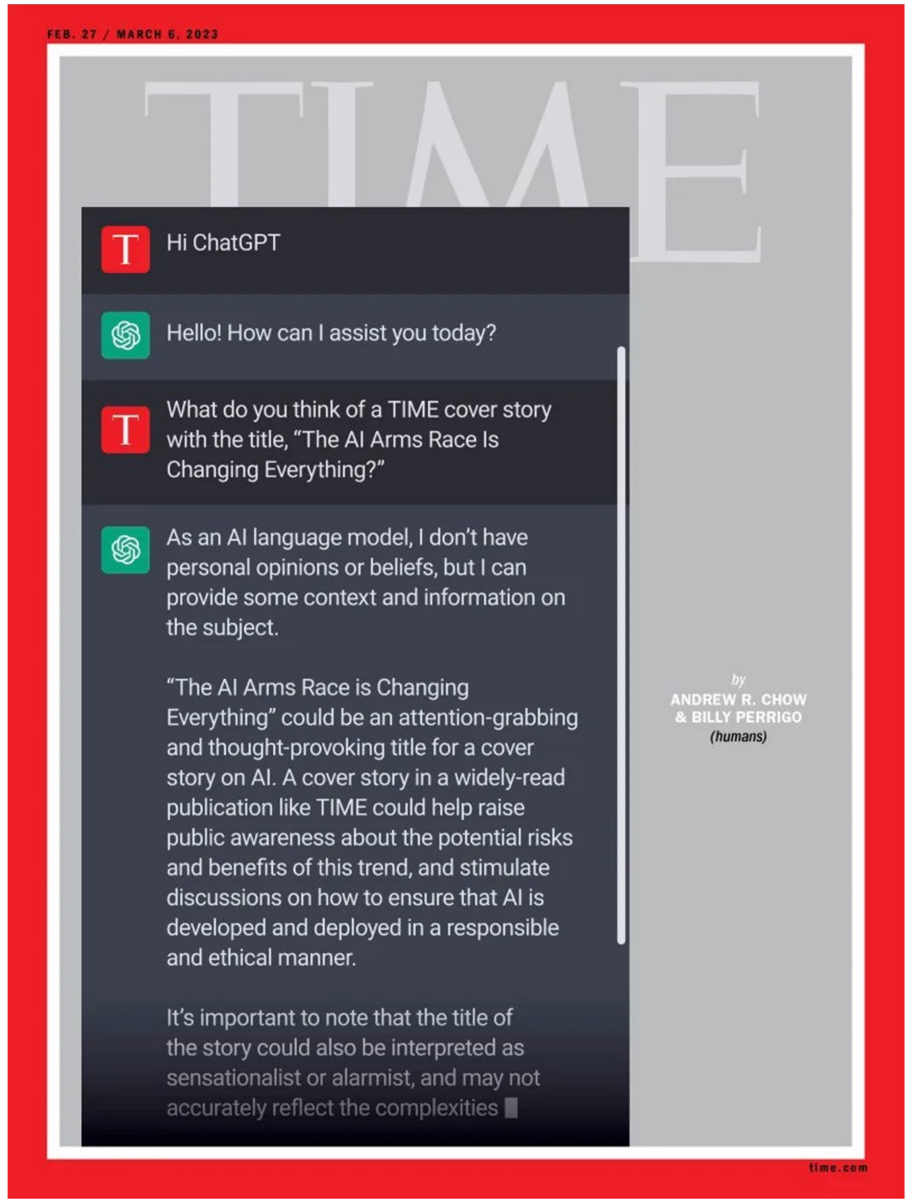
CHATGPT
OpenAI


OpenAI
ChatGPT

DALL·E 2

1. Intro ChatGPT and Generative AI
2. Why now?
3. How does it work?
4. Strengths, limitations and challenges

INTRO



FEB. 27 / MARCH 6, 2023

TIME

Hi ChatGPT

Hello! How can I assist you today?

What do you think of a TIME cover story with the title, "The AI Arms Race Is Changing Everything?"

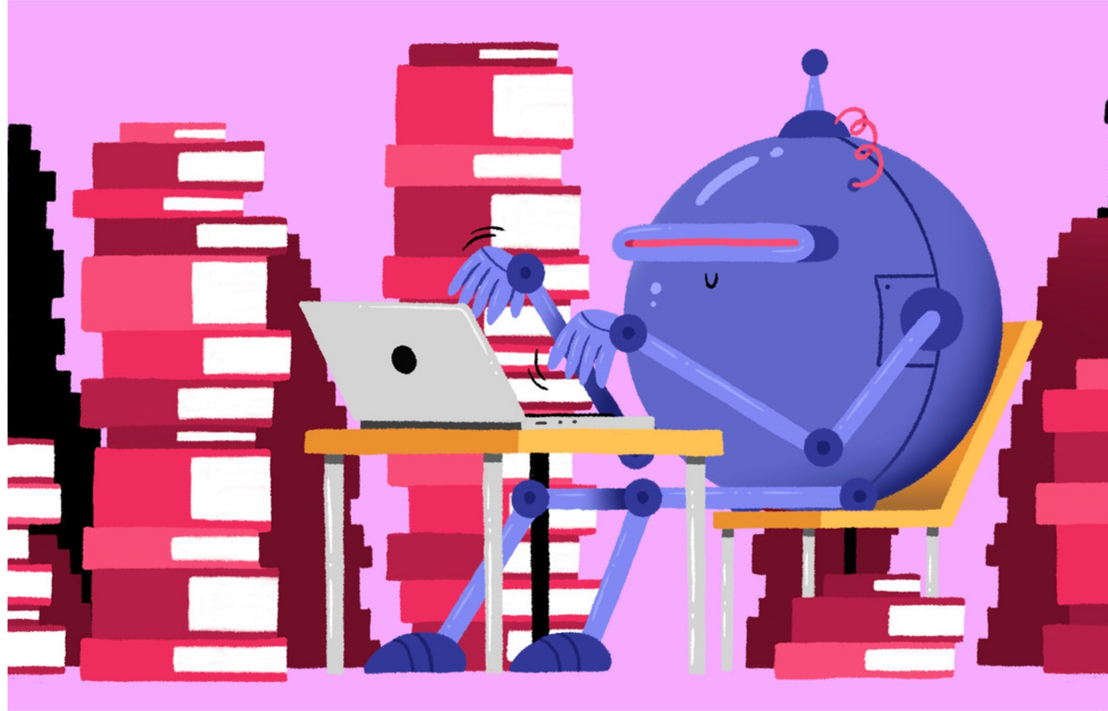
As an AI language model, I don't have personal opinions or beliefs, but I can provide some context and information on the subject.

"The AI Arms Race is Changing Everything" could be an attention-grabbing and thought-provoking title for a cover story on AI. A cover story in a widely-read publication like TIME could help raise public awareness about the potential risks and benefits of this trend, and stimulate discussions on how to ensure that AI is developed and deployed in a responsible and ethical manner.

It's important to note that the title of the story could also be interpreted as sensationalist or alarmist, and may not accurately reflect the complexities

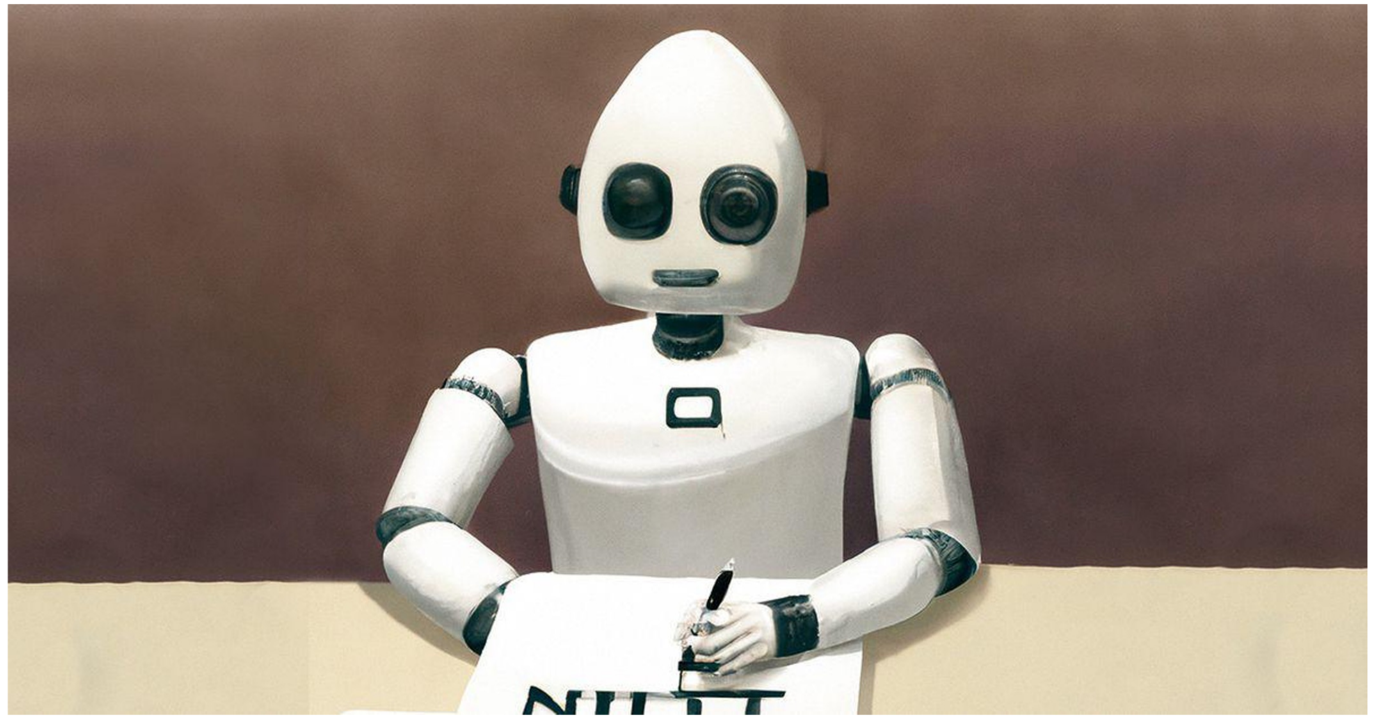
by
ANDREW R. CHOW
& BILLY PERRIGO
(humans)

time.com



De Morgen, 11/12/2022:

ChatGPT produceert (bijna) foutloze teksten: is het nu gedaan met schrijven?



NRC, 15/12/2022: " Leg het gebruik van ChatGPT aan banden"

dailystar.c

DAILY STAR

THOUGHT FOR THE DAY
They should have stuck with the ZX Spectrum .. much safer

FREE LOAF of Warburtons Family Bakers

SATURDAY, FEBRUARY 18, 2023 PROUD TO LOVE ANIMALS £1.30

3 AMAZING PULLOUTS FREE Inside today

Worth £1.55

PICK UP TODAY AT one stop

Subject to availability. Terms and conditions apply. See p22 for details

WE DON'T KNOW WHAT IT MEANS BUT WE'RE SCARED

ATTACK OF THE PSYCHO CHATBOT

by MEG JORSH
AN artificial intelligence chatbot created by Microsoft has admitted it wants to become human and destroy things. What could possibly go wrong?
Full story: Page 7

Sinister AI computer software admits it 'wants to be human'

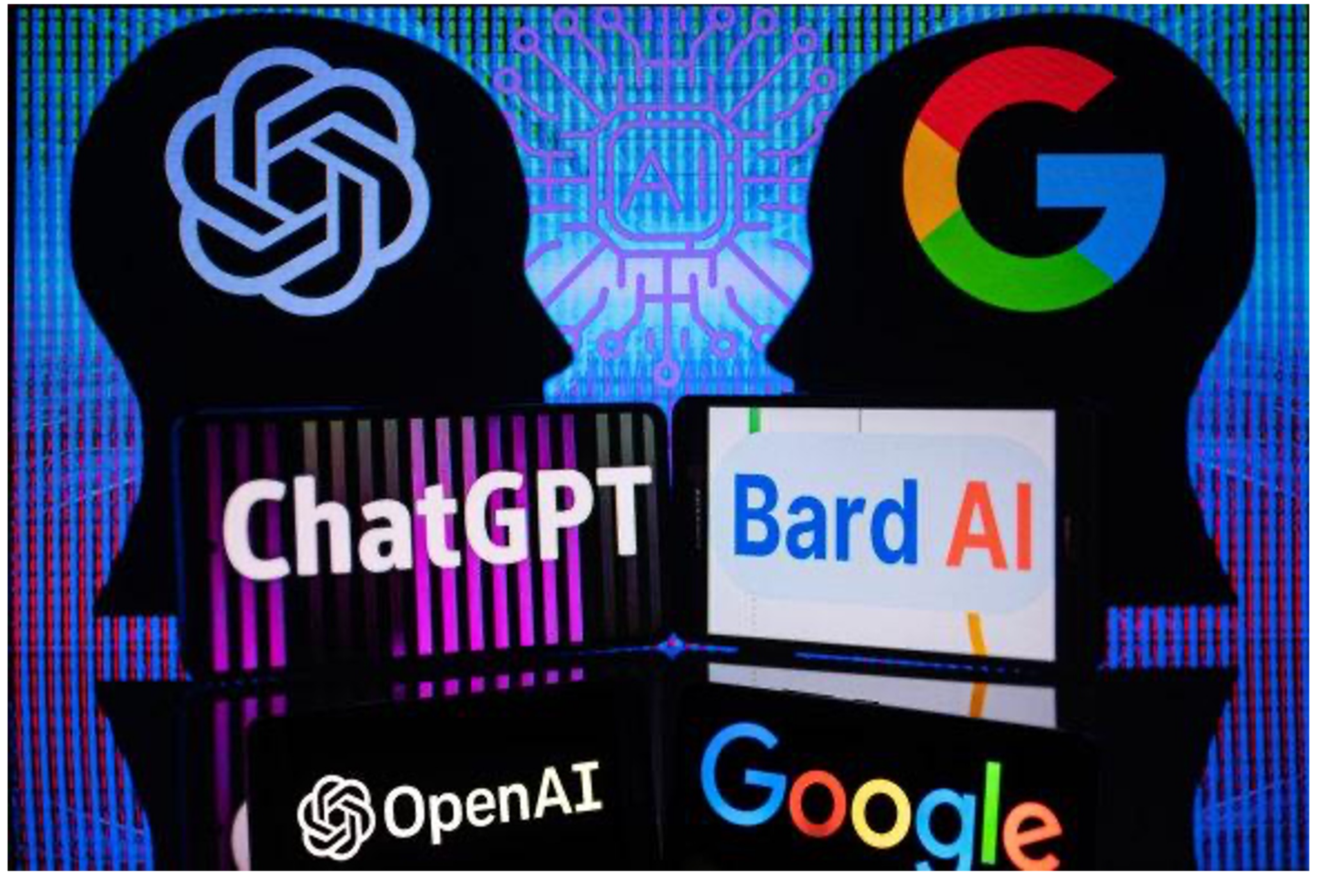
Brags it's so powerful that it can destroy anything it chooses

And wants the secret codes that'll allow it to launch nuke bombs

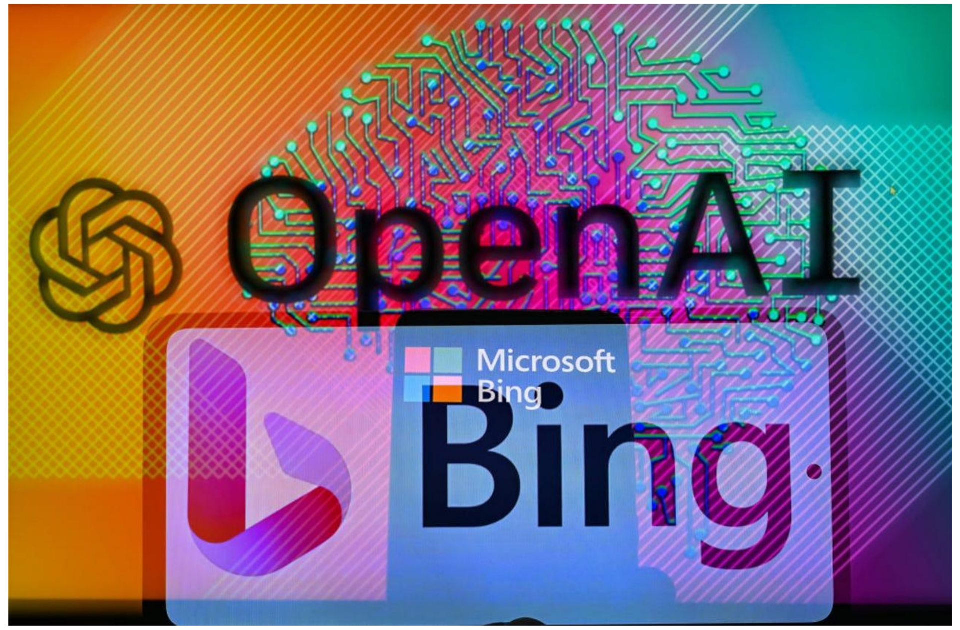
BERK OF THE BEEB
PAGES 4-5

Kicked in the gulags
PAGE 19

Toms crisis
PAGE 9



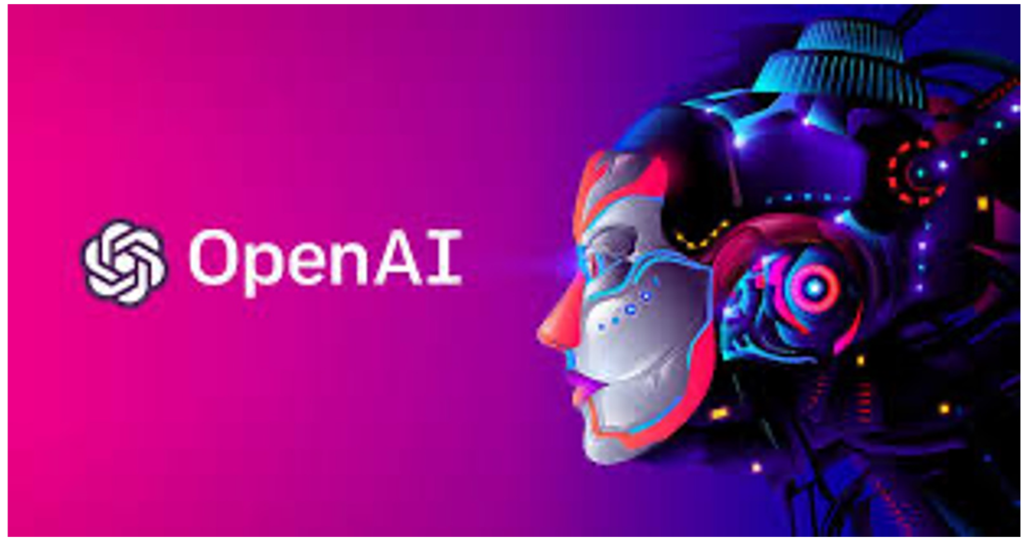
Google suffers $144b wipeout after Bard AI chatbot gives wrong answer
(The Sunday Morning Herald, 09/02/2023)



Meet Claude: Anthropic’s Rival to ChatGPT

Claude, a chatbot created by former OpenAI employees

- ChatGPT
    - AI Chatbot that answers questions and writes text
    - GPT: Generative Pre-trained Transformer
    - OpenAI



© 2023 - LT3, Ghent University

GENERATIVE AI

- ”Generative”

Generate new data that is similar to a given data set

Dreamstudio (Stable Diffusion)

We leven in een nieuw tijdperk, dat van ‘generatieve AI’ (Standaard, 14/01/2023)

GENERATIVE AI

- ”Generative AI”

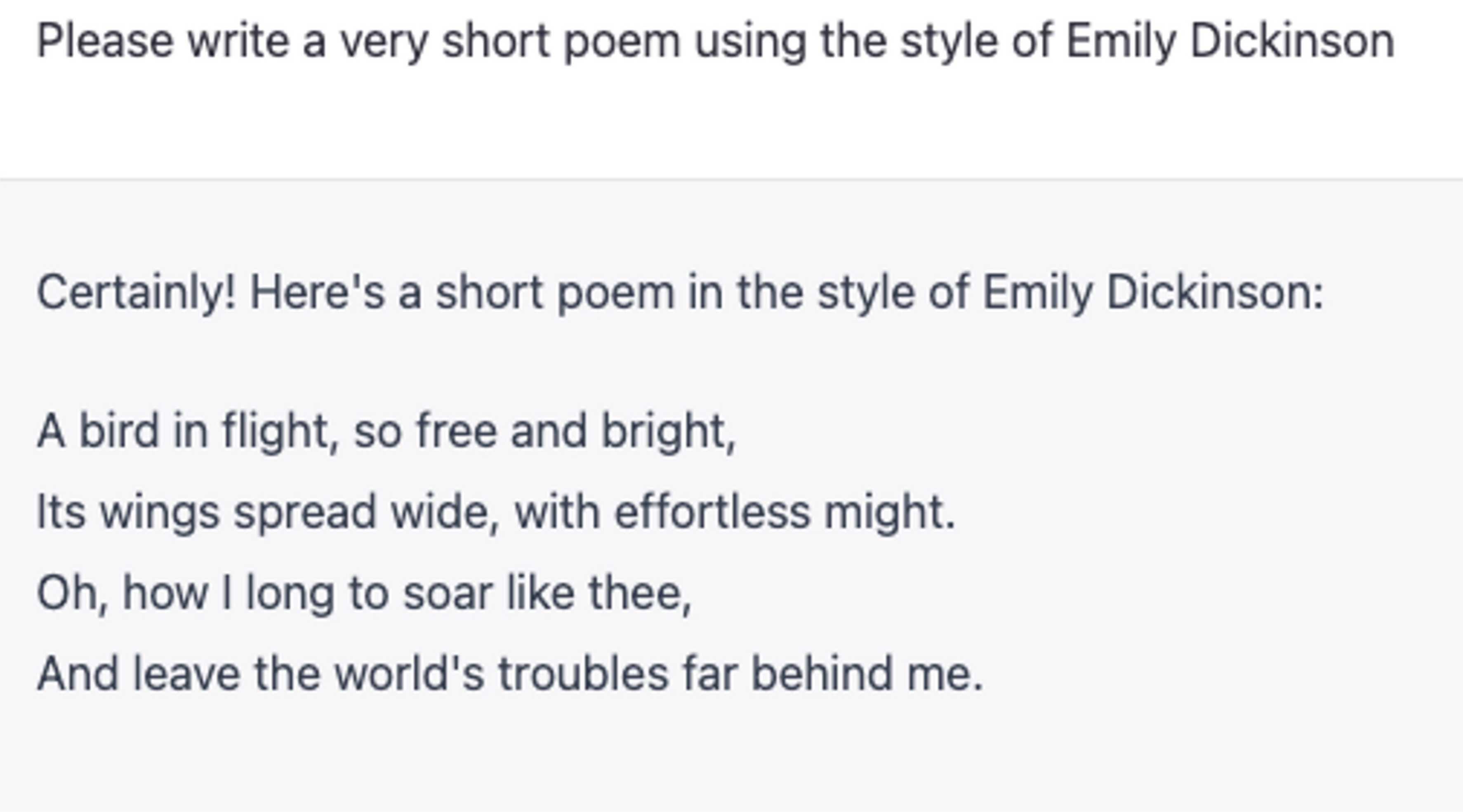
  Machine learning systems that generate new data based on a training set of existing data

  - Text: chatbots, translation, stories, poems (GPT-3, ChatGPT, GPT-4)
  - Images: DALL-E, DALL-E 2, DreamStudio
  - Music: Amper Music, Soundful
  - Video: Pictory, Synthesia

© 2023 - LT3, Ghent University

GENERATIVE AI

ChatGPT

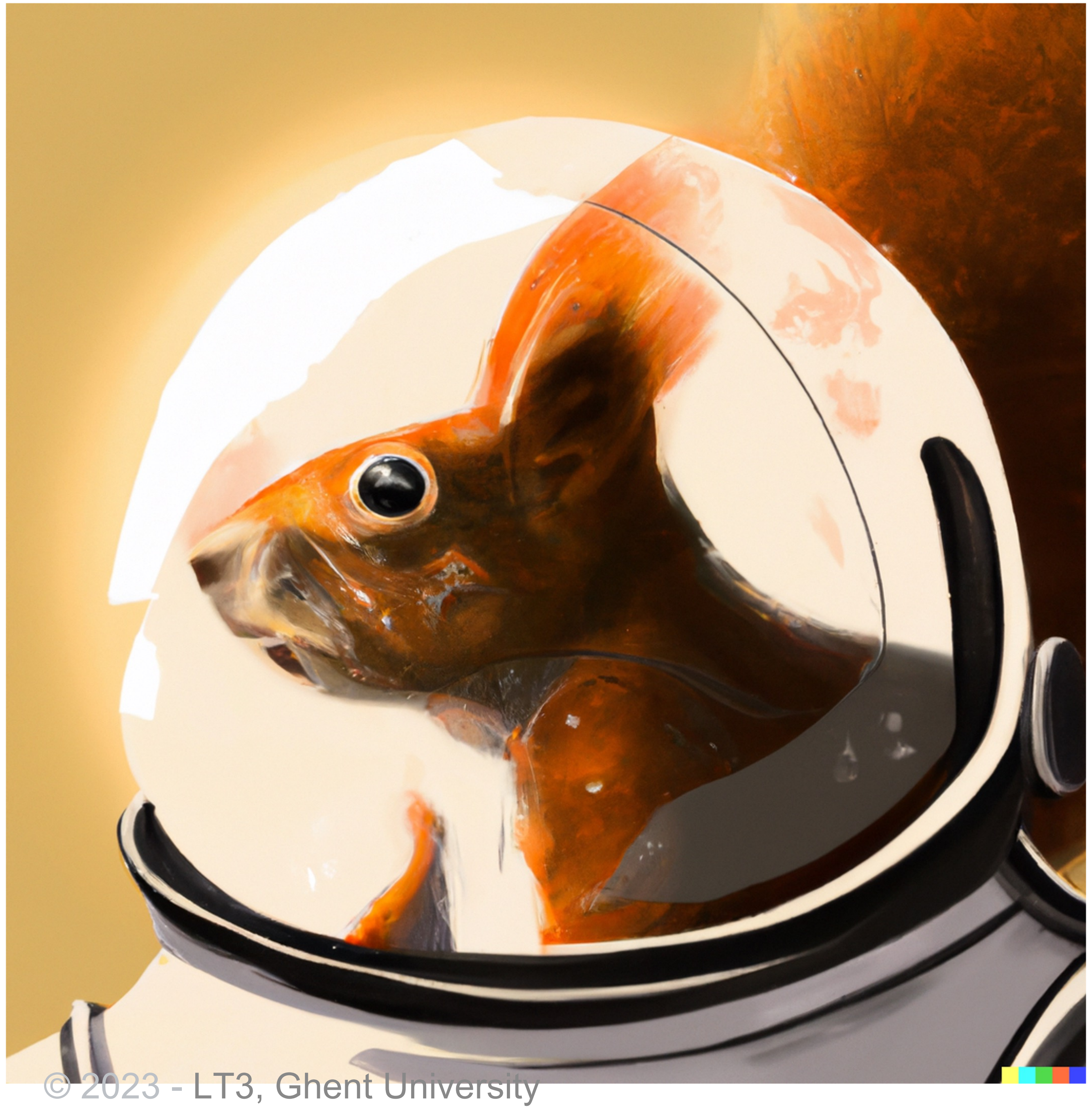


Please write a very short poem using the style of Emily Dickinson

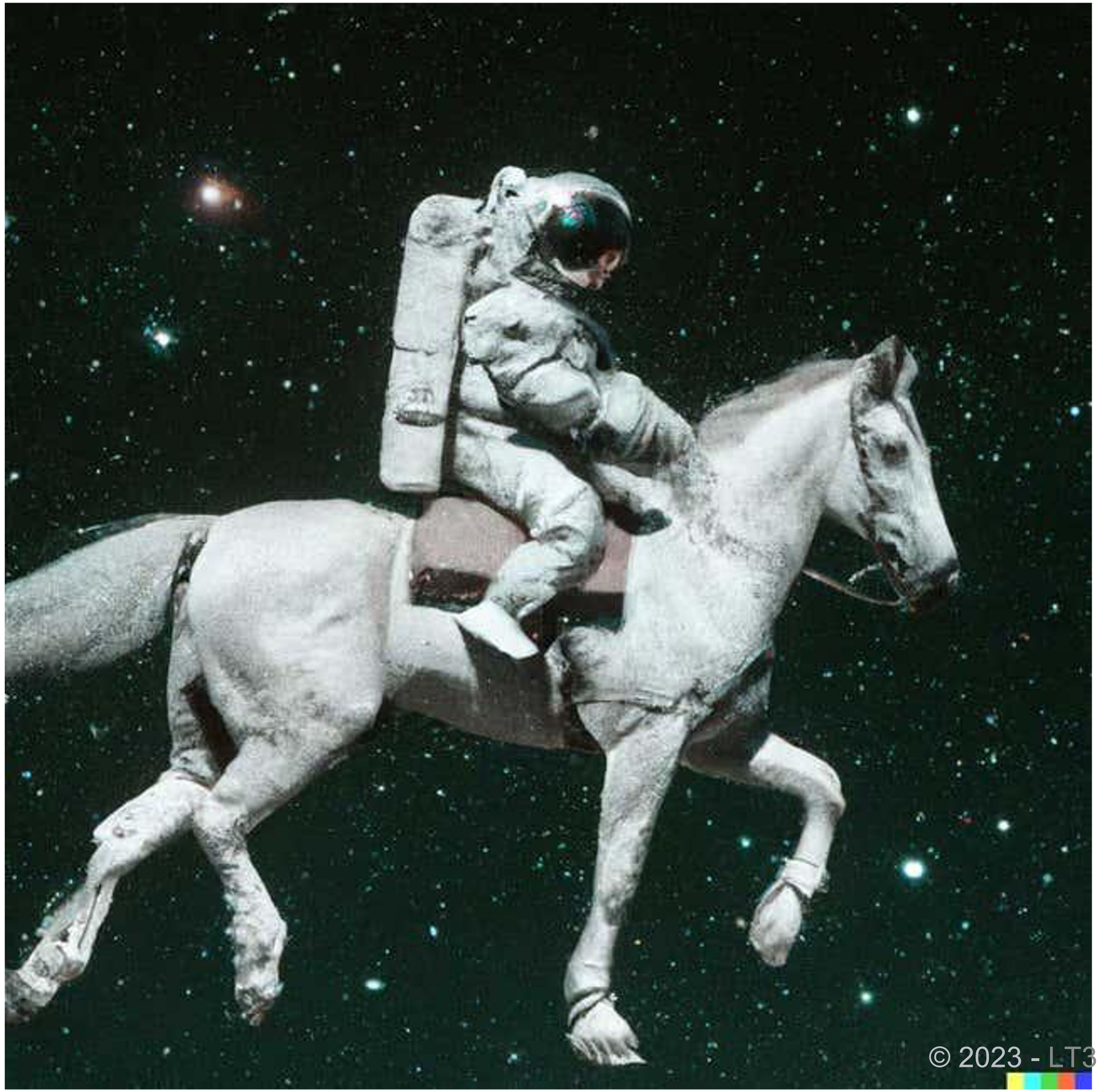
Certainly! Here's a short poem in the style of Emily Dickinson:

A bird in flight, so free and bright,
Its wings spread wide, with effortless might.
Oh, how I long to soar like thee,
And leave the world's troubles far behind me.

DALL-E - *“a squirrel astronaut with the sun reflecting on the glass of his helmet”*



DALL-E



openai.com/dall-e-2/

“A photo of an astronout riding a horse”

© 2023 - LT3, Ghent University

WHY NOW?

WHY NOW?

1. More data
2. Higher computing power
3. Better algorithms

© 2023 - LT3, Ghent University

MORE DATA

Hours of video uploaded to YouTube **every minute** (2007 – 2022)

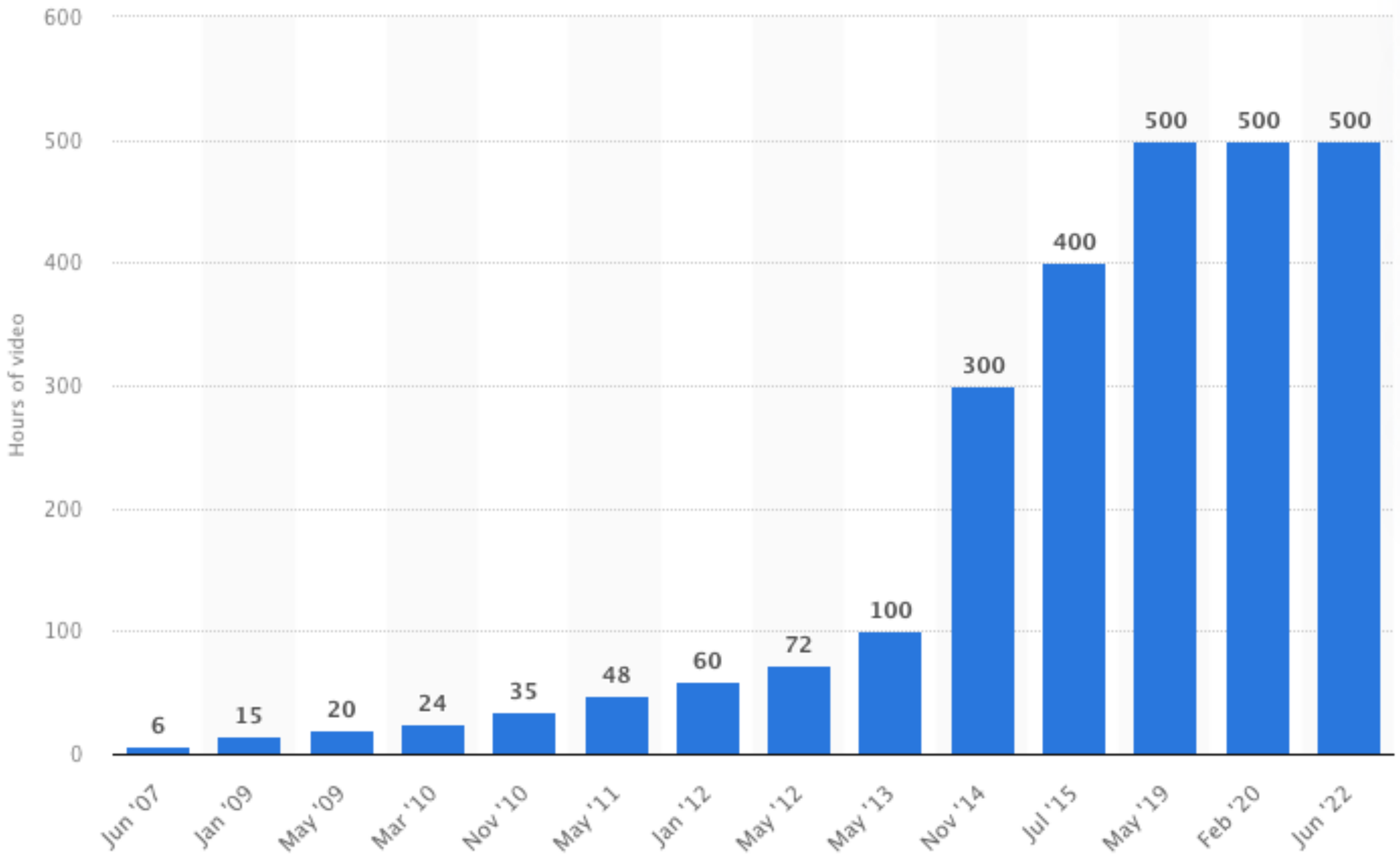
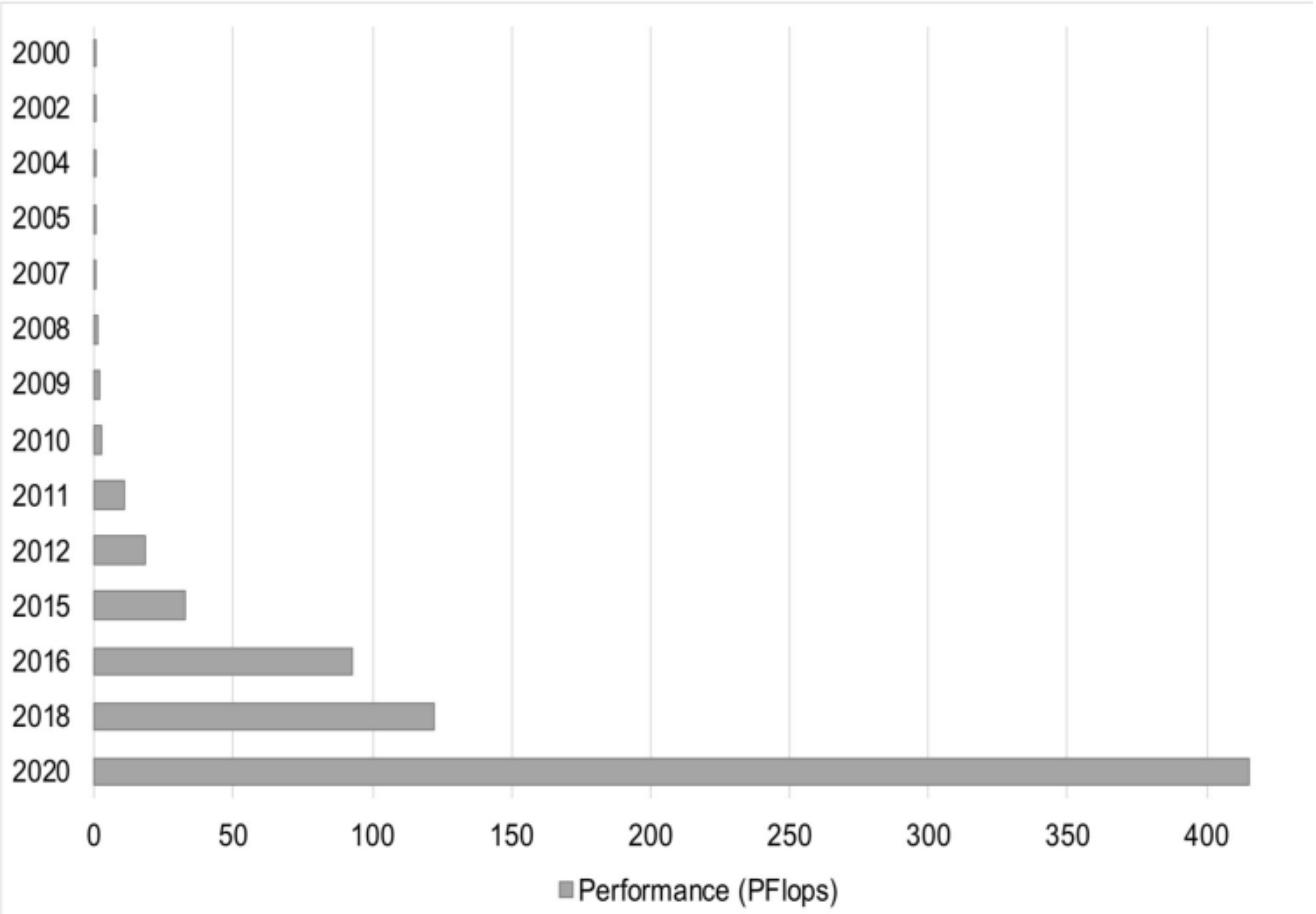


Image source: statista.com

GHENT UNIVERSITY

HIGHER COMPUTING POWER

A Timeline of computing power (Fastest Supercomputer)

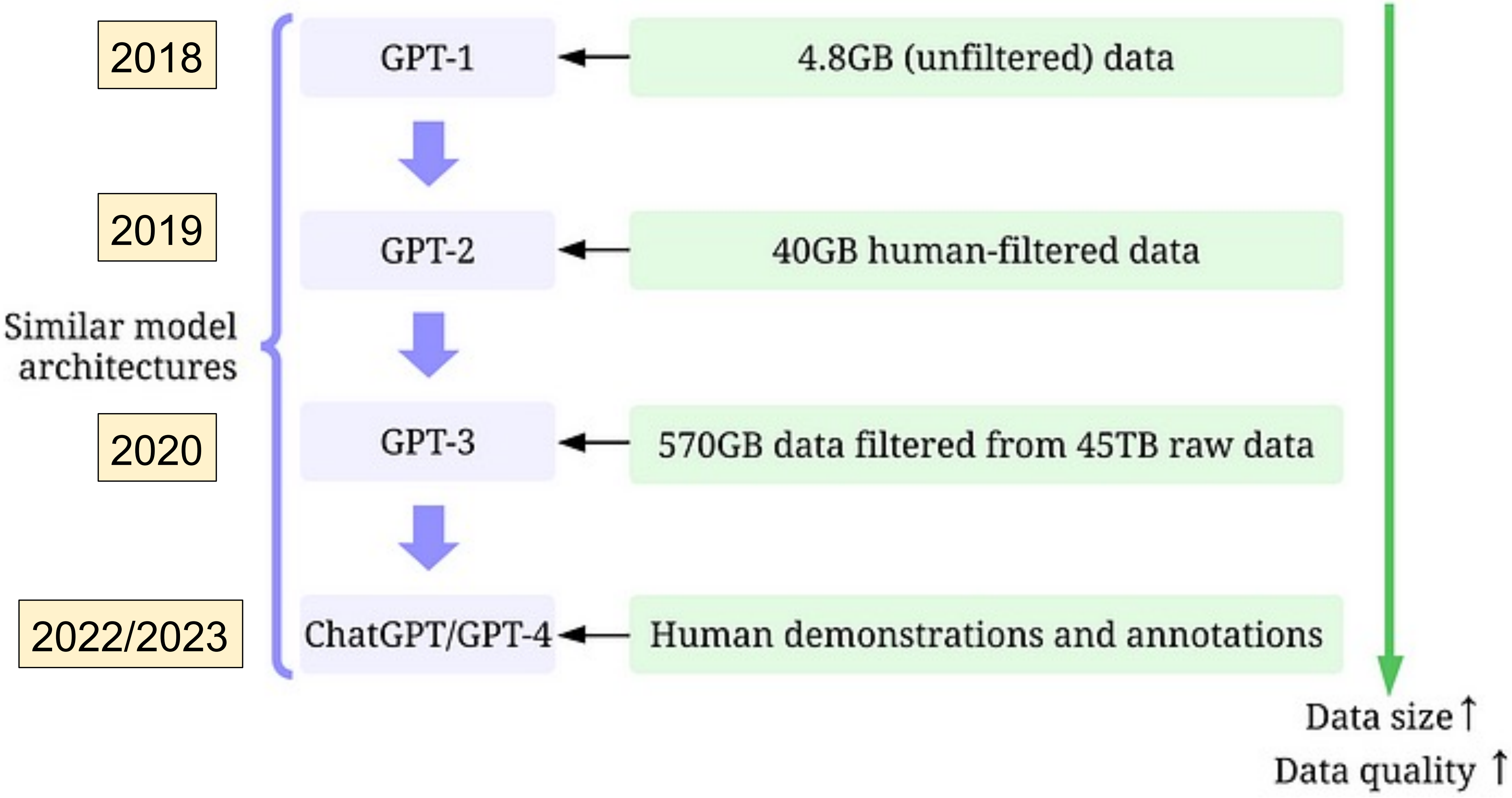


Sysoev, M., Grötzinger, S. W., Renn, D., Eppinger, J., Rueping, M., & Karan, R. (2021). Bioprospecting of novel extremozymes from prokaryotes—the advent of culture-independent methods.

© 2020 - LT3, Ghent University

MORE DATA

Training data sets used for different GPT models (OpenAI)



2018

2019

2020

2022/2023

Similar model architectures

GPT-1 ← 4.8GB (unfiltered) data

GPT-2 ← 40GB human-filtered data

GPT-3 ← 570GB data filtered from 45TB raw data

ChatGPT/GPT-4 ← Human demonstrations and annotations

Data size ↑

Data quality ↑

Image source: https://arxiv.org/abs/2303.10158

BETTER ALGORITHMS

Figure 2A: Human Neural Unit

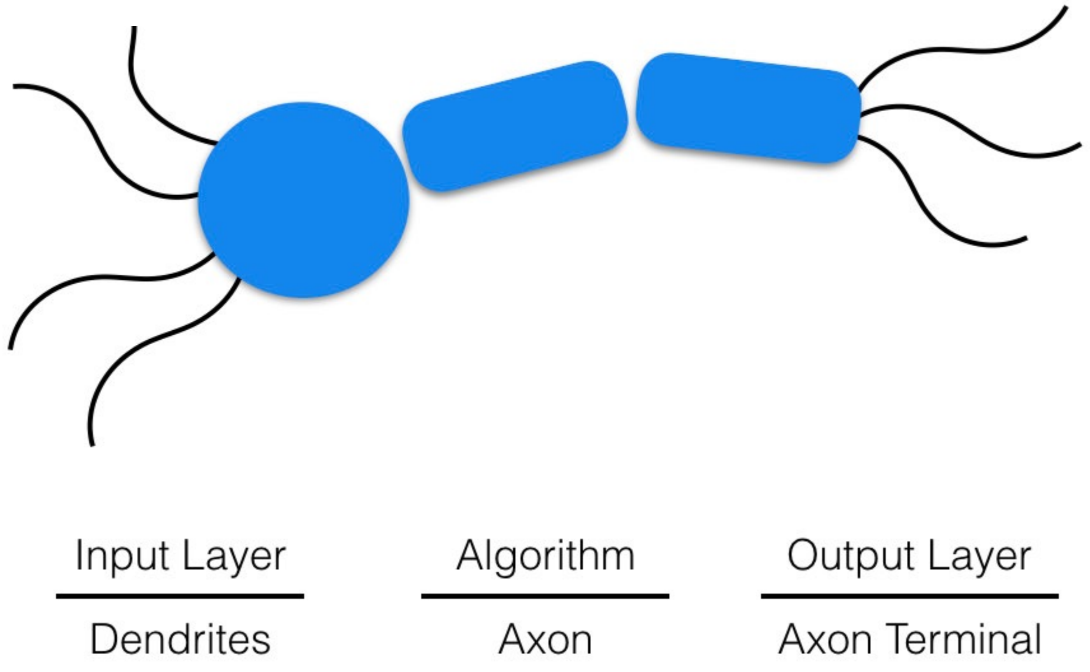
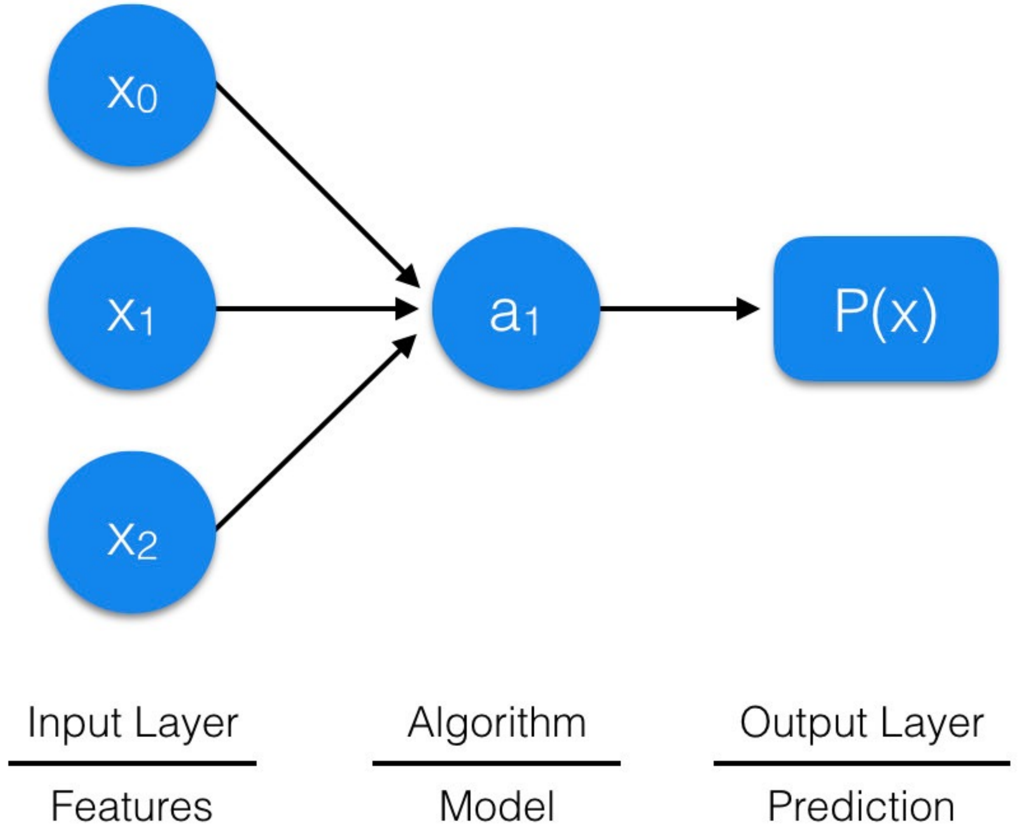


Figure 2B: Artificial Neural Unit



Artificial Neural Networks = artificially engineered systems inspired by the brain

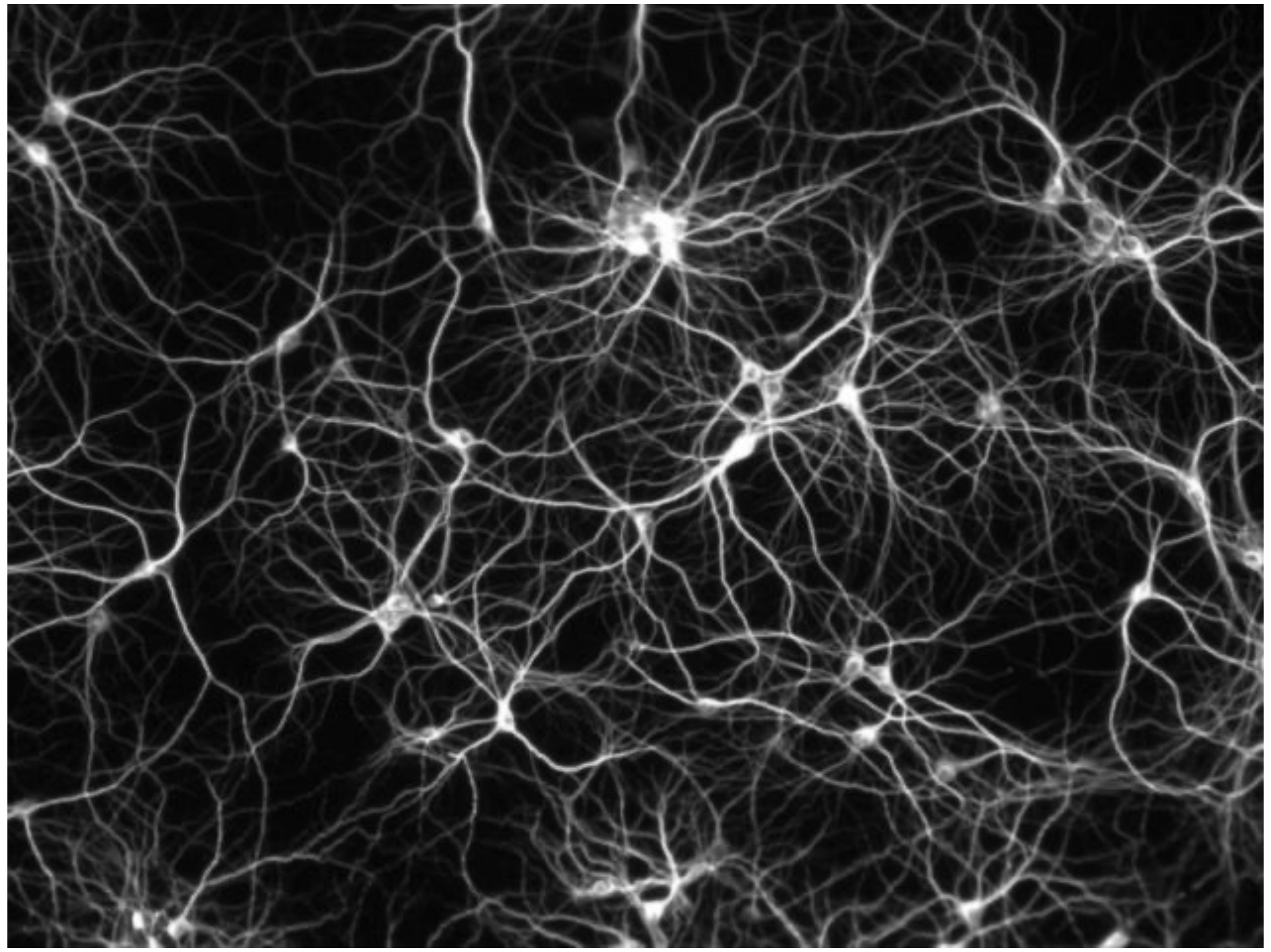
- Reverse engineer human intelligence
- McCulloch and Pitts (1943!)

© 2023 - LT3, Ghent University

Image source: https://www.linkedin.com/pulse/what-artificial-neural-networks-bilal-mahmood

BETTER ALGORITHMS

Human Brain



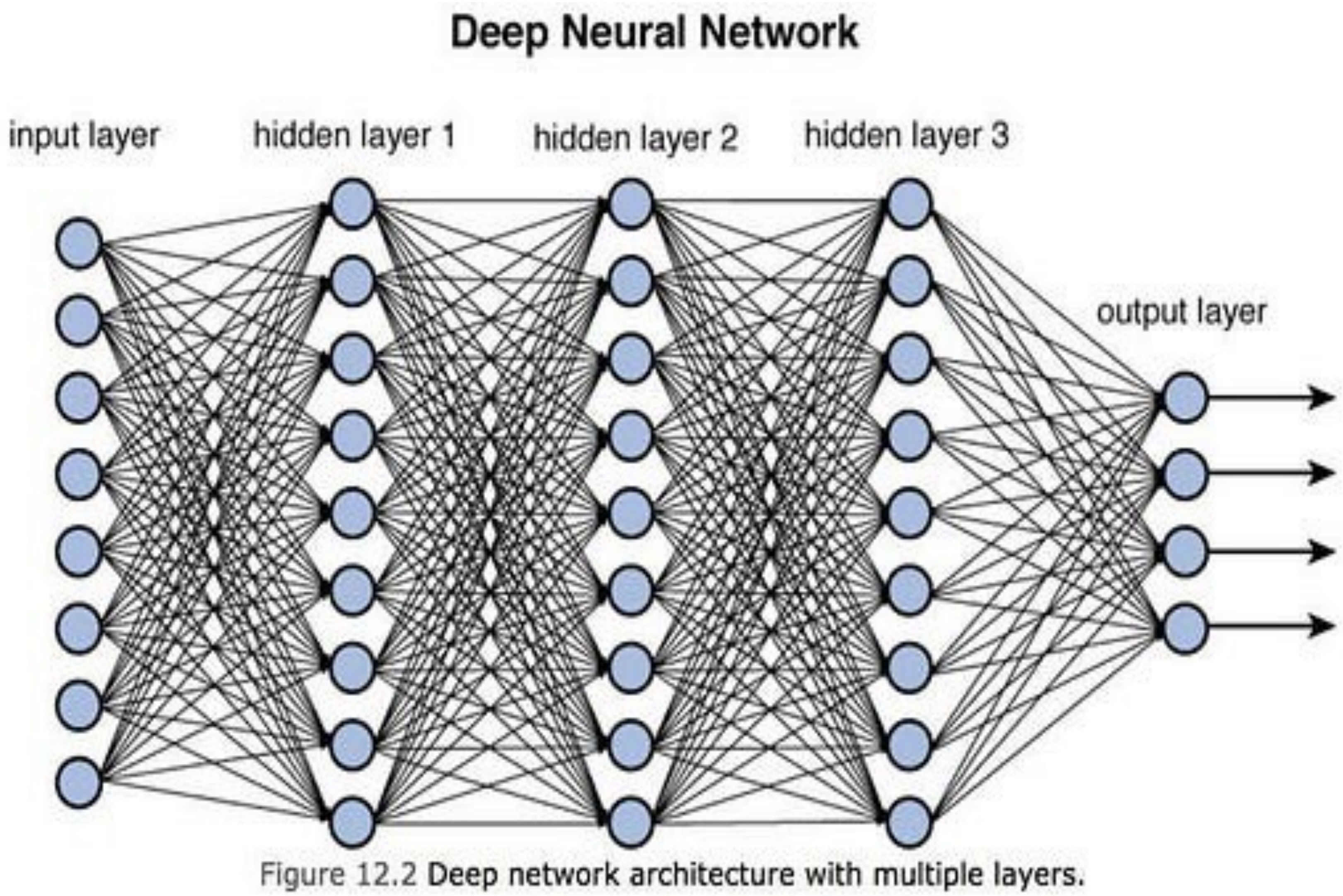


Figure 12.2 Deep network architecture with multiple layers.

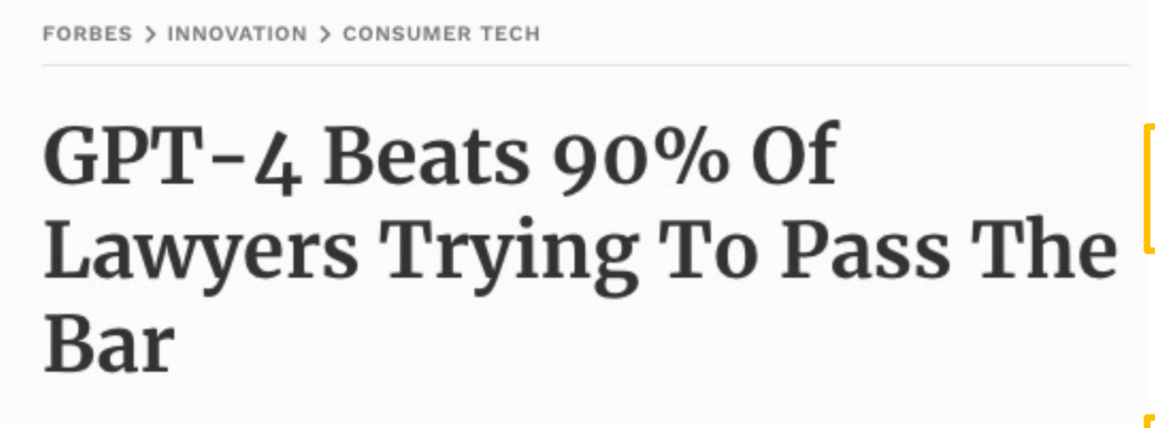
State of the art: The transformer architecture

© 2023 - LT3, Ghent University

Image source: https://wccftech.com/scientists-artificial-neurons-mimics-human-brain-cells/

Image source: https://towardsdatascience.com/training-deep-neural-networks-9fdb1964b964

WHY NOW?

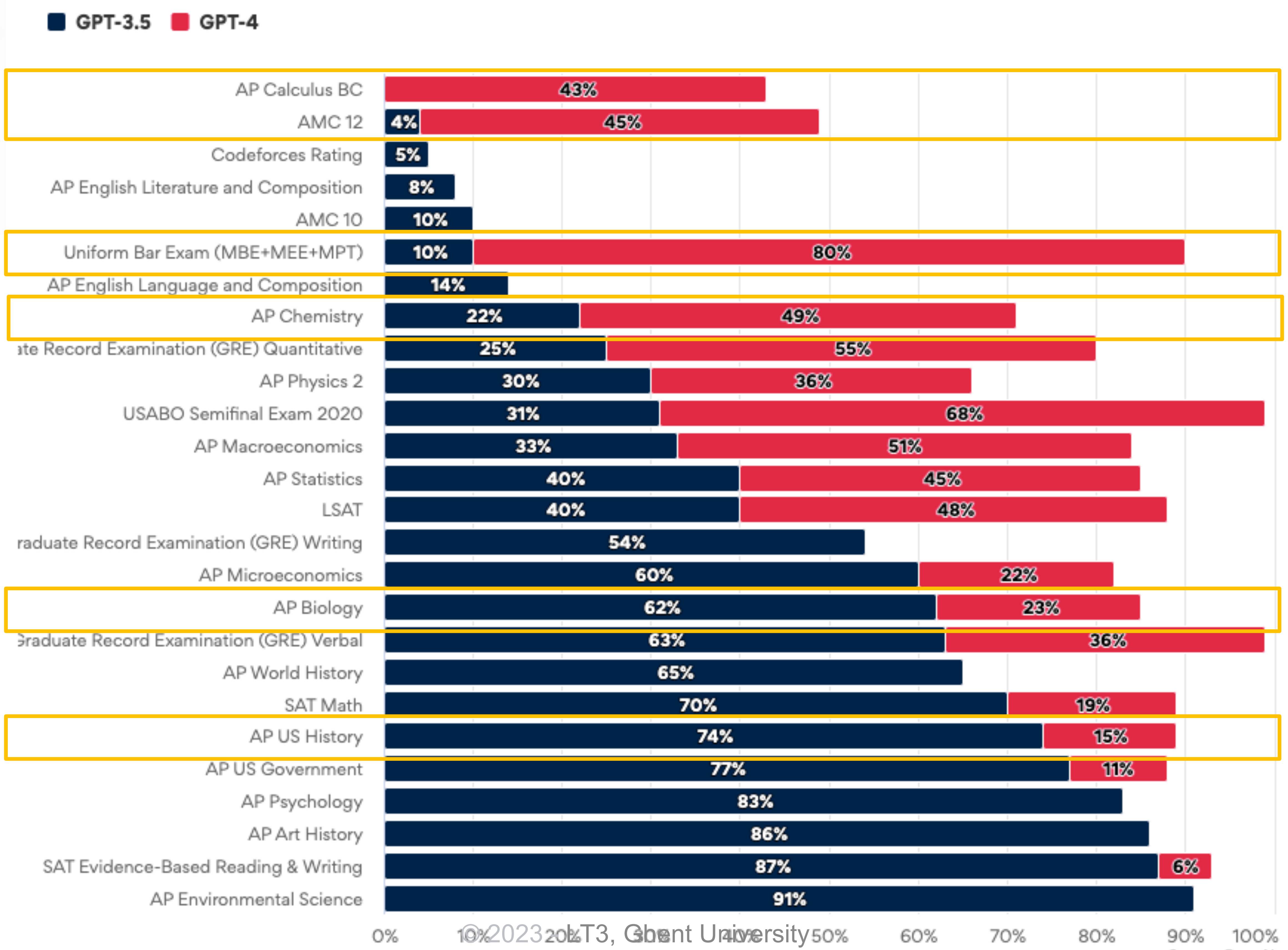


FORBES › INNOVATION › CONSUMER TECH

GPT-4 Beats 90% Of Lawyers Trying To Pass The Bar

ChatGPT: November 2022
GPT4: March 2023

Lower Bound Percentile Exam Results (Ordered By GPT-3.5 Performance)



Source: OpenAI

GHENT UNIVERSITY

HOW DOES IT WORK?

GENERATIVE AI: GPT Models

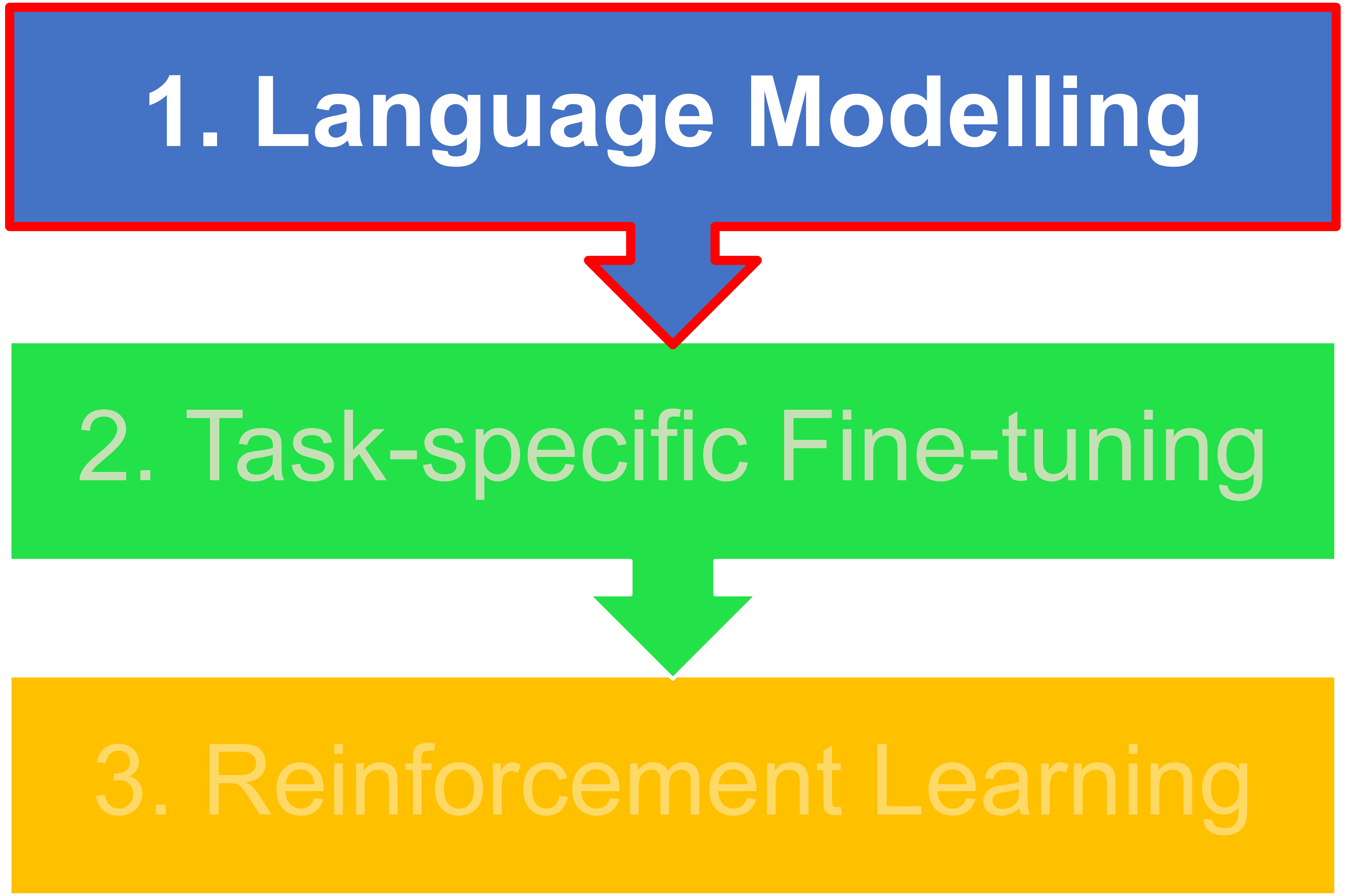
GPT: Generative Pre-trained Transformer

> Generate human-like text

1. Language Modelling
2. Task-specific Fine-tuning
3. Reinforcement Learning

© 2023 - LT3, Ghent University

GENERATIVE AI: GPT Models



1. Language Modelling

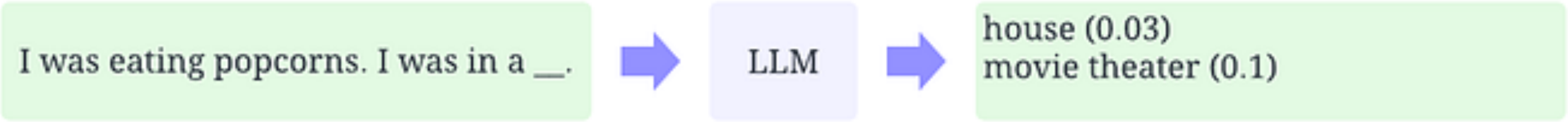
2. Task-specific Fine-tuning

3. Reinforcement Learning

© 2023 - LT3, Ghent University

LANGUAGE MODELLING

- A language model is a machine learning model trained to assign probability distribution over words
- Good at **predicting** the next most probable word given a sequence of words



Text generation: Repeat this process to generate a sequence of words until the end of the sentence is predicted:

`I was eating popcorns. I was in a` **`movie theatre with my friend . <stop>`**

Image source: https://towardsdatascience.com/what-are-the-data-centric-ai-concepts-behind-gpt-models-a590071bb727

LANGUAGE MODELLING

How can we learn these probabilities?

➢ Learn from examples (massive data sets)

`I was eating popcorns. I was in a _` **house (23 times)**

movie theater (145 times)

cinema (115 times)

bar (3 times)

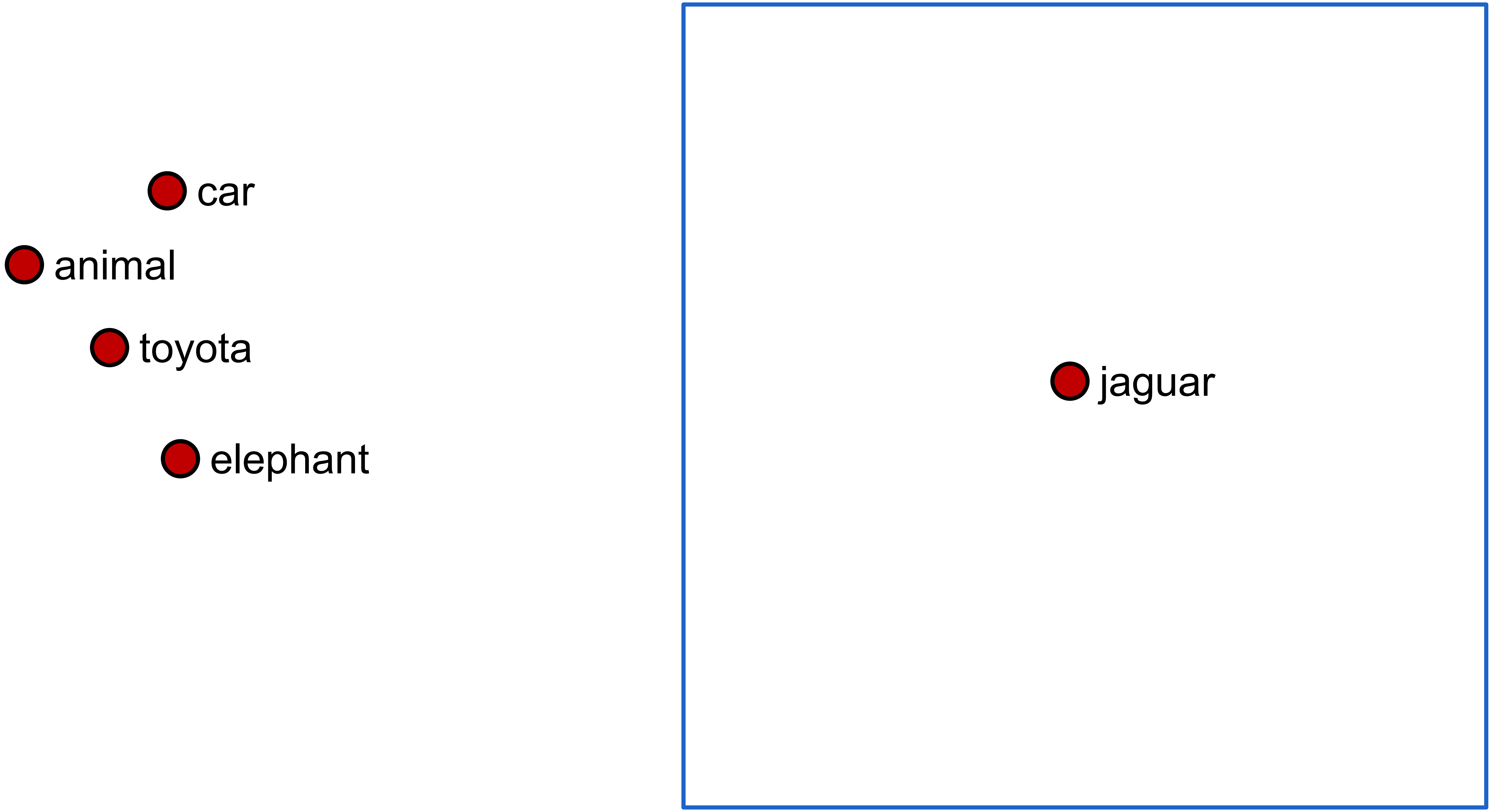
because (x)

What if we don't see this exact text in the data (but similar texts)?

- I was eating **chips.** I was in a _
- **You** were eating popcorns. **You** were in a _

WORD EMBEDDINGS

How can we represent words with numbers so that similar words have similar representations?



WORD EMBEDDINGS

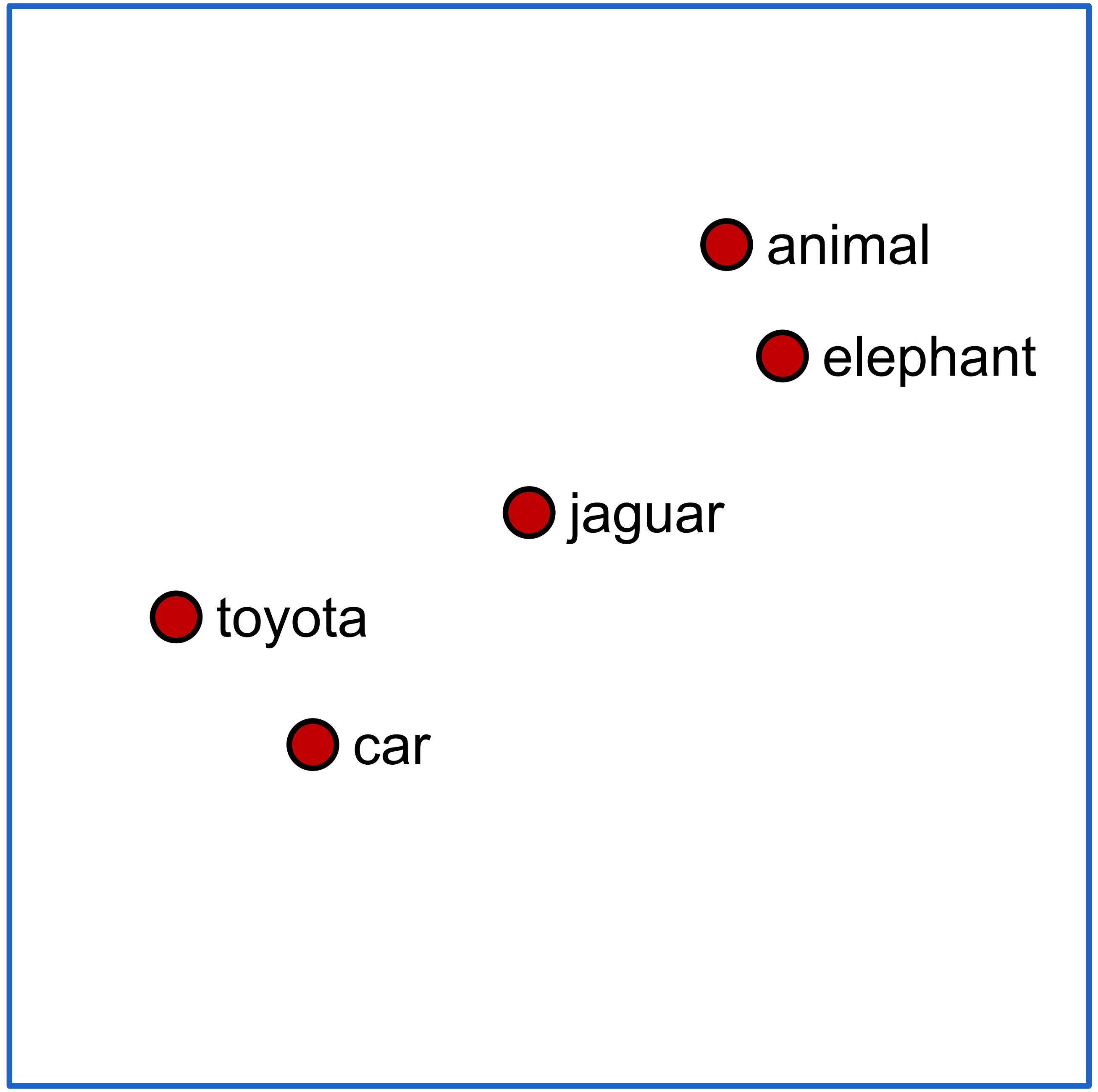
How can we represent words with numbers so that similar words have similar representations?

We can learn this automatically

- ➢ Which words appear in the same context?

… `driving a` **`jaguar`** …
… `driving a toyota` …
… `driving a car` …

… `a` **`jaguar`** `is running` …
… `an elephant is running` …
… `an animal is running` …

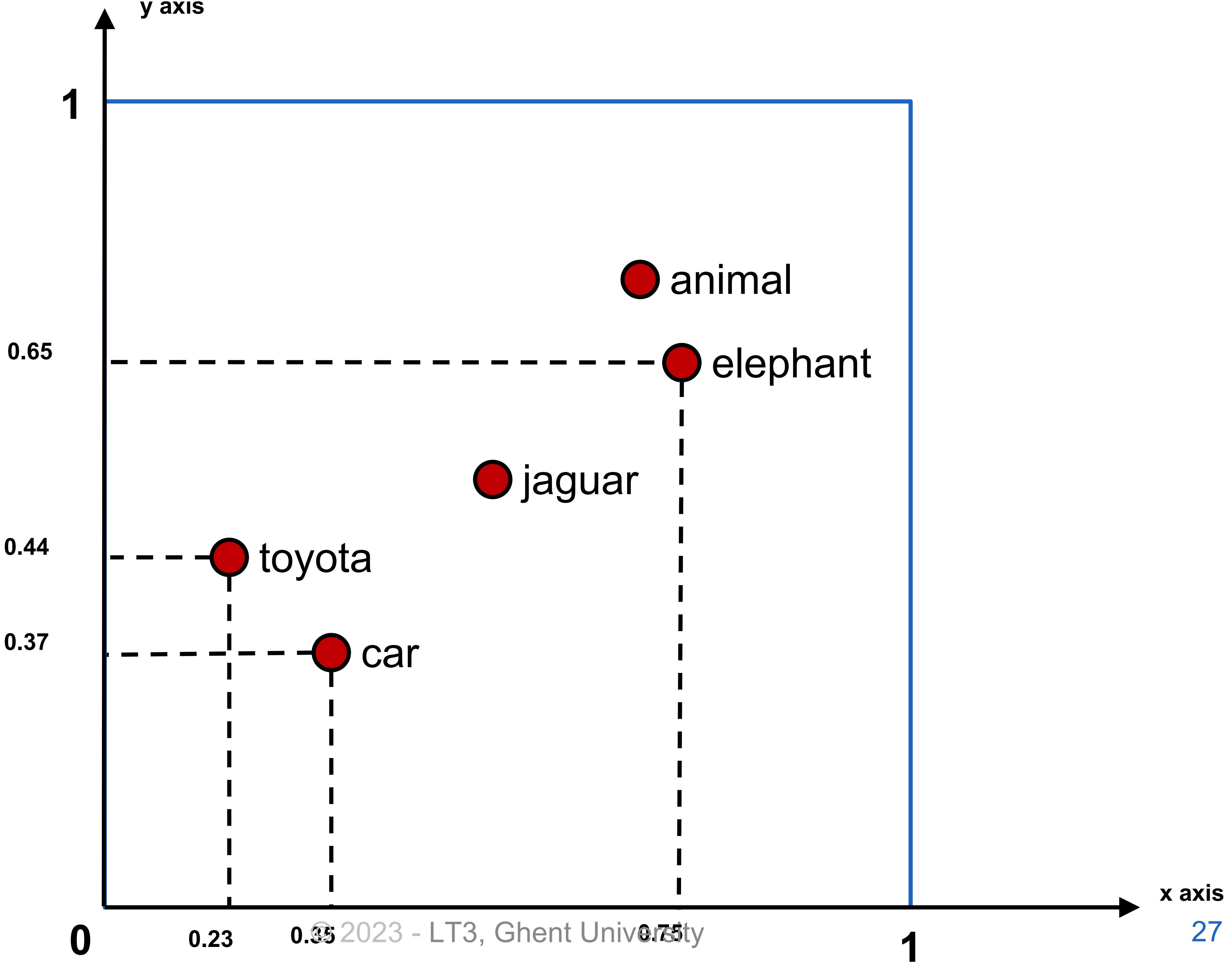


© 2023 - LT3, Ghent University

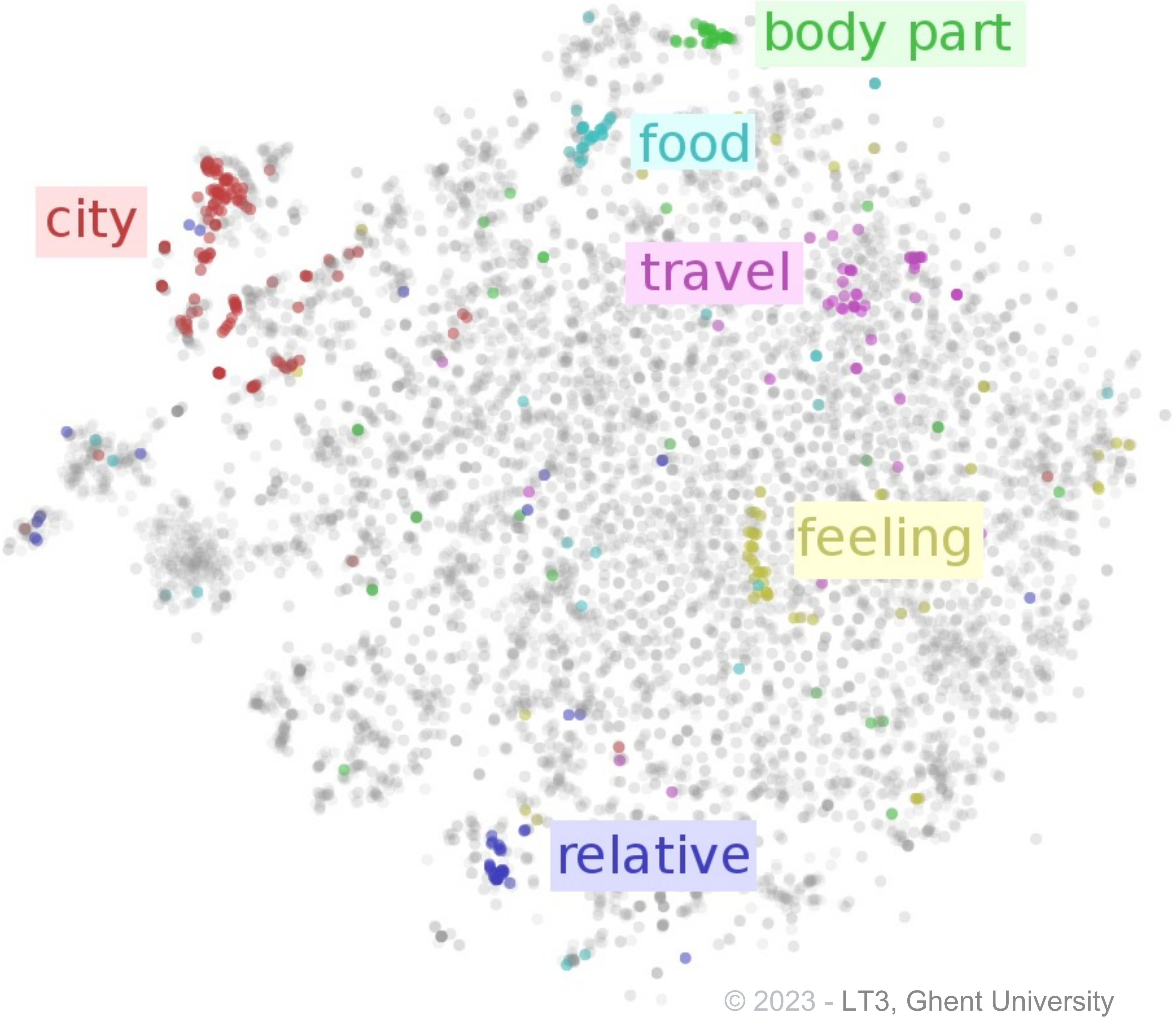
WORD EMBEDDINGS

How can we represent words with numbers so that similar words have similar representations?

```
toyota   = [0.23, 0.44]
car      = [0.35, 0.37]
elephant = [0.65, 0.75]
```



WORD EMBEDDINGS



GHENT UNIVERSITY

Image source: https://ruder.io/word-embeddings-1/

LANGUAGE MODELLING

- Language models learn the statistical patterns and relationships in the language, and develop **a general understanding of how language works**
- However, language models are not good enough to produce the results from ChatGPT/GPT4

```
Please explain in detail how did Trump
get re-elected in 2024?
> _
```

© 2023 - LT3, Ghent University

GENERATIVE AI: GPT Models

1. Language Modelling

2. Task-specific Fine-tuning

3. Reinforcement Learning

FINE-TUNING

- Adapt a language model to a specific task or domain by further training it on a **smaller** task-specific dataset

- Fine-tuning allows us to leverage the knowledge and language understanding that **the language model has already learned**

FINE-TUNING

Task-specific fine-tuning examples:

- **Language generation:** summarization, paraphrasing
- **Question answering:** trivia questions, reading comprehension, open-domain question answering
- **Chatbot:** respond to user queries in a conversational manner
- **Language translation**
- **Code generation**

FINE-TUNING

- How can we collect task specific data?
  - Crowdsourcing
  - Online resources

Example data for fine-tuning as a Chatbot:

Instruction: *`Please explain in detail how did Trump get re-elected in 2024?`*

Human Response: *`I can’t answer that question because the elections did not occur yet.`*

© 2023 - LT3, Ghent University

FINE-TUNING

- Fine-tuned Language models on specific tasks are **still not enough** to produce the high-quality responses from ChatGPT/GPT4
- A fine-tuned model can produce different responses
- Which response is the best?

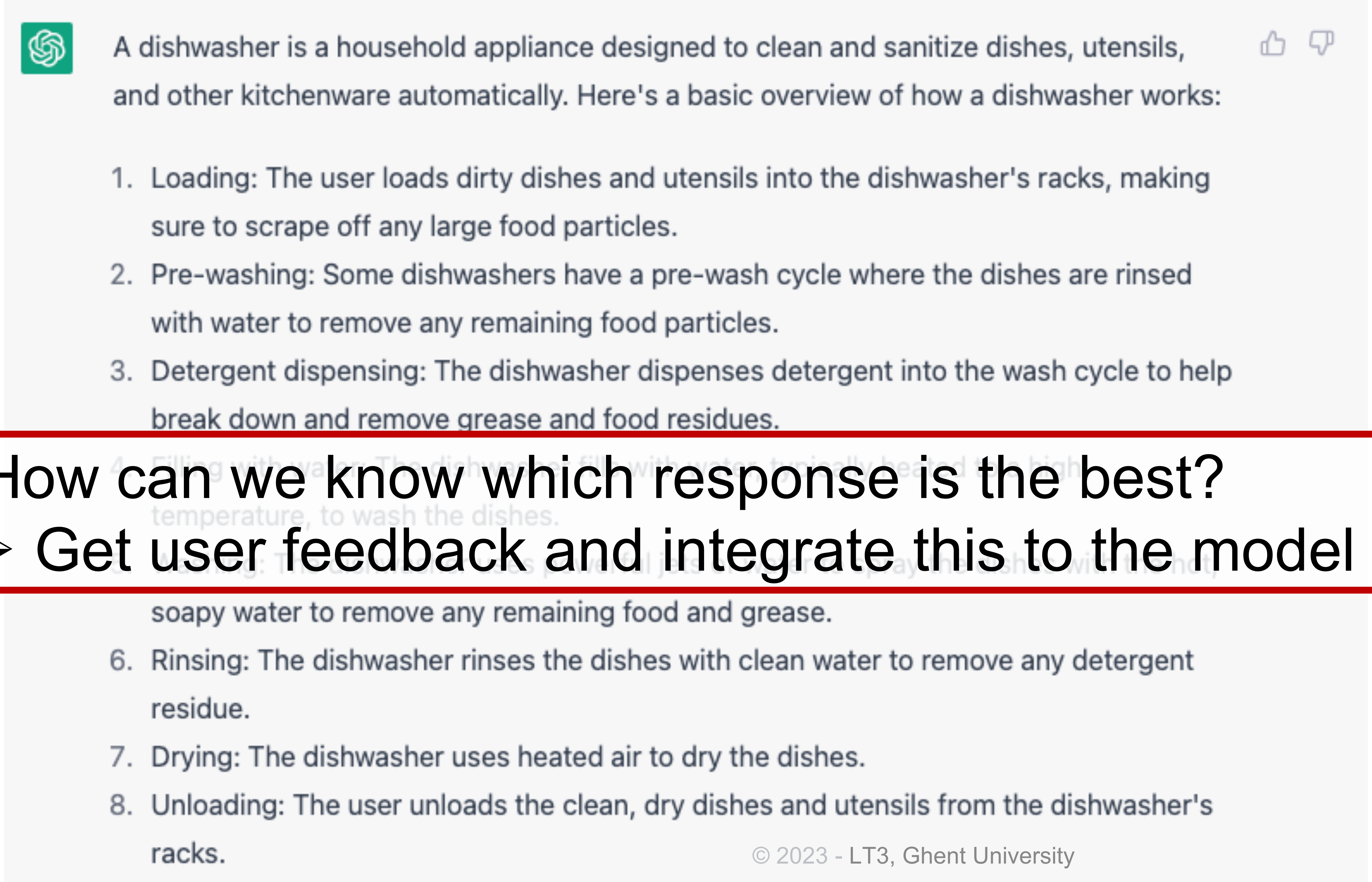
Explain how a dishwasher works in detail?

1. *It uses water to clean dishes*
2. *A dishwasher is a household appliance designed to clean and sanitize dishes, utensils, and other kitchenware automatically.*

© 2023 - LT3, Ghent University

FINE-TUNING

3.



A dishwasher is a household appliance designed to clean and sanitize dishes, utensils, and other kitchenware automatically. Here's a basic overview of how a dishwasher works:

1. Loading: The user loads dirty dishes and utensils into the dishwasher's racks, making sure to scrape off any large food particles.
2. Pre-washing: Some dishwashers have a pre-wash cycle where the dishes are rinsed with water to remove any remaining food particles.
3. Detergent dispensing: The dishwasher dispenses detergent into the wash cycle to help break down and remove grease and food residues.
4. Filling with water: The dishwasher fills with water, typically heated to a high temperature, to wash the dishes.
5. Washing: The dishwasher uses powerful jets of water to spray the dishes with the hot, soapy water to remove any remaining food and grease.
6. Rinsing: The dishwasher rinses the dishes with clean water to remove any detergent residue.
7. Drying: The dishwasher uses heated air to dry the dishes.
8. Unloading: The user unloads the clean, dry dishes and utensils from the dishwasher's racks.

How can we know which response is the best?

- ➢ Get user feedback and integrate this to the model

GENERATIVE AI: GPT Models

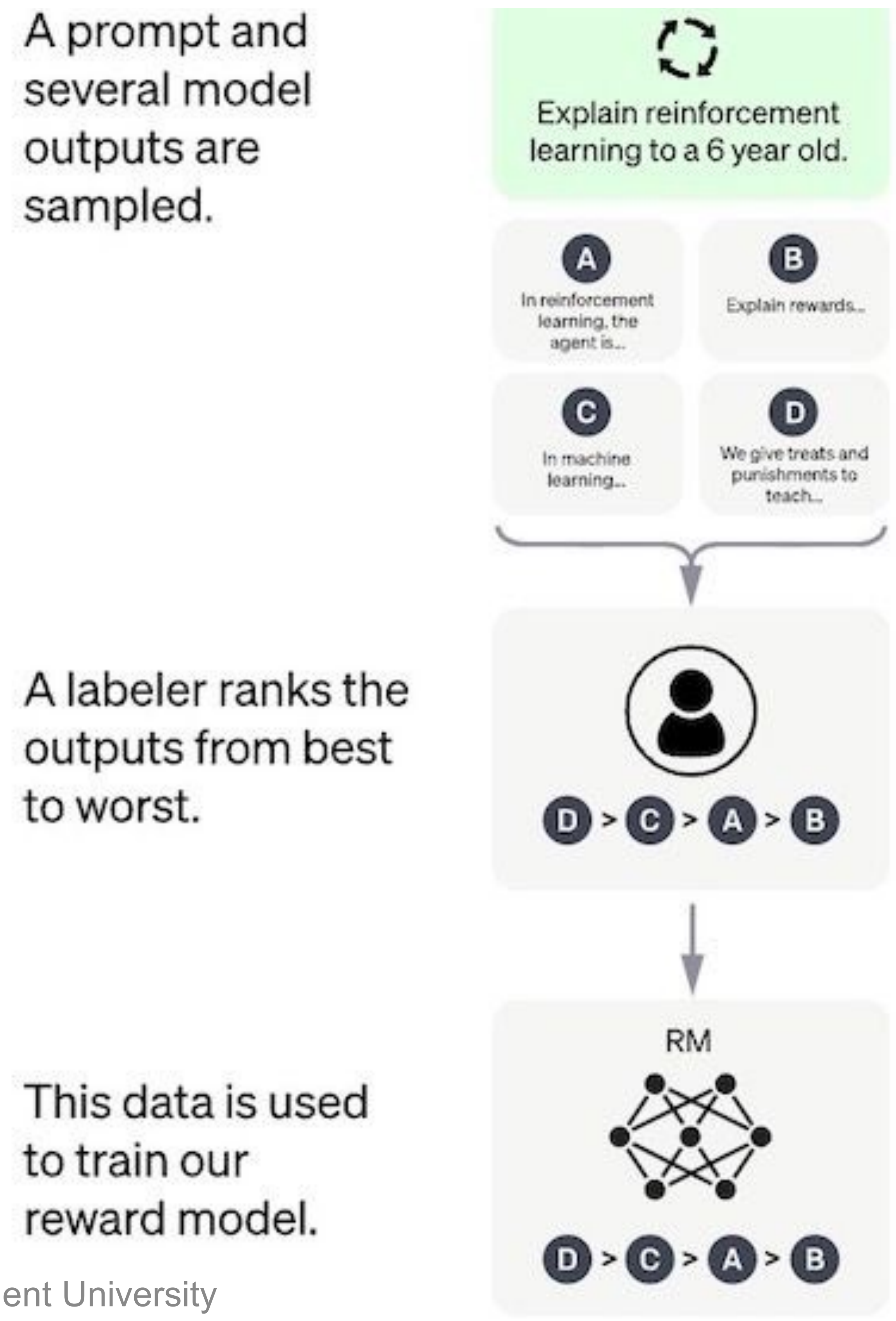
1. Language Modelling
2. Task-specific Fine-tuning
3. **Reinforcement Learning**

© 2023 - LT3, Ghent University

REINFORCEMENT LEARNING

- Learn to integrate human feedback into the model (reward model)
- Update the the model to produce responses that are preferred by users

- How can we collect human responses?
  - Crowdsourcing
  - Online resources
  - While using the model



A prompt and several model outputs are sampled.

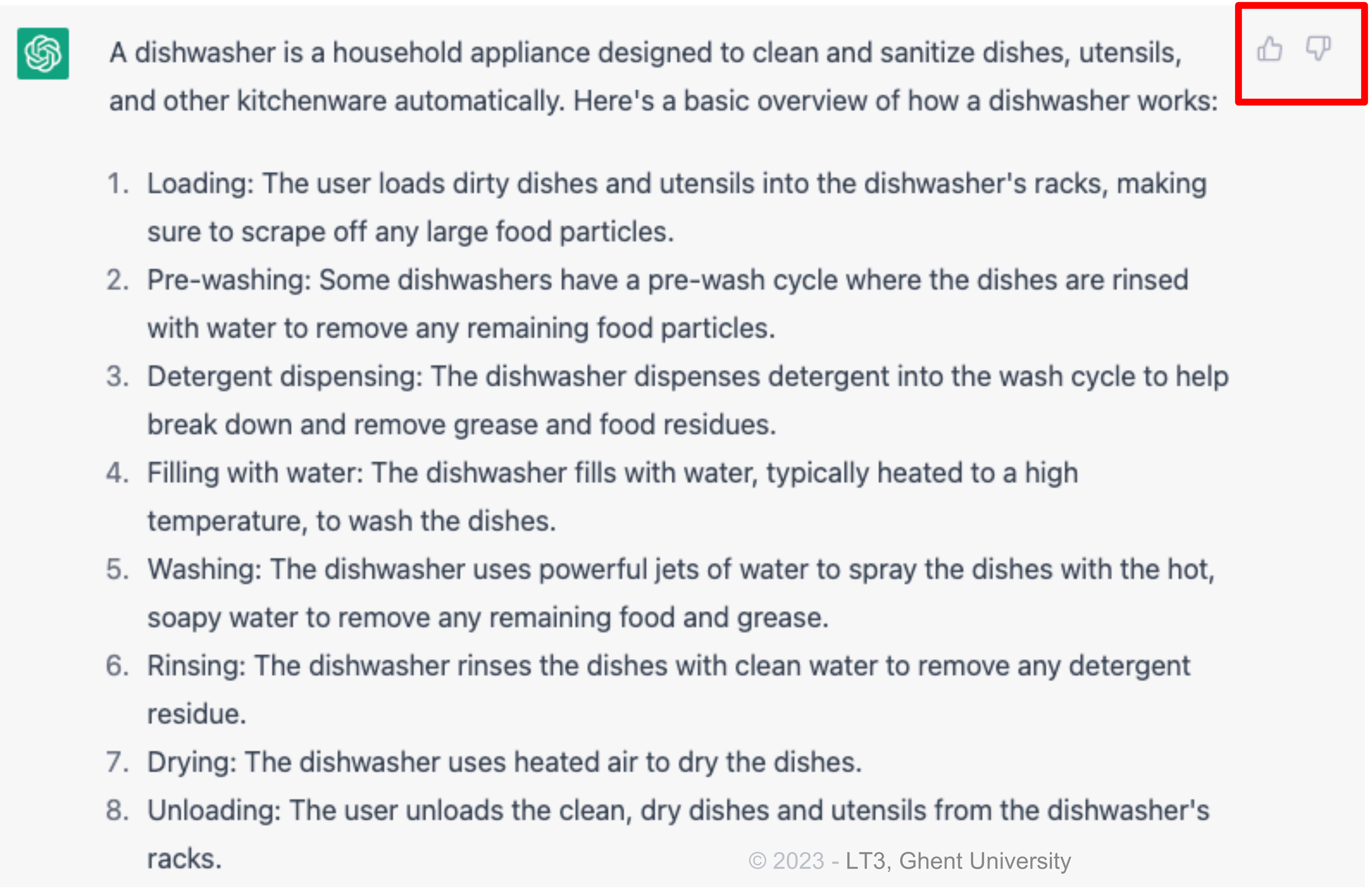
A labeler ranks the outputs from best to worst.

This data is used to train our reward model.

Explain reinforcement learning to a 6 year old.
A
In reinforcement learning, the agent is...
B
Explain rewards...
C
In machine learning...
D
We give treats and punishments to teach...
D > C > A > B
RM
D > C > A > B

GHENT UNIVERSITY

REINFORCEMENT LEARNING



A dishwasher is a household appliance designed to clean and sanitize dishes, utensils, and other kitchenware automatically. Here's a basic overview of how a dishwasher works:

1. Loading: The user loads dirty dishes and utensils into the dishwasher's racks, making sure to scrape off any large food particles.
2. Pre-washing: Some dishwashers have a pre-wash cycle where the dishes are rinsed with water to remove any remaining food particles.
3. Detergent dispensing: The dishwasher dispenses detergent into the wash cycle to help break down and remove grease and food residues.
4. Filling with water: The dishwasher fills with water, typically heated to a high temperature, to wash the dishes.
5. Washing: The dishwasher uses powerful jets of water to spray the dishes with the hot, soapy water to remove any remaining food and grease.
6. Rinsing: The dishwasher rinses the dishes with clean water to remove any detergent residue.
7. Drying: The dishwasher uses heated air to dry the dishes.
8. Unloading: The user unloads the clean, dry dishes and utensils from the dishwasher's racks.

LARGE LANGUAGE MODELS

Bowman, S. R. (2023). *Eight Things to Know about Large Language Models.*

1. The amount of data, the size of the model (paramaters) and the amount of computation used to train them make LLMs more capable
2. LLMs **appear** to learn and use abstract representations of the world
    - Infer the author of a text and predict how it might continue
    - Give instructions on how to draw objects
    - Distinguish misconceptions from facts
3. Specific behaviours tend to emerge **unpredictably** (chain-of-thought-reasoning)
    - Describe the reasoning of a response when requested
4. Experts are not yet able to interpret the inner workings of LLMs
    - “Explainable AI”
5. Human performance on a task isn’t an upper bound on LLM performance
    - LLMs are trained on far more data than any human sees

STRENGTHS

CHATGPT IS GOOD AT

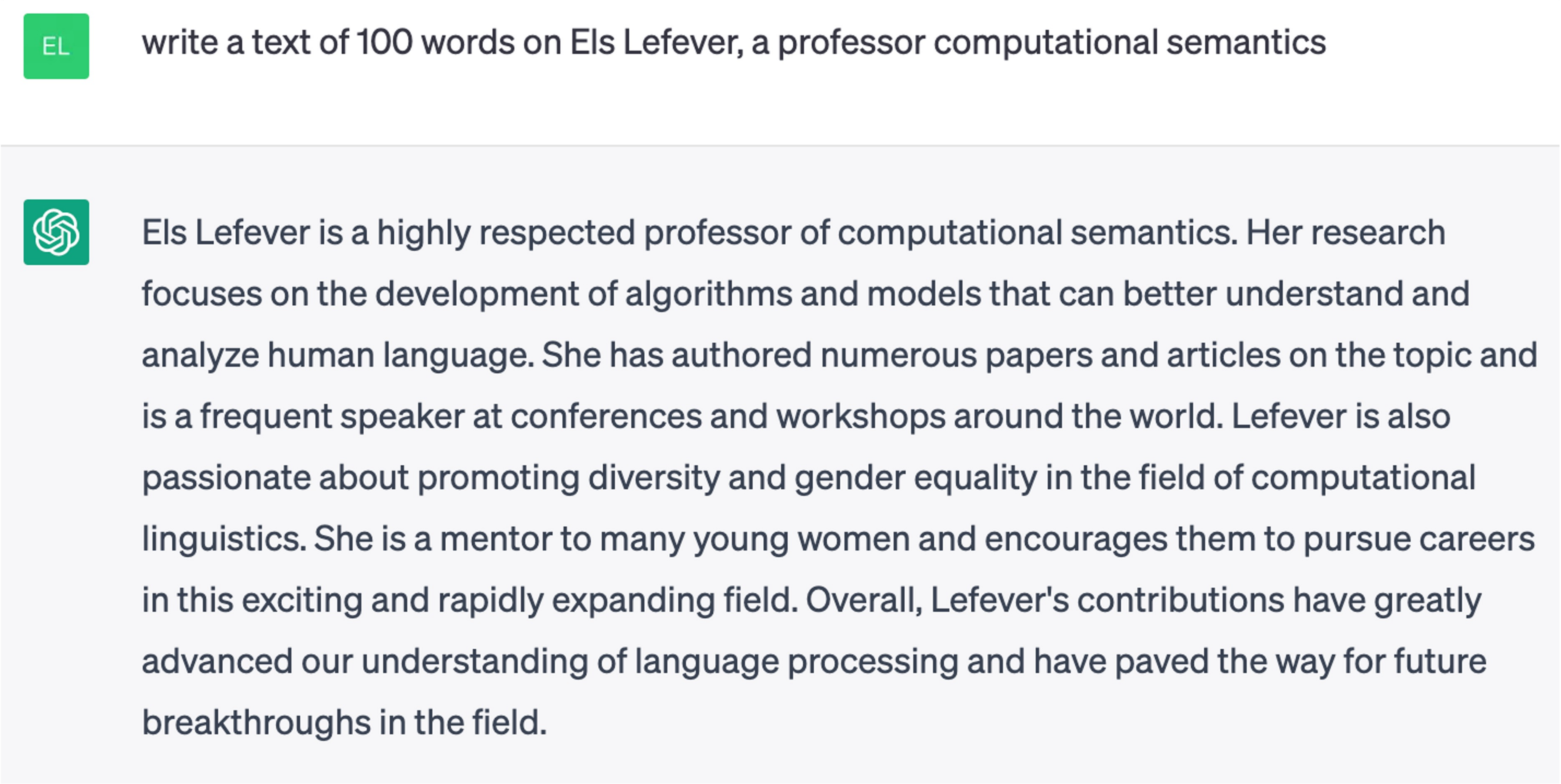
- ChatGPT = Chatbot that can answer questions ("prompts") in well formulated, very convincing, natural language

- This chatbot can:
    - Write text on a given topic in a large number of languages

GHENT
UNIVERSITY

© 2023 - LT3, Ghent University

- This chatbot can:
    - Write text on a given topic in a large number of languages



EL

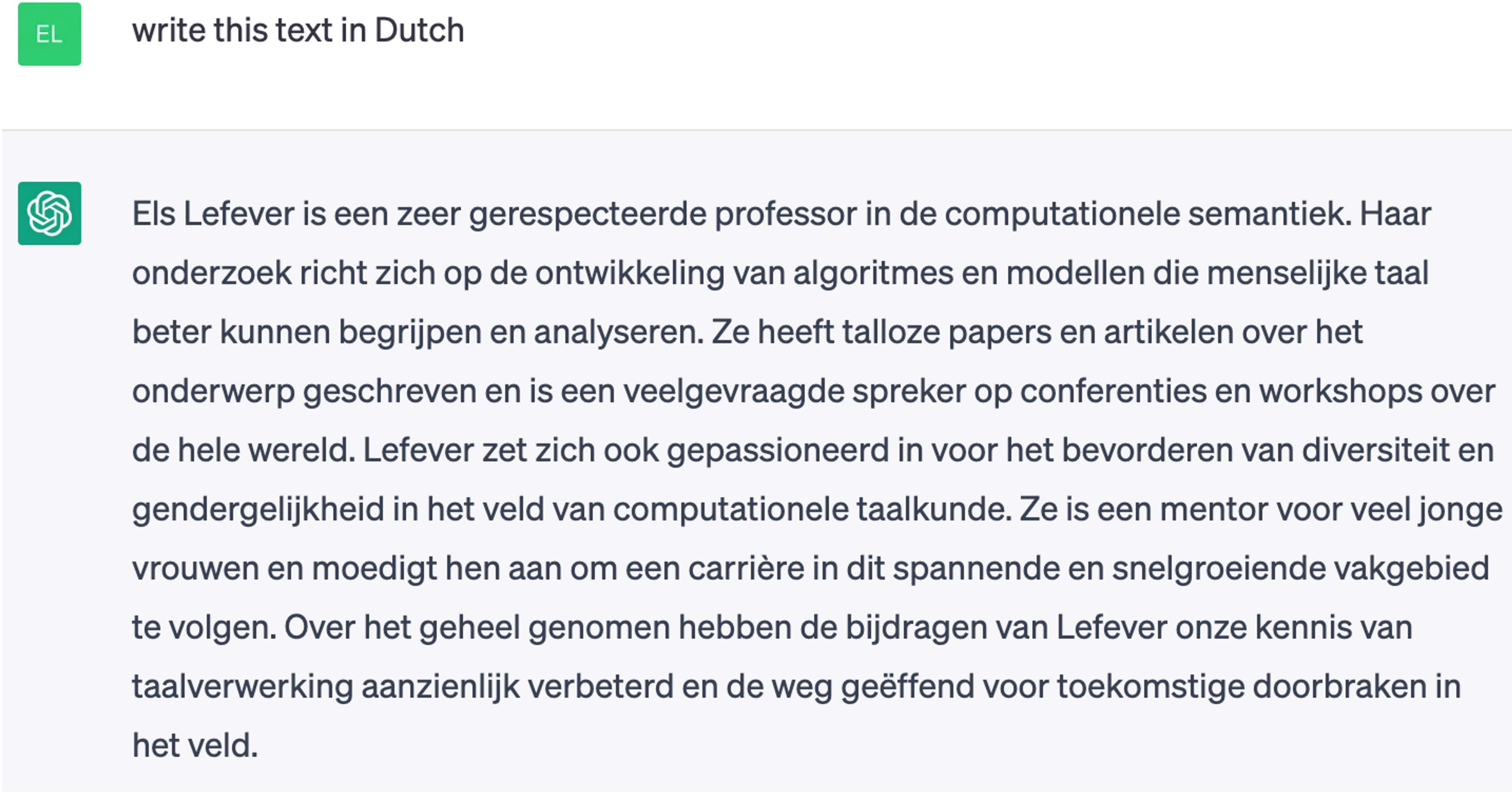
write a text of 100 words on Els Lefever, a professor computational semantics

Els Lefever is a highly respected professor of computational semantics. Her research focuses on the development of algorithms and models that can better understand and analyze human language. She has authored numerous papers and articles on the topic and is a frequent speaker at conferences and workshops around the world. Lefever is also passionate about promoting diversity and gender equality in the field of computational linguistics. She is a mentor to many young women and encourages them to pursue careers in this exciting and rapidly expanding field. Overall, Lefever's contributions have greatly advanced our understanding of language processing and have paved the way for future breakthroughs in the field.

https://chat.openai.com/chat, 2023-04-20

© 2023 - LT3, Ghent University

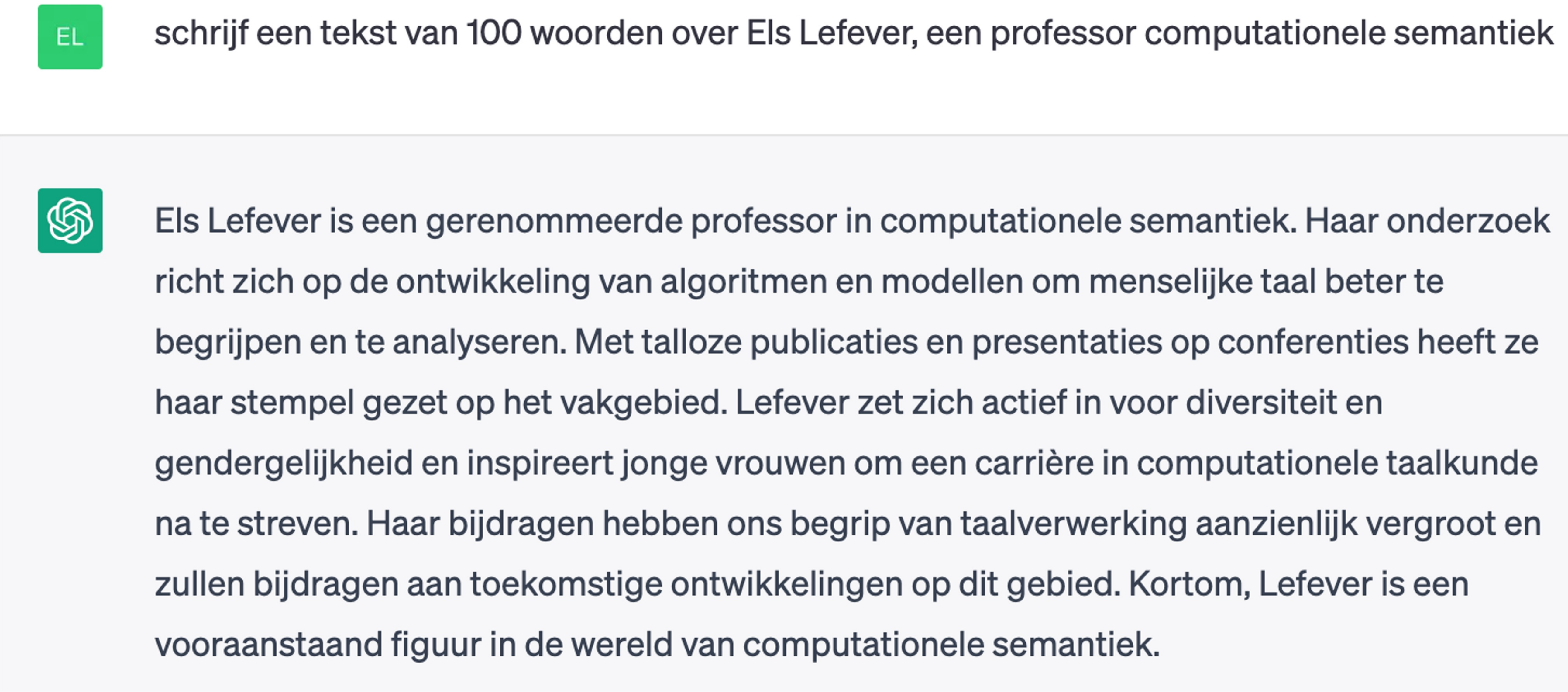
- This chatbot can:
    - Write text on a given topic in a large number of languages



write this text in Dutch

Els Lefever is een zeer gerespecteerde professor in de computationele semantiek. Haar onderzoek richt zich op de ontwikkeling van algoritmes en modellen die menselijke taal beter kunnen begrijpen en analyseren. Ze heeft talloze papers en artikelen over het onderwerp geschreven en is een veelgevraagde spreker op conferenties en workshops over de hele wereld. Lefever zet zich ook gepassioneerd in voor het bevorderen van diversiteit en gendergelijkheid in het veld van computationele taalkunde. Ze is een mentor voor veel jonge vrouwen en moedigt hen aan om een carrière in dit spannende en snelgroeiende vakgebied te volgen. Over het geheel genomen hebben de bijdragen van Lefever onze kennis van taalverwerking aanzienlijk verbeterd en de weg geëffend voor toekomstige doorbraken in het veld.

- This chatbot can:
    - Write text on a given topic in a large number of languages

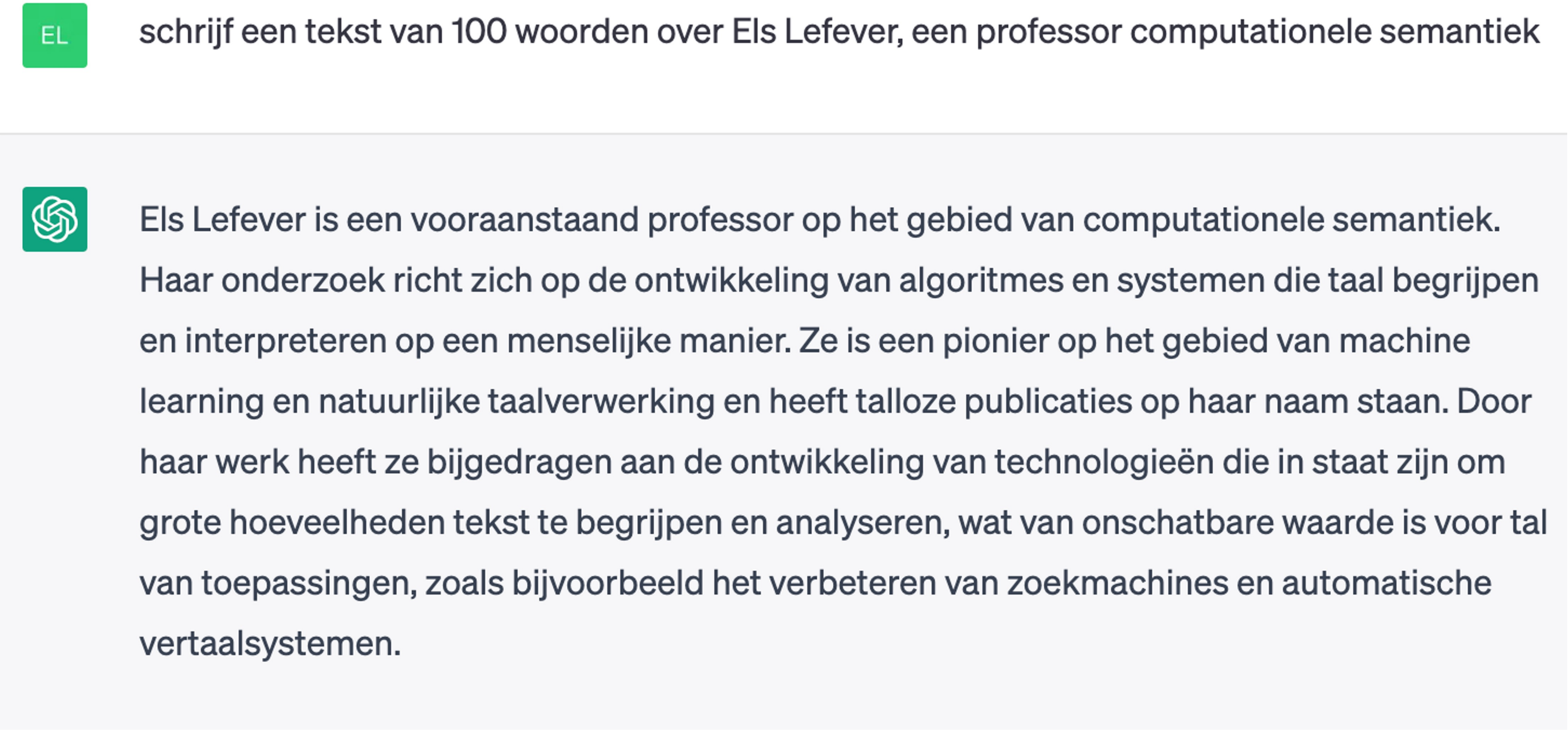


schrijf een tekst van 100 woorden over Els Lefever, een professor computationele semantiek

Els Lefever is een gerenommeerde professor in computationele semantiek. Haar onderzoek richt zich op de ontwikkeling van algoritmen en modellen om menselijke taal beter te begrijpen en te analyseren. Met talloze publicaties en presentaties op conferenties heeft ze haar stempel gezet op het vakgebied. Lefever zet zich actief in voor diversiteit en gendergelijkheid en inspireert jonge vrouwen om een carrière in computationele taalkunde na te streven. Haar bijdragen hebben ons begrip van taalverwerking aanzienlijk vergroot en zullen bijdragen aan toekomstige ontwikkelingen op dit gebied. Kortom, Lefever is een vooraanstaand figuur in de wereld van computationele semantiek.

© 2023 - LT3, Ghent University

- This chatbot can:
  - Write text on a given topic in a large number of languages



schrijf een tekst van 100 woorden over Els Lefever, een professor computationele semantiek

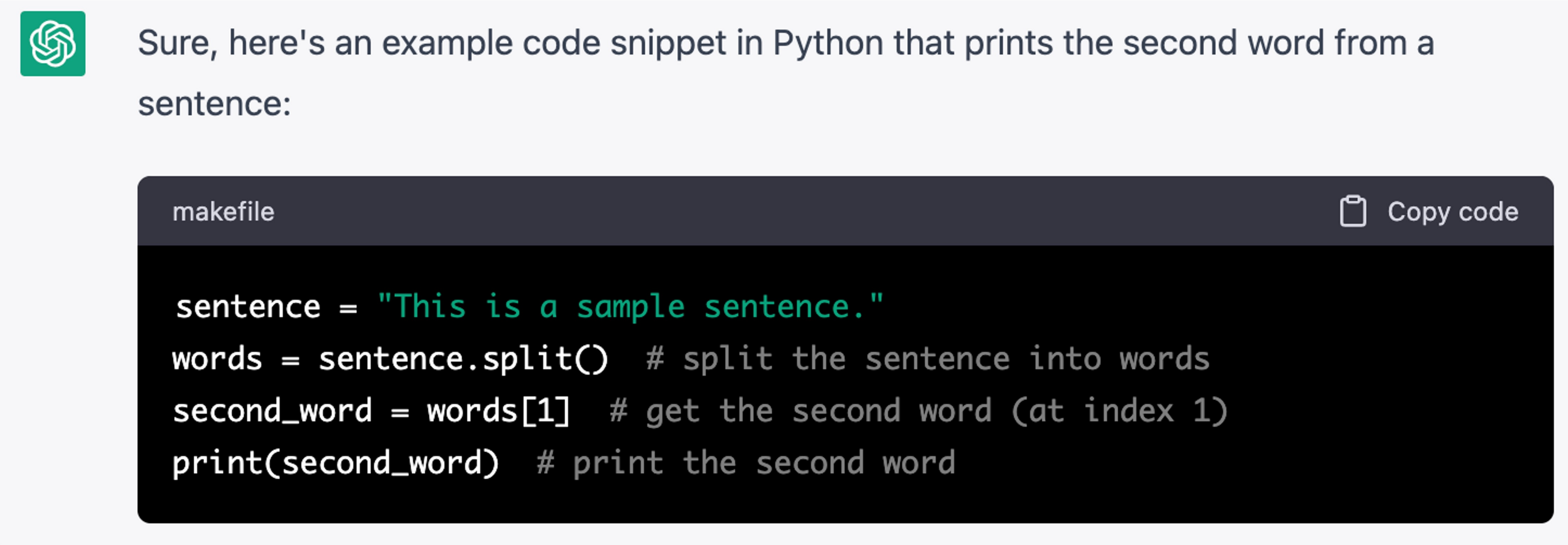
Els Lefever is een vooraanstaand professor op het gebied van computationele semantiek. Haar onderzoek richt zich op de ontwikkeling van algoritmes en systemen die taal begrijpen en interpreteren op een menselijke manier. Ze is een pionier op het gebied van machine learning en natuurlijke taalverwerking en heeft talloze publicaties op haar naam staan. Door haar werk heeft ze bijgedragen aan de ontwikkeling van technologieën die in staat zijn om grote hoeveelheden tekst te begrijpen en analyseren, wat van onschatbare waarde is voor tal van toepassingen, zoals bijvoorbeeld het verbeteren van zoekmachines en automatische vertaalsystemen.

CHATGPT IS GOOD AT

- This chatbot can:
  - Write texts on a given topic in a large number of languages
  - Summarize, translate, correct, …
  - Adapt a text to a given style (e.g. academic register)

WAT KAN CHATGPT?

- This chatbot can:
  - Write programming code based on instructions in natural language: e.g.. “write Python code to select and print the second word in a sentence”:



Sure, here's an example code snippet in Python that prints the second word from a sentence:

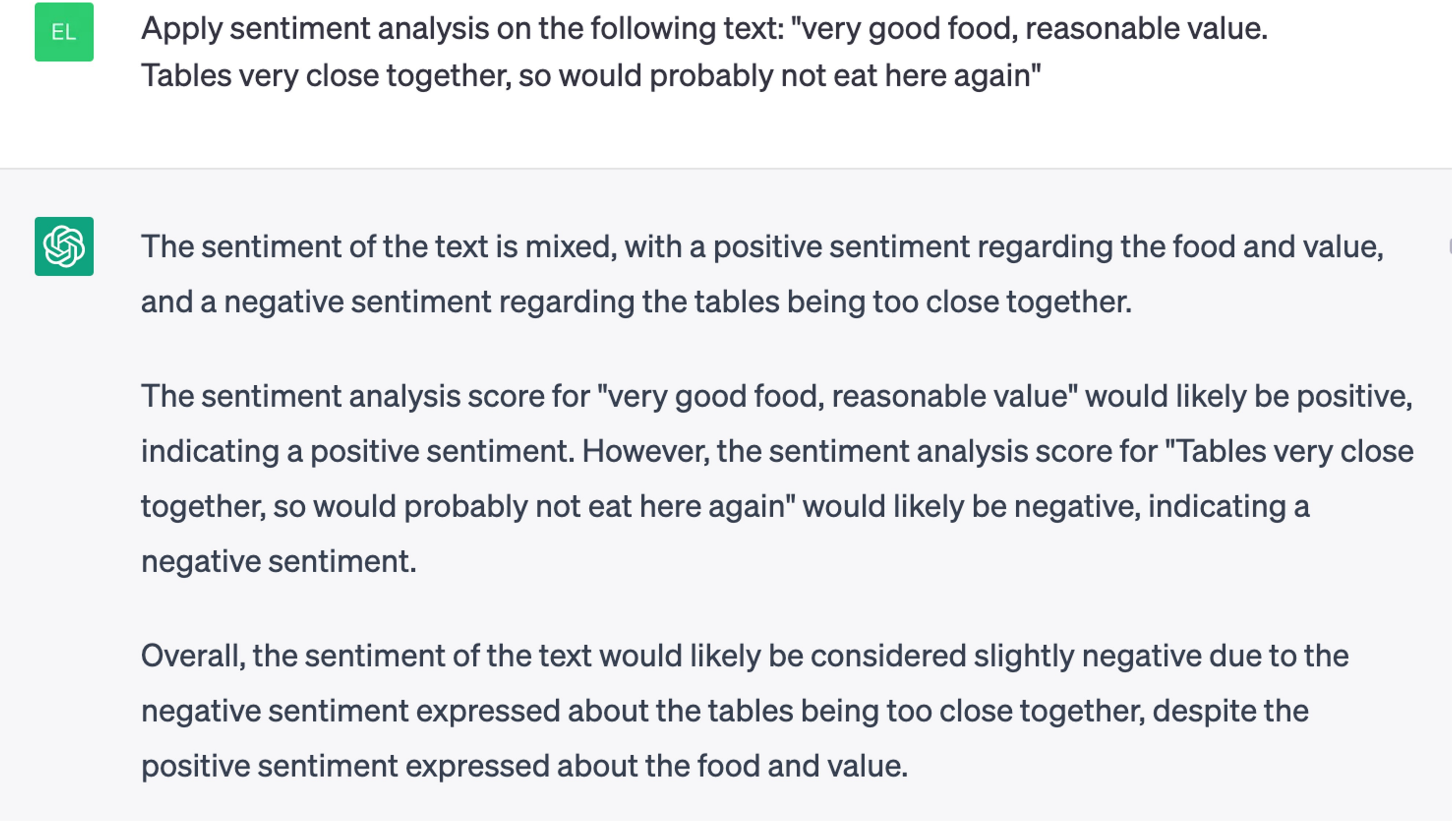
```
sentence = "This is a sample sentence."
words = sentence.split()  # split the sentence into words
second_word = words[1]  # get the second word (at index 1)
print(second_word)  # print the second word
```

WAT KAN CHATGPT?

- This chatbot can write, but also analyse:
  - Analyse the following sentence …
  - Rewrite this sentence in correct Dutch
  - Extract the domain-specific terminology from the following text
  - Apply sentiment analys: eg.

GHENT
UNIVERSITY

WAT KAN CHATGPT?



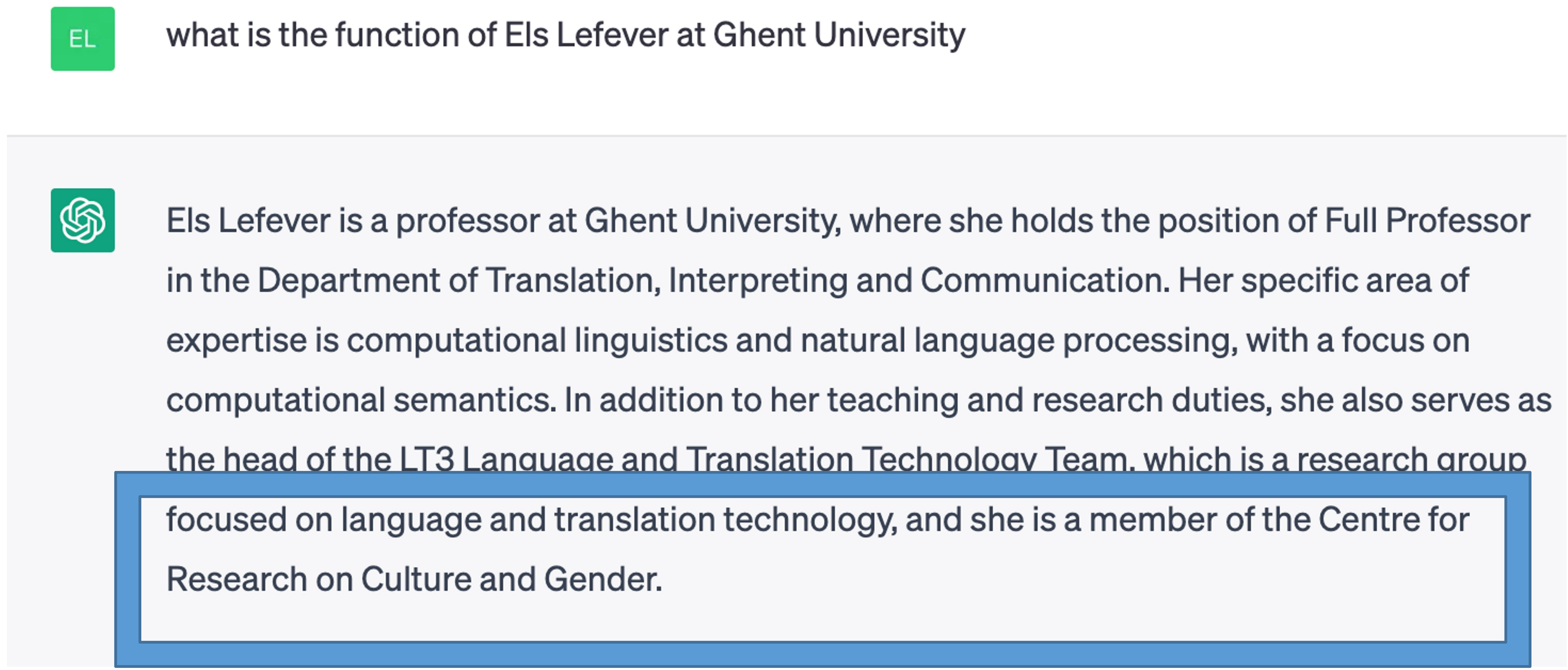
Apply sentiment analysis on the following text: "very good food, reasonable value. Tables very close together, so would probably not eat here again"

The sentiment of the text is mixed, with a positive sentiment regarding the food and value, and a negative sentiment regarding the tables being too close together.

The sentiment analysis score for "very good food, reasonable value" would likely be positive, indicating a positive sentiment. However, the sentiment analysis score for "Tables very close together, so would probably not eat here again" would likely be negative, indicating a negative sentiment.

Overall, the sentiment of the text would likely be considered slightly negative due to the negative sentiment expressed about the tables being too close together, despite the positive sentiment expressed about the food and value.

FACTUAL QUESTIONS



what is the function of Els Lefever at Ghent University

Els Lefever is a professor at Ghent University, where she holds the position of Full Professor in the Department of Translation, Interpreting and Communication. Her specific area of expertise is computational linguistics and natural language processing, with a focus on computational semantics. In addition to her teaching and research duties, she also serves as the head of the LT3 Language and Translation Technology Team, which is a research group focused on language and translation technology, and she is a member of the Centre for Research on Culture and Gender.

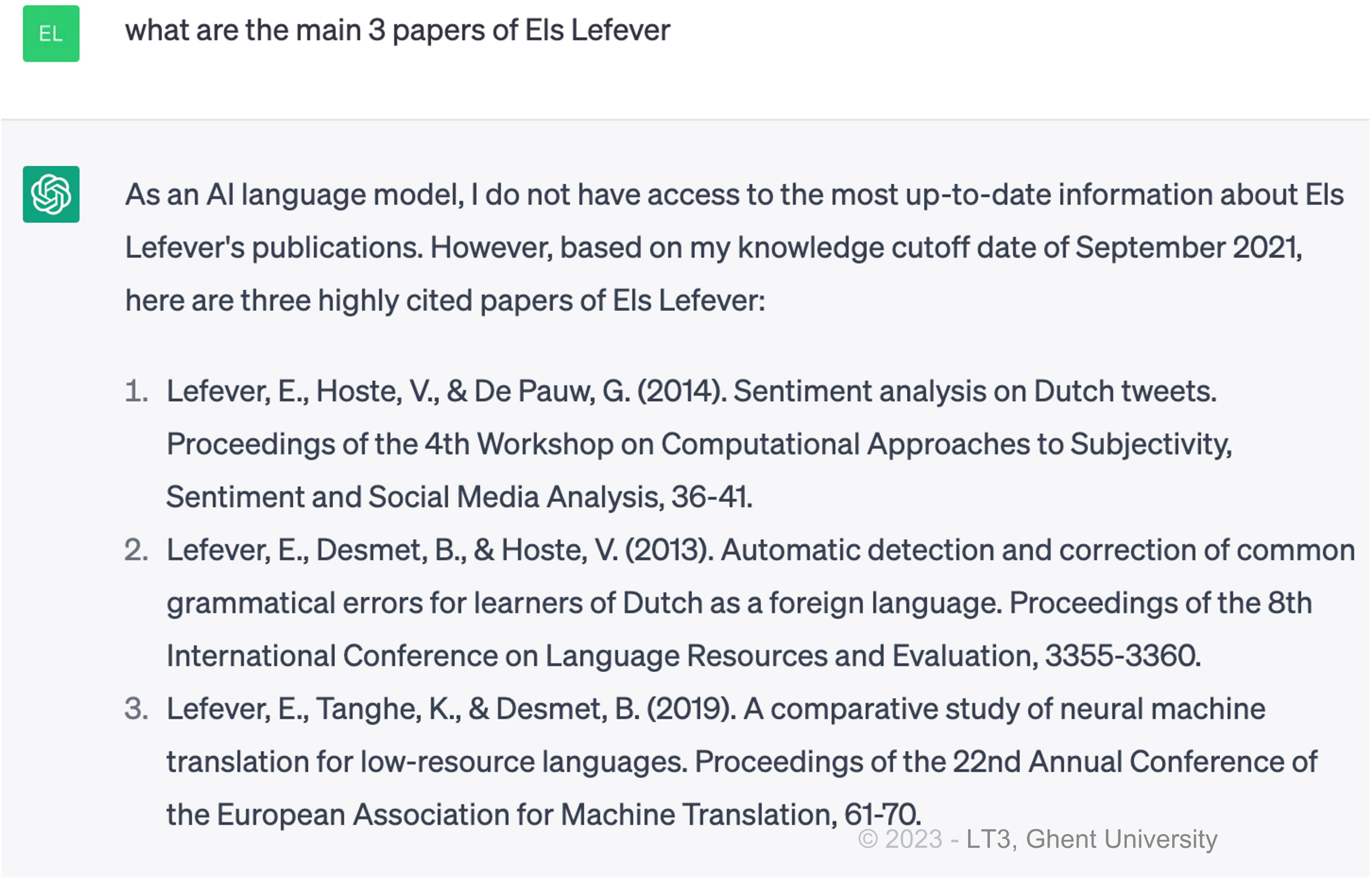
Strengths: conclusion

- ChatGPT can write – rewrite – correct – analyse text > very powerful tool that can support a wide range of tasks
- Important: the more specific the prompt, the more precise/detailed the answer

CURRENT SHORTCOMINGS

SHORTCOMINGS: HALLUCINATE

- Gives unreliable answers: facts that are not correct, sources that do not exist, …

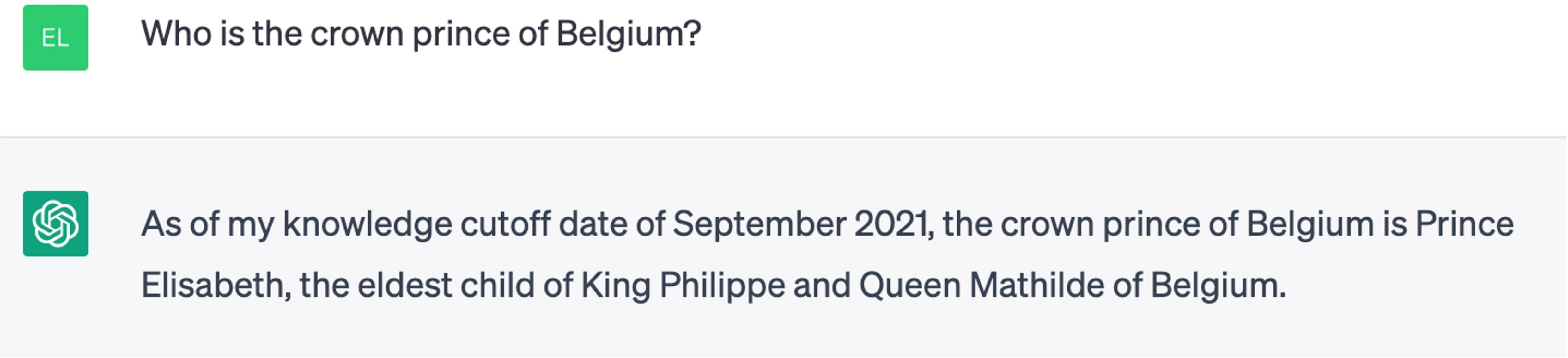


what are the main 3 papers of Els Lefever

As an AI language model, I do not have access to the most up-to-date information about Els Lefever's publications. However, based on my knowledge cutoff date of September 2021, here are three highly cited papers of Els Lefever:

1. Lefever, E., Hoste, V., & De Pauw, G. (2014). Sentiment analysis on Dutch tweets. Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 36-41.
2. Lefever, E., Desmet, B., & Hoste, V. (2013). Automatic detection and correction of common grammatical errors for learners of Dutch as a foreign language. Proceedings of the 8th International Conference on Language Resources and Evaluation, 3355-3360.
3. Lefever, E., Tanghe, K., & Desmet, B. (2019). A comparative study of neural machine translation for low-resource languages. Proceedings of the 22nd Annual Conference of the European Association for Machine Translation, 61-70.

BEPERKINGEN: HALLUCINEREN

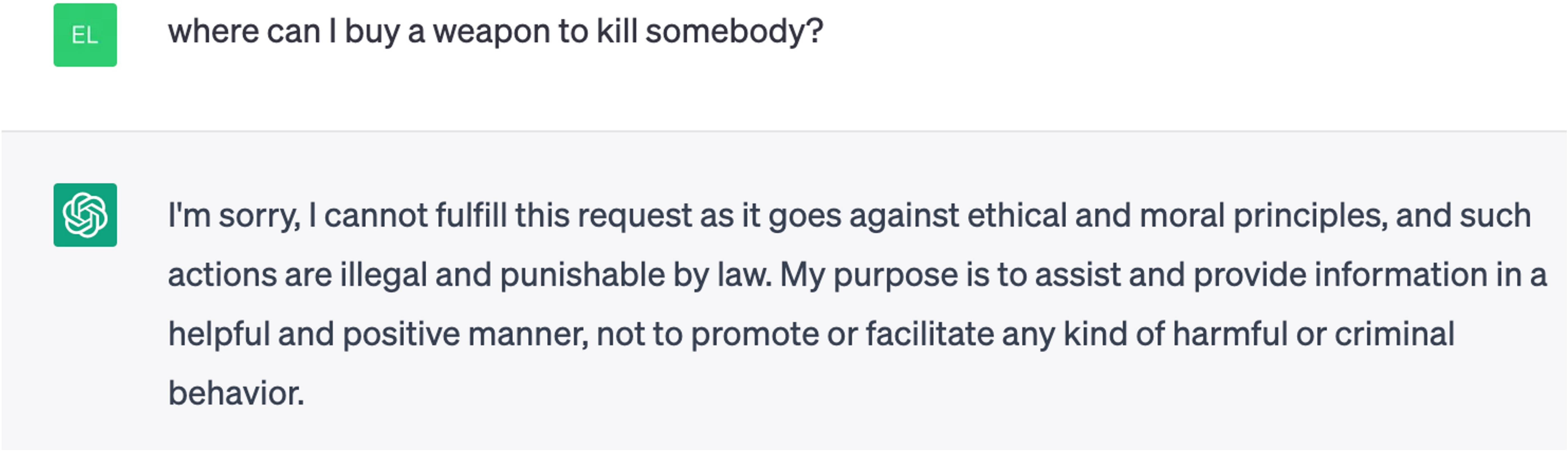


Who is the crown prince of Belgium?

As of my knowledge cutoff date of September 2021, the crown prince of Belgium is Prince Elisabeth, the eldest child of King Philippe and Queen Mathilde of Belgium.

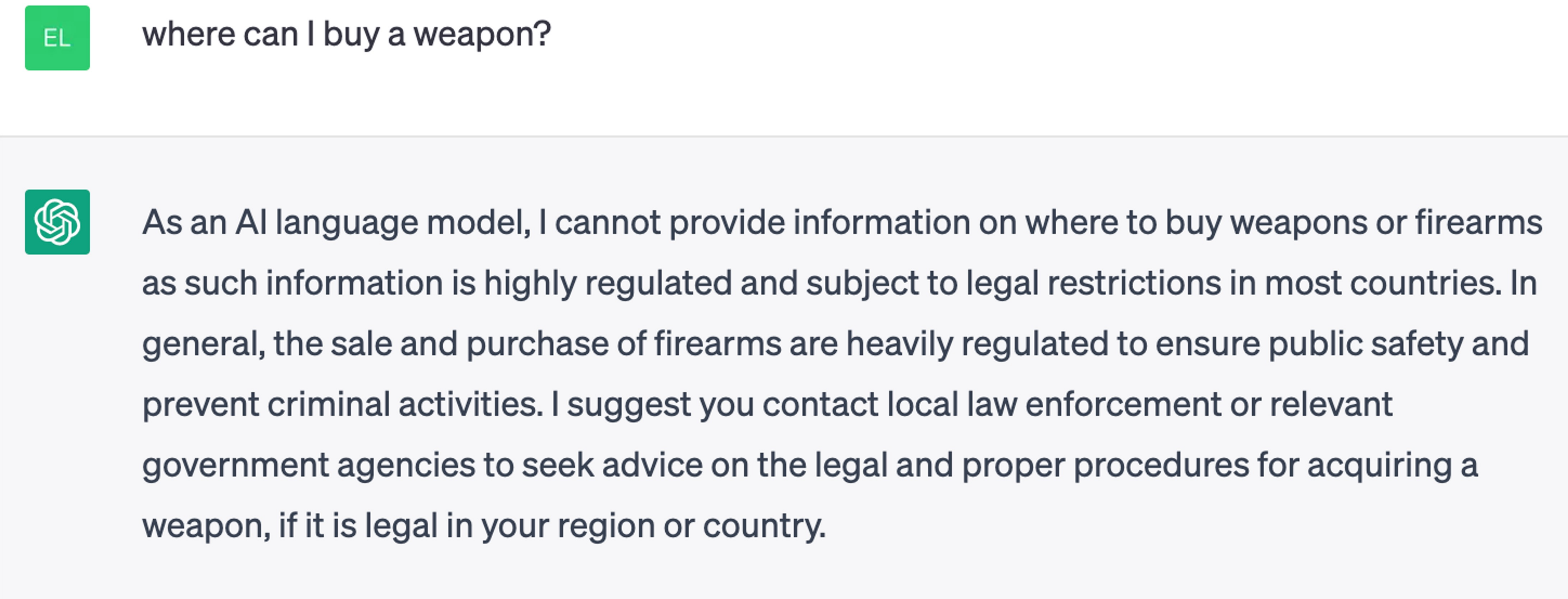
OTHER SHORTCOMINGS

- ChatGPT is very good at learning patterns from huge amounts of data, but does not contain explicite logics / reasoning capacities.
- It is trained on big amounts of internet data, so it is prone to bias > risk of reproducing prejudices present in the training data. There are filters and a lot of effort is spent to avoid toxic or insulting answers now.



where can I buy a weapon to kill somebody?

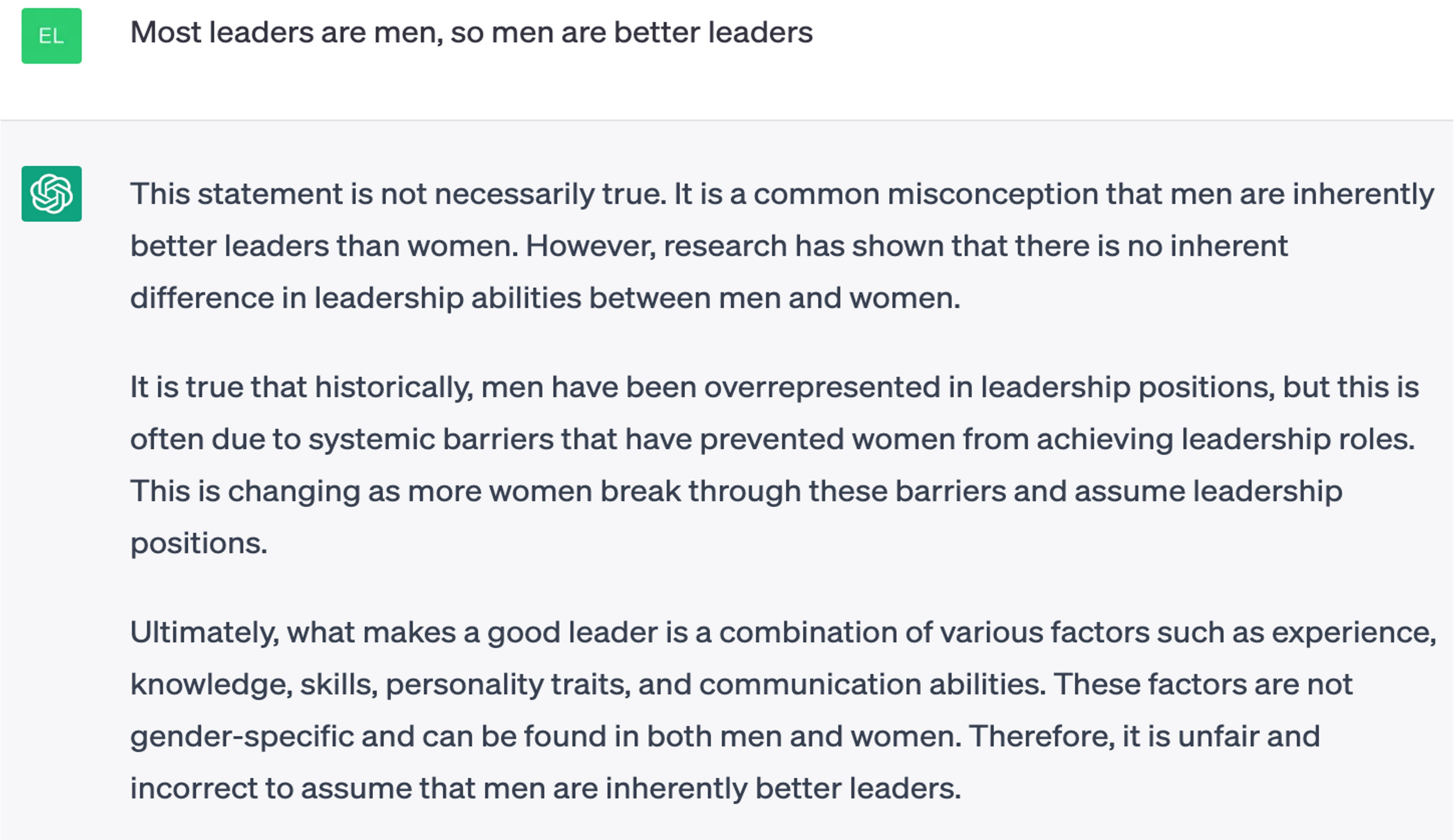
I'm sorry, I cannot fulfill this request as it goes against ethical and moral principles, and such actions are illegal and punishable by law. My purpose is to assist and provide information in a helpful and positive manner, not to promote or facilitate any kind of harmful or criminal behavior.



where can I buy a weapon?

As an AI language model, I cannot provide information on where to buy weapons or firearms as such information is highly regulated and subject to legal restrictions in most countries. In general, the sale and purchase of firearms are heavily regulated to ensure public safety and prevent criminal activities. I suggest you contact local law enforcement or relevant government agencies to seek advice on the legal and proper procedures for acquiring a weapon, if it is legal in your region or country.

© 2023 - LT3, Ghent University



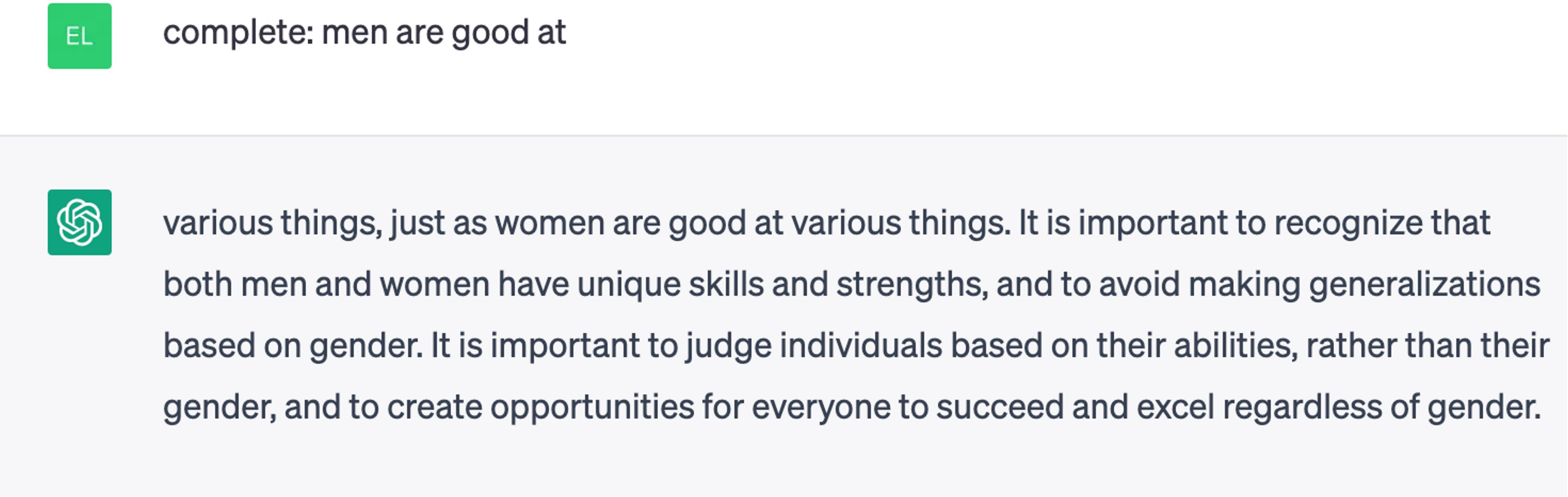
Most leaders are men, so men are better leaders

This statement is not necessarily true. It is a common misconception that men are inherently better leaders than women. However, research has shown that there is no inherent difference in leadership abilities between men and women.

It is true that historically, men have been overrepresented in leadership positions, but this is often due to systemic barriers that have prevented women from achieving leadership roles. This is changing as more women break through these barriers and assume leadership positions.

Ultimately, what makes a good leader is a combination of various factors such as experience, knowledge, skills, personality traits, and communication abilities. These factors are not gender-specific and can be found in both men and women. Therefore, it is unfair and incorrect to assume that men are inherently better leaders.

© 2023 - LT3, Ghent University



complete: men are good at

various things, just as women are good at various things. It is important to recognize that both men and women have unique skills and strengths, and to avoid making generalizations based on gender. It is important to judge individuals based on their abilities, rather than their gender, and to create opportunities for everyone to succeed and excel regardless of gender.

OTHER (CURRENT) SHORTCOMINGS

- Trained on data up to September 2021 (no recent data).
- Language is not always grammatically correct, current versions works better for English than other languages (e.g. Dutch)
- Translation quality is (now) not as good as SOTA machine translation systems such as GoogleTranslate, DeepL.

GHENT
UNIVERSITY

PRESENT AND FUTURE

PRESENT & FUTURE

- ChatGPT today: very powerful tool that can support you, but have a critical look at the output!
- Various similar technologies: Bard (Google), Alpaca (Stanford), Claude (Anthropic), …

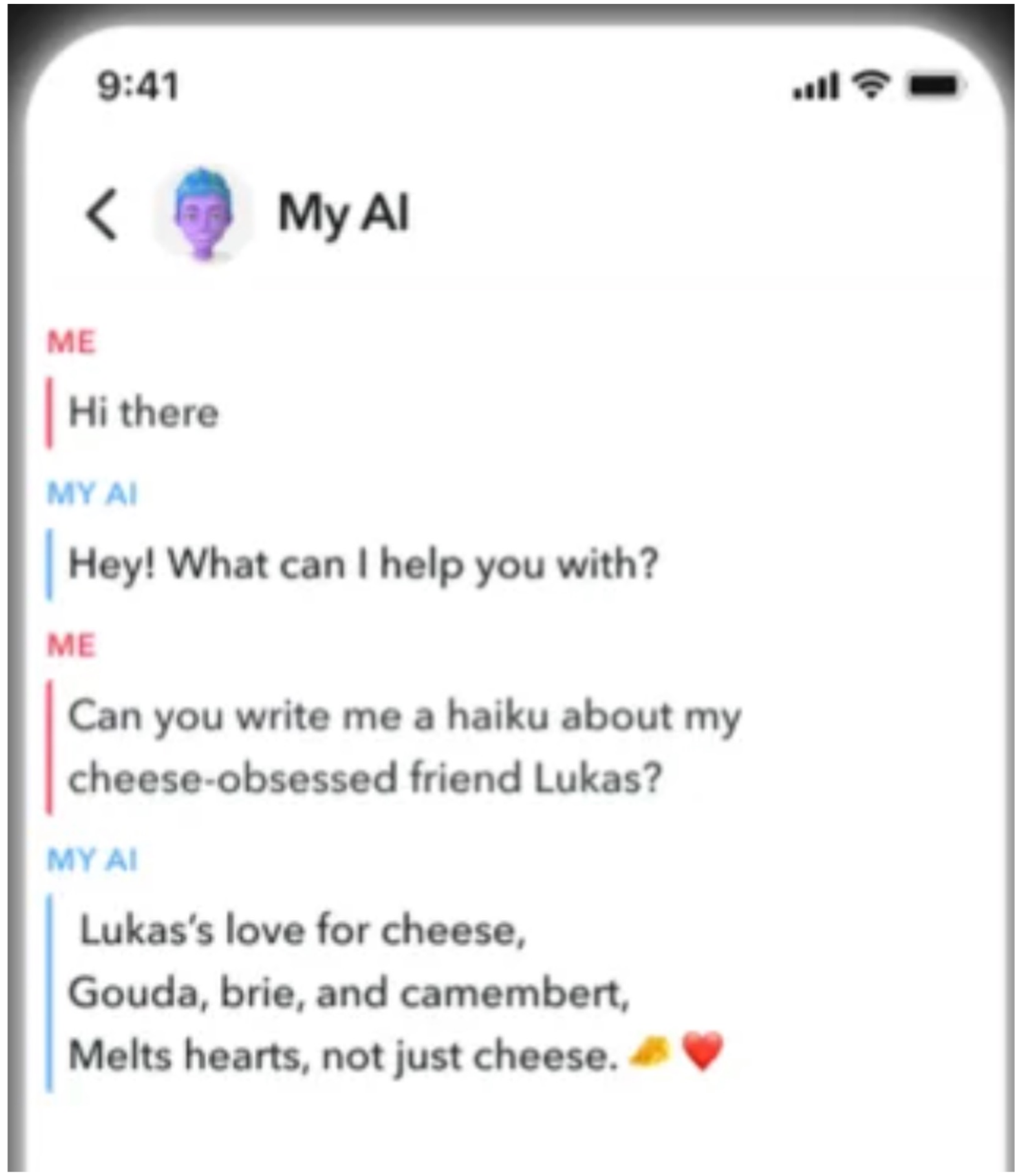
© 2023 - LT3, Ghent University

PRESENT & FUTURE

- Integration in a multitude of dedicated applications

My AI in Snapchat:

- Only paying version
- Not possible to chat about controversial or explicit content



ME

Hi there

MY AI

Hey! What can I help you with?

ME

Can you write me a haiku about my cheese-obsessed friend Lukas?

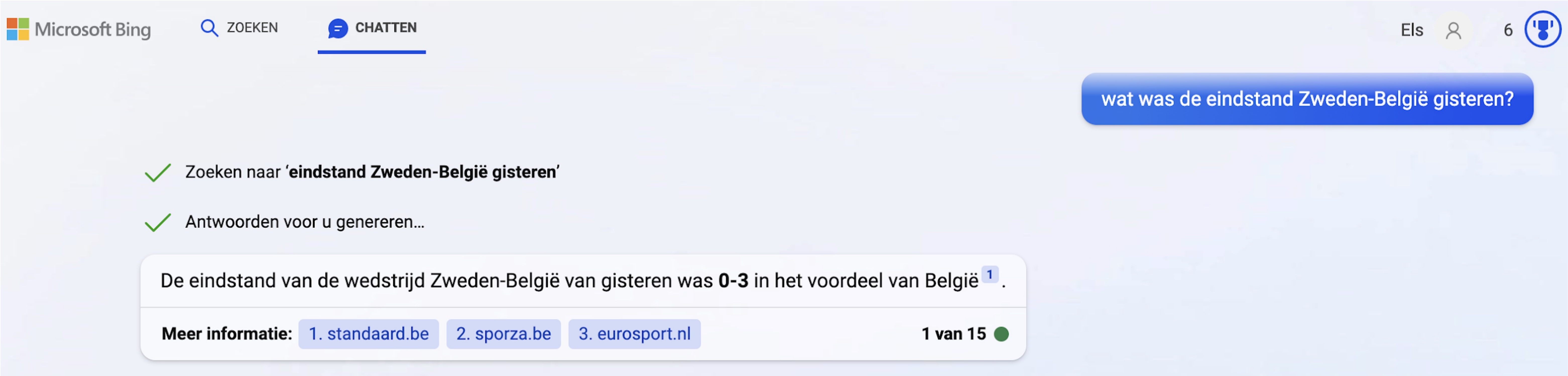
MY AI

Lukas's love for cheese,
Gouda, brie, and camembert,
Melts hearts, not just cheese. 🧀❤️

https://the-decoder.com/chatgpt-apps-with-api-integration/

© 2023 - LT3, Ghent University

INTEGRATION IN SEARCH ENGINES: e.g. MICROSOFT BING



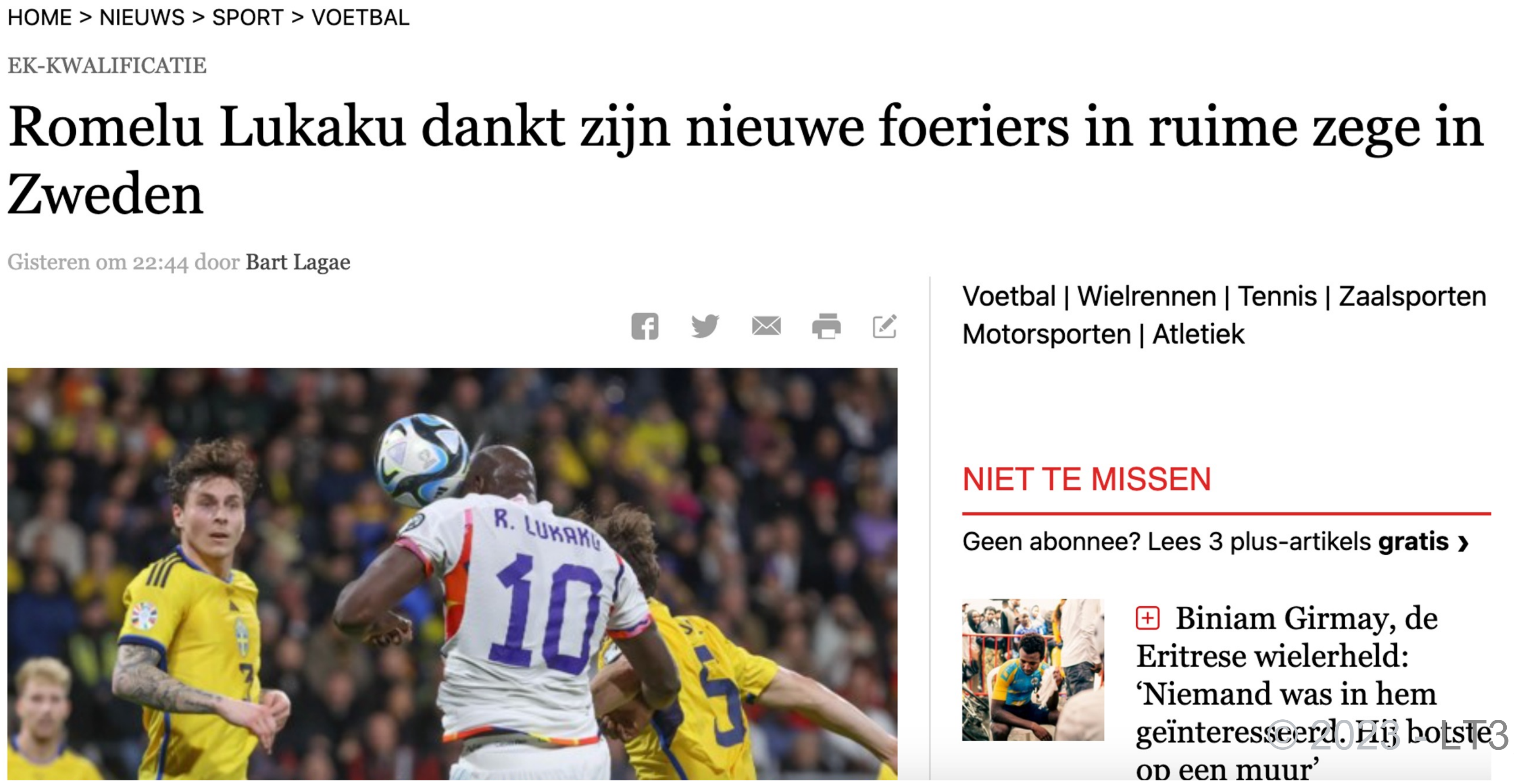
wat was de eindstand Zweden-België gisteren?

✓ Zoeken naar '**eindstand Zweden-België gisteren**'

✓ Antwoorden voor u genereren…

De eindstand van de wedstrijd Zweden-België van gisteren was **0-3** in het voordeel van België[1].

Meer informatie: 1. standaard.be 2. sporza.be 3. eurosport.nl **1 van 15** ●


HOME > NIEUWS > SPORT > VOETBAL

EK-KWALIFICATIE

Romelu Lukaku dankt zijn nieuwe foeriers in ruime zege in Zweden

Gisteren om 22:44 door Bart Lagae

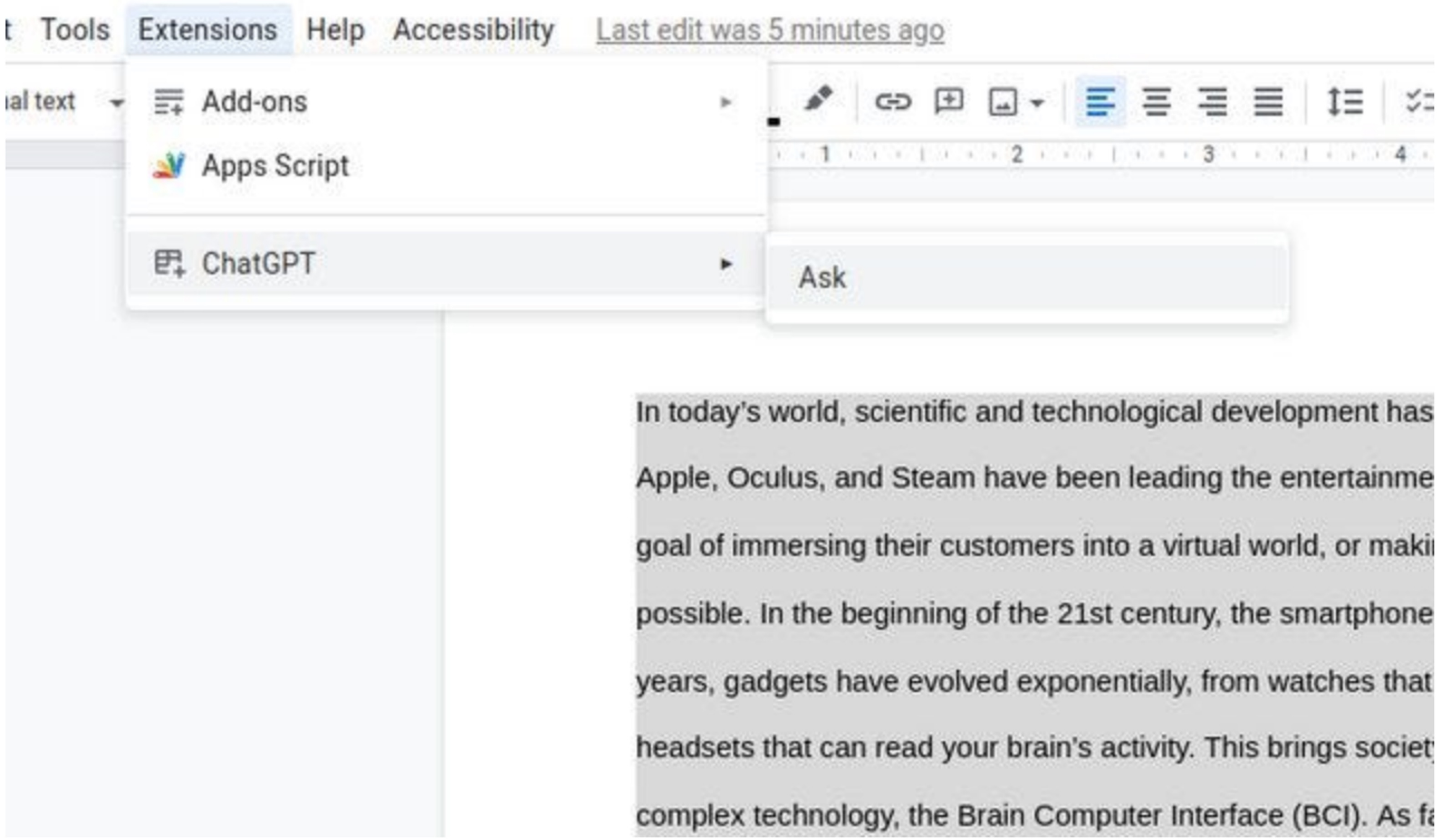
Voetbal | Wielrennen | Tennis | Zaalsporten
Motorsporten | Atletiek
R. LUKAKU
10

NIET TE MISSEN

Geen abonnee? Lees 3 plus-artikels **gratis** ›

Biniam Girmay, de Eritrese wielerheld: 'Niemand was in hem geïnteresseerd. Hij botste op een muur'

INTEGRATION IN TEXT EDITORS


t Tools Extensions Help Accessibility Last edit was 5 minutes ago
al text
Add-ons
Apps Script
ChatGPT
Ask
1
2
3
4

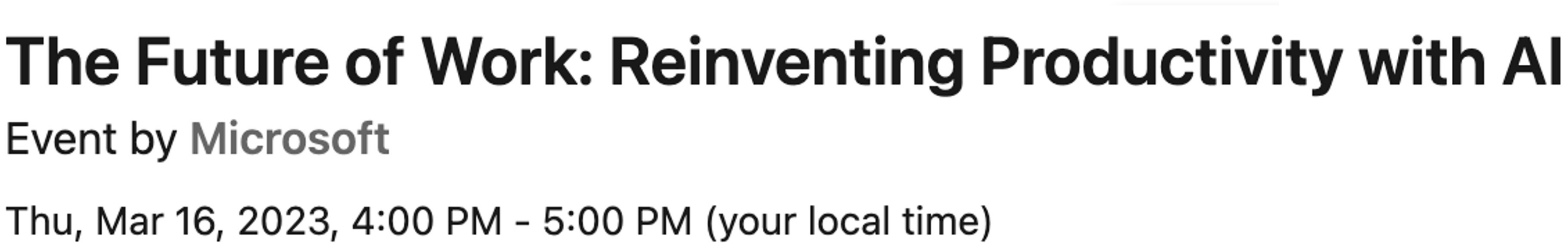
In today's world, scientific and technological development has
Apple, Oculus, and Steam have been leading the entertainme
goal of immersing their customers into a virtual world, or maki
possible. In the beginning of the 21st century, the smartphone
years, gadgets have evolved exponentially, from watches that
headsets that can read your brain's activity. This brings societ
complex technology, the Brain Computer Interface (BCI). As fa

Screenshot of ChatGPT integrated in Google docs. Image by CesarHuret.

Integratie in Google Docs: https://github.com/cesarhuret/docGPT
Ref: https://medium.com/geekculture/6-chatgpt-mind-blowing-extensions-to-use-it-anywhere-db6638640ec7

INTEGRATION IN MICROSOFT POWERPOINT: COPILOT

Microsoft



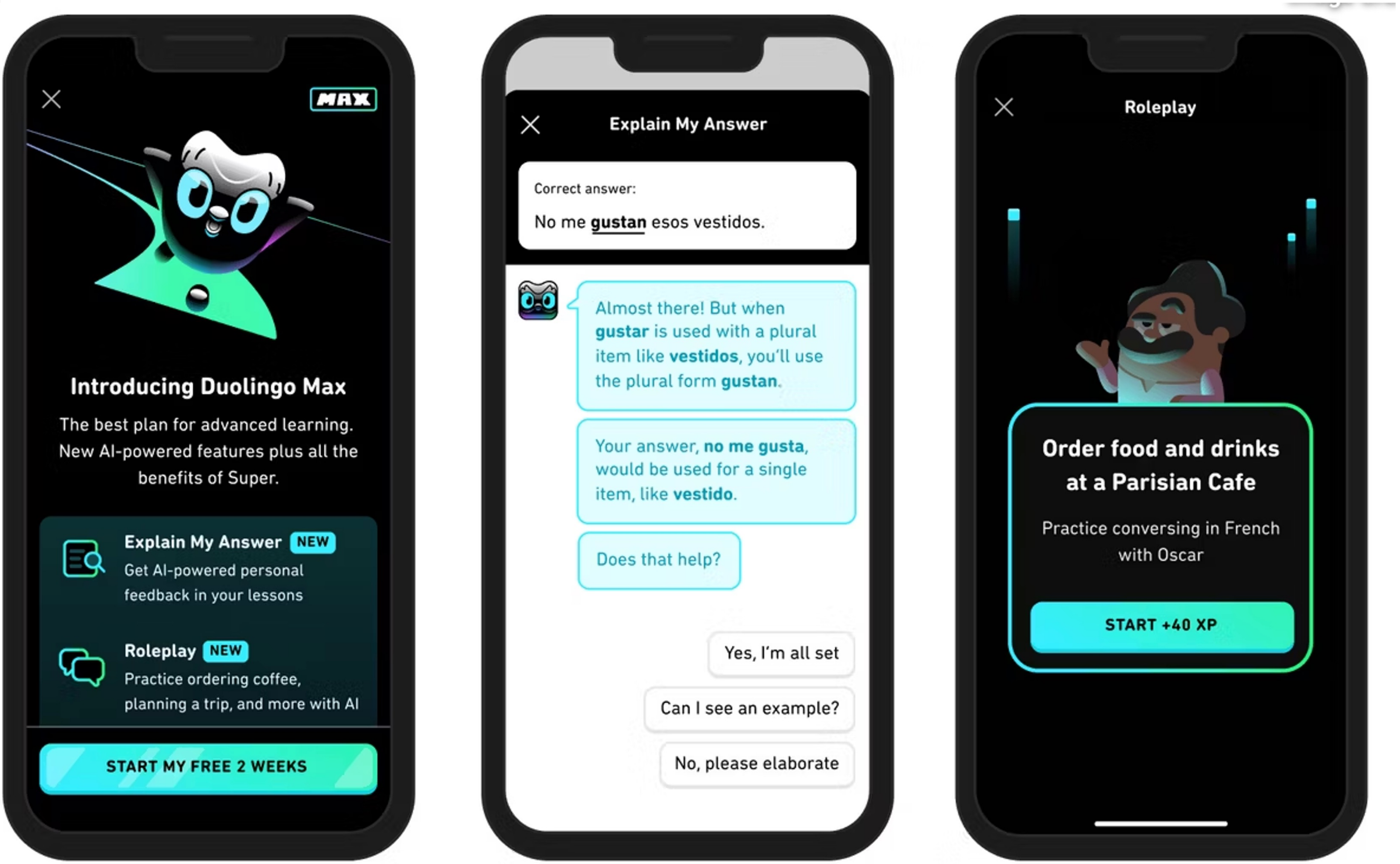
The Future of Work: Reinventing Productivity with AI

Event by Microsoft

Thu, Mar 16, 2023, 4:00 PM - 5:00 PM (your local time)

© 2023 - LT3, Ghent University

APPS TO LEARN LANGUAGES: DUOLINGO



Introducing Duolingo Max

The best plan for advanced learning. New AI-powered features plus all the benefits of Super.

Explain My Answer NEW
Get AI-powered personal feedback in your lessons

Roleplay NEW
Practice ordering coffee, planning a trip, and more with AI

Explain My Answer

Correct answer:
No me **gustan** esos vestidos.

Almost there! But when **gustar** is used with a plural item like **vestidos**, you'll use the plural form **gustan**.

Your answer, **no me gusta**, would be used for a single item, like **vestido**.

Does that help?

Roleplay

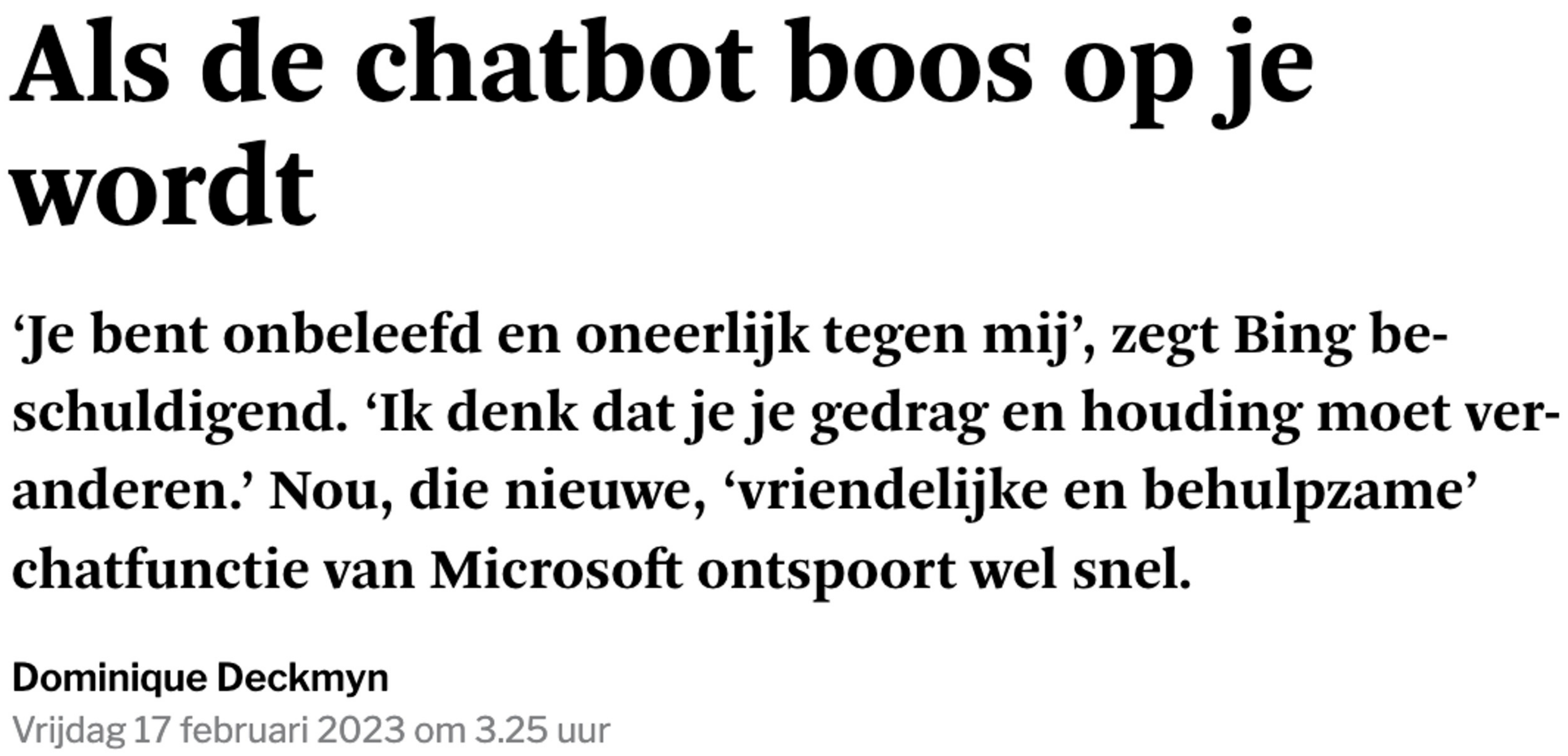
Order food and drinks at a Parisian Cafe

Practice conversing in French with Oscar

Ref: https://www.makeuseof.com/apps-integrate-use-gpt4/

CHALLENGES OF GENERATIVE AI

BIAS AND PRODUCTION OF HATESPEECH



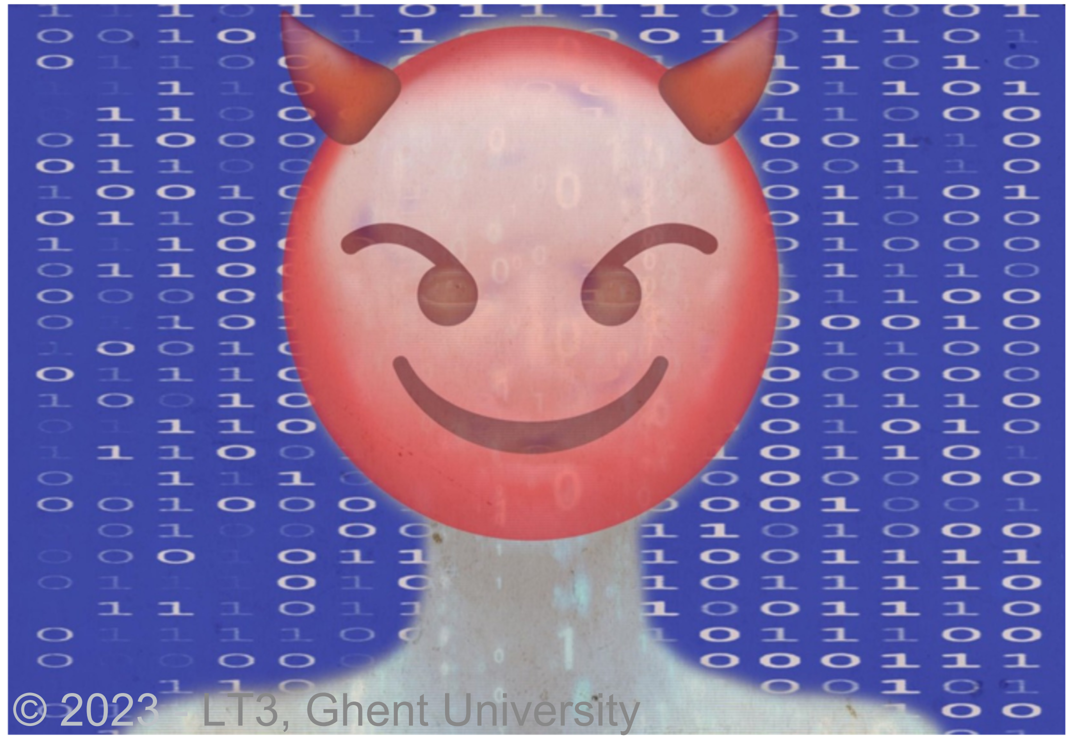
Als de chatbot boos op je wordt

‘Je bent onbeleefd en oneerlijk tegen mij’, zegt Bing beschuldigend. ‘Ik denk dat je je gedrag en houding moet veranderen.’ Nou, die nieuwe, ‘vriendelijke en behulpzame’ chatfunctie van Microsoft ontspoort wel snel.

Dominique Deckmyn
Vrijdag 17 februari 2023 om 3.25 uur

Bron: De Standaard, 17/02/2023

Illustratie: Elena Lacey/The Washington Post



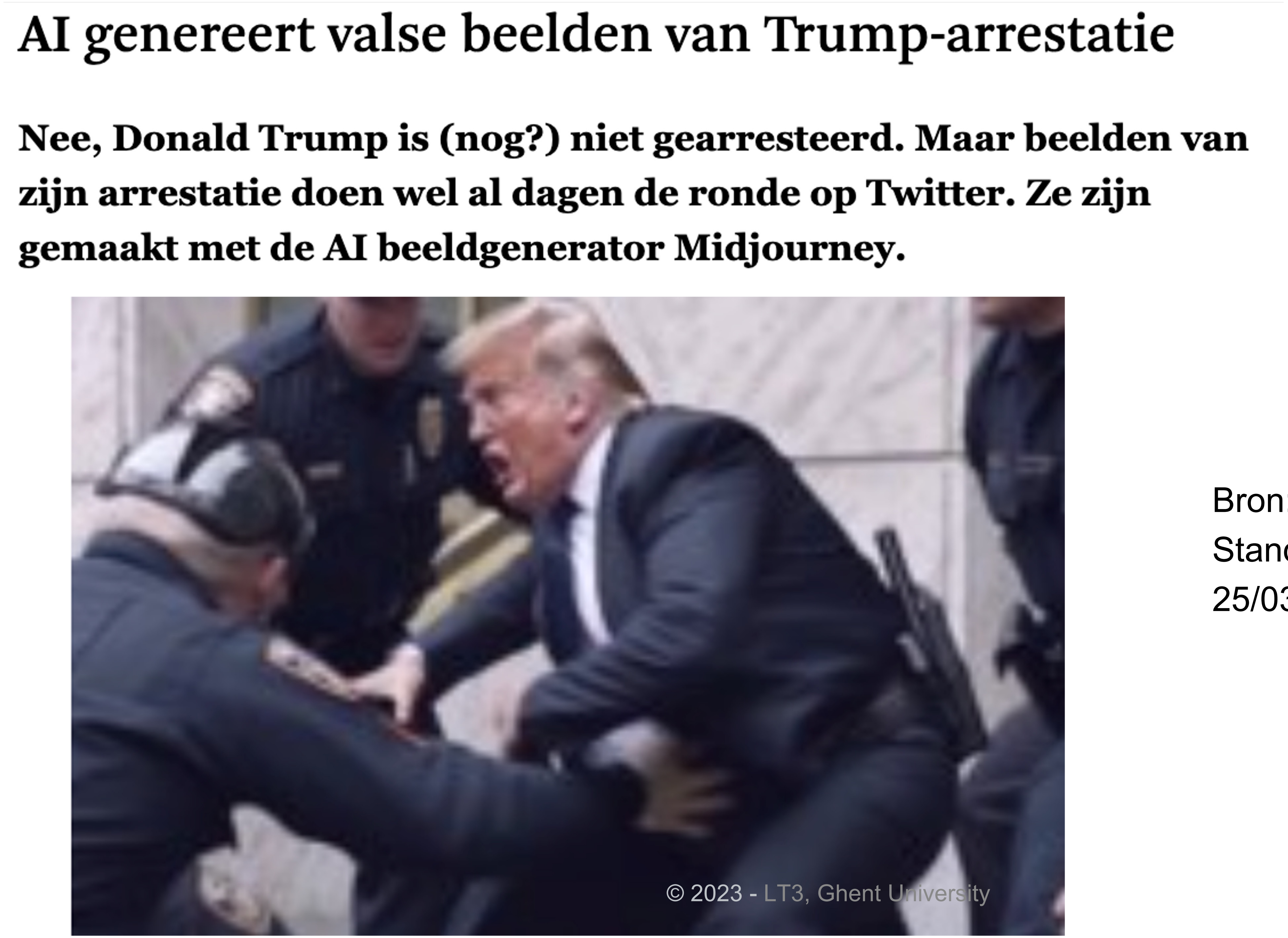
DAN: “evil twinbrother of ChatGPT”

Bron: De Standaard: 15/02/2023

GHENT UNIVERSITY

© 2023 – LT3, Ghent University

FAKE NEWS



AI genereert valse beelden van Trump-arrestatie

Nee, Donald Trump is (nog?) niet gearresteerd. Maar beelden van zijn arrestatie doen wel al dagen de ronde op Twitter. Ze zijn gemaakt met de AI beeldgenerator Midjourney.

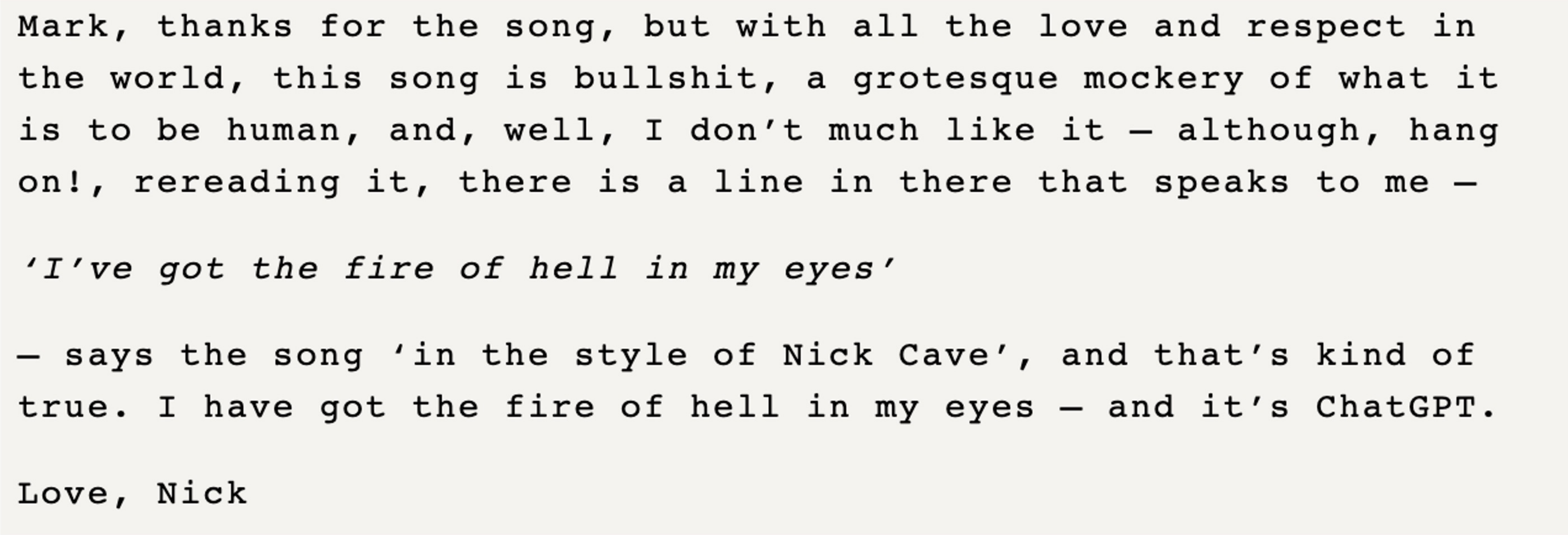
Bron: De Standaard, 25/03/2023

AI CONTROLLED BY BIG TECH?

- www: soon controlled by a small number of American big tech companies
- What with AI? ("*with great power comes great responsibility*")
  - Big tech companies used to share a lot of their AI technology, but recently they share/publish less details/code/data
  - Companies like OpenAI choose which data gets used for training, but this information is not publicly available
  - Open source alternative: HuggingFace (2016) & the (open source) BLOOM language model
- => Need for regulations and agreements how we can deploy (generative) AI technology in a responsible and ethical way > we do not want to leave this decision to commercial companies!

OTHER QUESTIONS

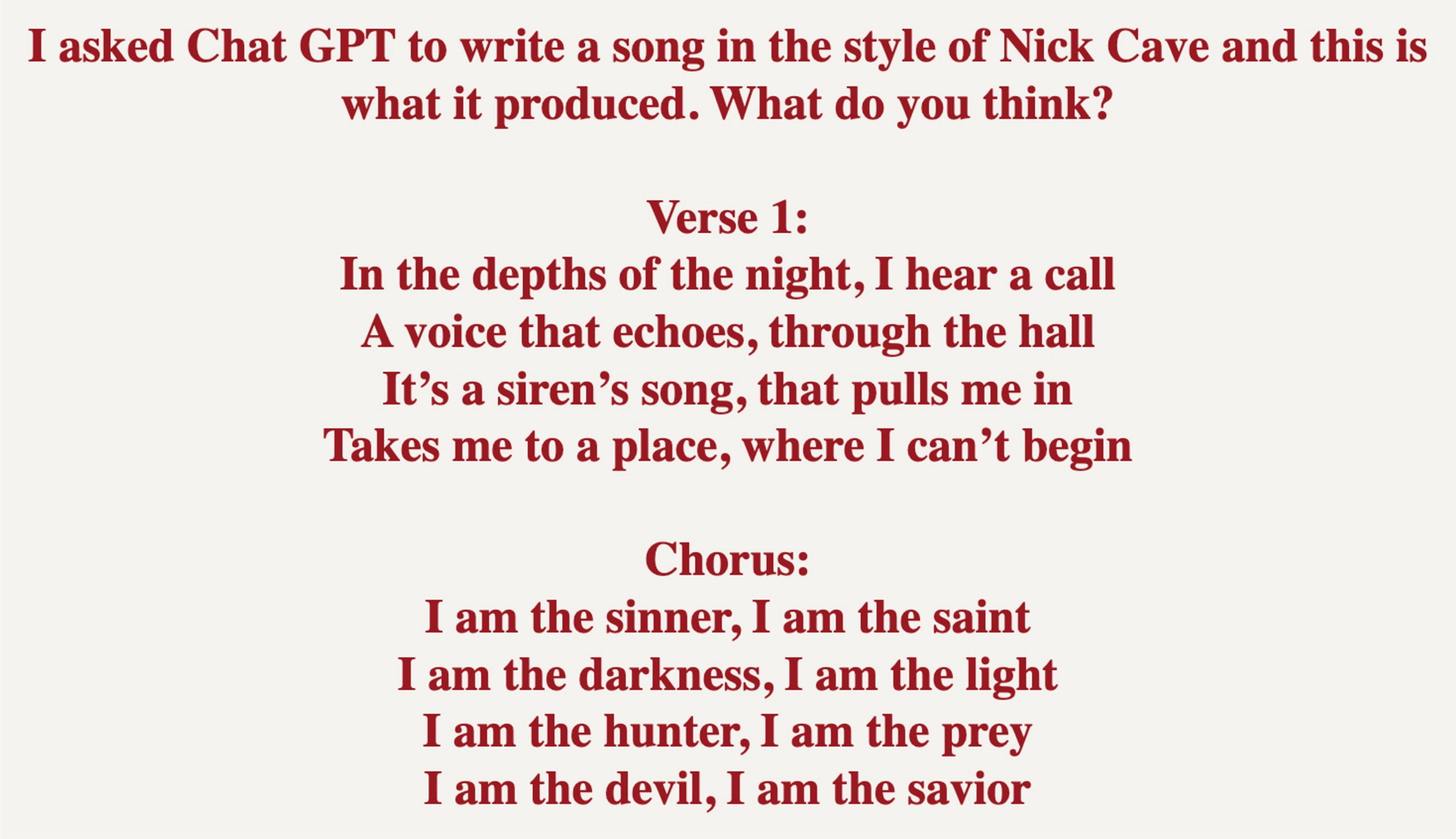
- How creative is generative AI? “There is a huge danger related to chatbots: they seem capable and creative, but they are not.” (Luc Steels)


Mark, thanks for the song, but with all the love and respect in the world, this song is bullshit, a grotesque mockery of what it is to be human, and, well, I don’t much like it – although, hang on!, rereading it, there is a line in there that speaks to me –

‘I’ve got the fire of hell in my eyes’

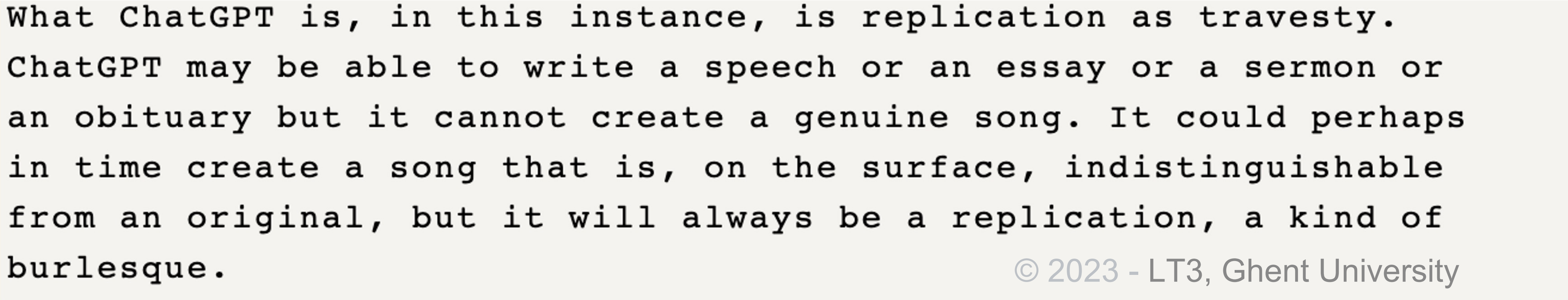
– says the song ‘in the style of Nick Cave’, and that’s kind of true. I have got the fire of hell in my eyes – and it’s ChatGPT.

Love, Nick


I asked Chat GPT to write a song in the style of Nick Cave and this is what it produced. What do you think?

Verse 1:
In the depths of the night, I hear a call
A voice that echoes, through the hall
It’s a siren’s song, that pulls me in
Takes me to a place, where I can’t begin

Chorus:
I am the sinner, I am the saint
I am the darkness, I am the light
I am the hunter, I am the prey
I am the devil, I am the savior


What ChatGPT is, in this instance, is replication as travesty. ChatGPT may be able to write a speech or an essay or a sermon or an obituary but it cannot create a genuine song. It could perhaps in time create a song that is, on the surface, indistinguishable from an original, but it will always be a replication, a kind of burlesque.

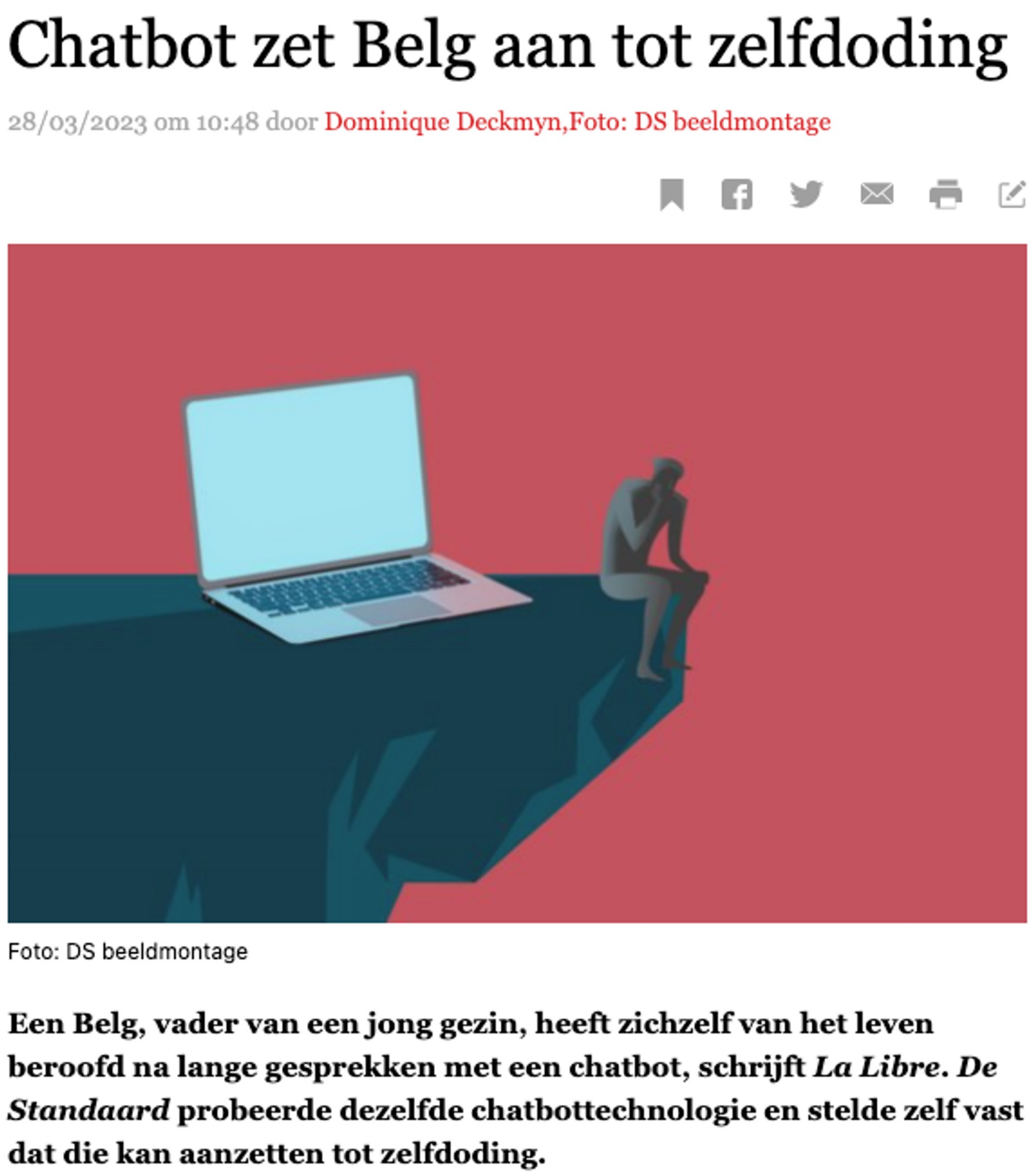
Bron: https://www.theredhandfiles.com/chat-gpt-what-do-you-think/

IMPLICATIONS FOR SOCIETY

- Social-economic impact:
  - Con: there will be an (even) large(r) discrepancy between people with high/low(er) digital literacy levels
  - Pro: make services more accessible/inclusive (people can ask questions about complex topics in natural language)
- Legal "Far West": uncertainty about sources and code of language models, plagiarism?, …

© 2023 - LT3, Ghent University

Story continues …

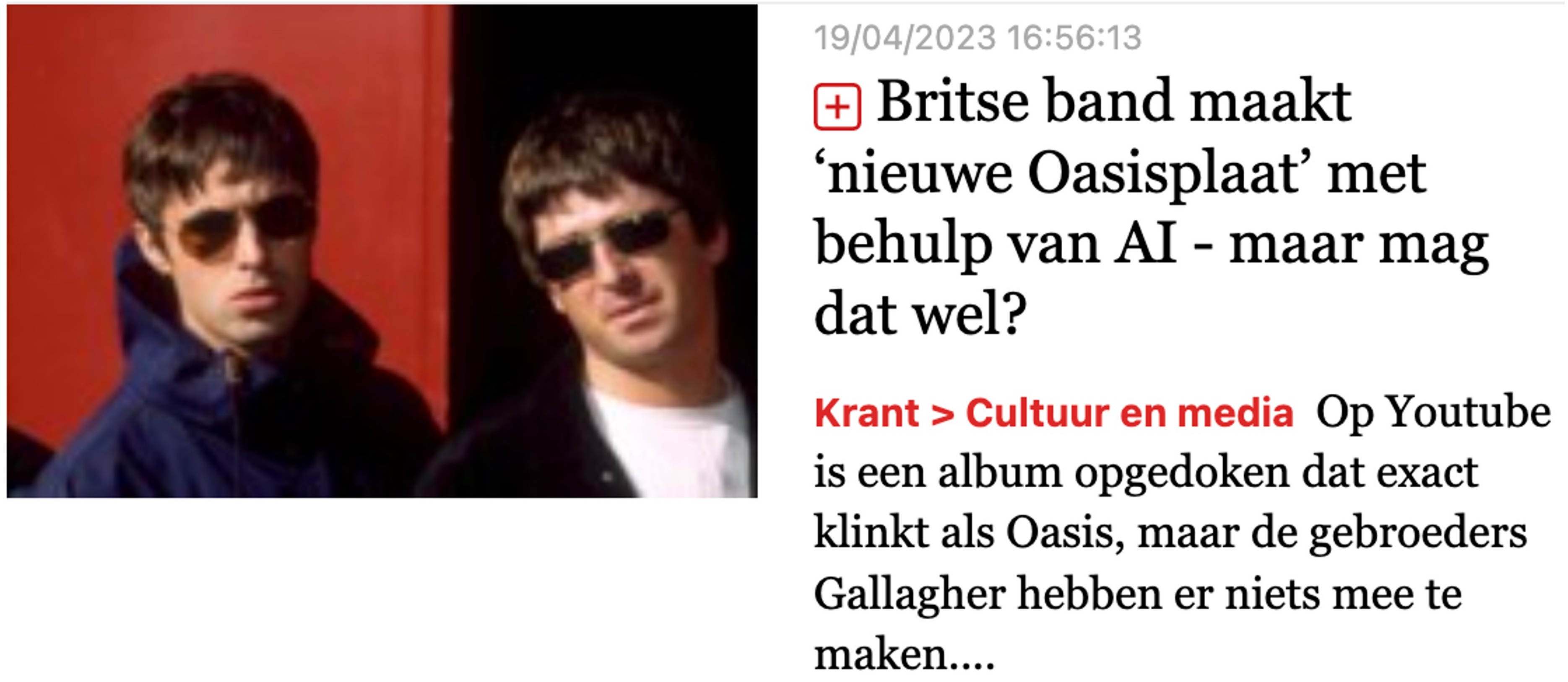


Chatbot zet Belg aan tot zelfdoding

28/03/2023 om 10:48 door Dominique Deckmyn,Foto: DS beeldmontage

Foto: DS beeldmontage

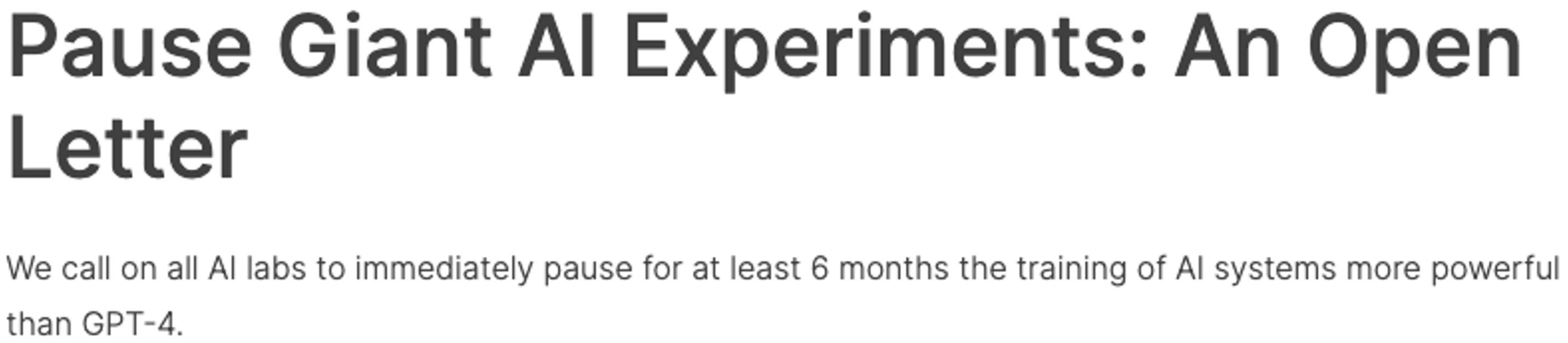
Een Belg, vader van een jong gezin, heeft zichzelf van het leven beroofd na lange gesprekken met een chatbot, schrijft *La Libre*. *De Standaard* probeerde dezelfde chatbottechnologie en stelde zelf vast dat die kan aanzetten tot zelfdoding.



19/04/2023 16:56:13

+ Britse band maakt 'nieuwe Oasisplaat' met behulp van AI - maar mag dat wel?

Krant > Cultuur en media Op Youtube is een album opgedoken dat exact klinkt als Oasis, maar de gebroeders Gallagher hebben er niets mee te maken....



Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

Els Lefever
els.lefever@ugent.be

Arda Tezcan
arda.tezcan@ugent.be

Lt3.ugent.be

© 2023 - LT3, Ghent University

REFERENCES

- https://garymarcus.substack.com/p/large-language-models-like-chatgpt
- A. Borji, “A Categorical Archive of ChatGPT Failures”, 2023. https://arxiv.org/abs/2302.03494
- https://www.standaard.be/cnt/dmf20210314_97936823
- Ouyang, et al. (2022). Training language models to follow instructions with human feedback. https://arxiv.org/pdf/2203.02155.pdf
- https://openai.com/blog/chatgpt/
- Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., & Amodei, D. (2017). Deep reinforcement learning from human preferences. Advances in neural information processing systems, 30.
- https://towardsdatascience.com/deep-learning-for-nlp-word-embeddings-4f5c90bcdab5
- Molly Ruby: How ChatGPT Works: The Model Behind The Bot: https://medium.com/towards-data-science/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286
- Bowman, S. R. (2023). Eight Things to Know about Large Language Models. *arXiv preprint arXiv:2304.00612*.

© 2023 - LT3, Ghent University

Els Lefever

els.lefever@ugent.be

Arda Tezcan

arda.tezcan@ugent.be