

Wetenschapscommunicatie on steroids

Waar komt GPT vandaan, en hoe gebruik je het optimaal?

Thomas Winters

PhD Student at KU Leuven & FWO Fellow

@thomas_wint

thomaswinters.be

Ben Verhoeven

ERLNMYR

@_ERLNMYR

erlnmyr.be

Slides: thomaswinters.be/nwd

Wie heeft al tekstgenerator (ChatGPT, Bing Assistant...) gebruikt?



Wie heeft tekstgenerator voor wetenschapscommunicatie gebruikt?



Wie gebotst op taken dat tekstgenerator niet kon?



Thomas Winters

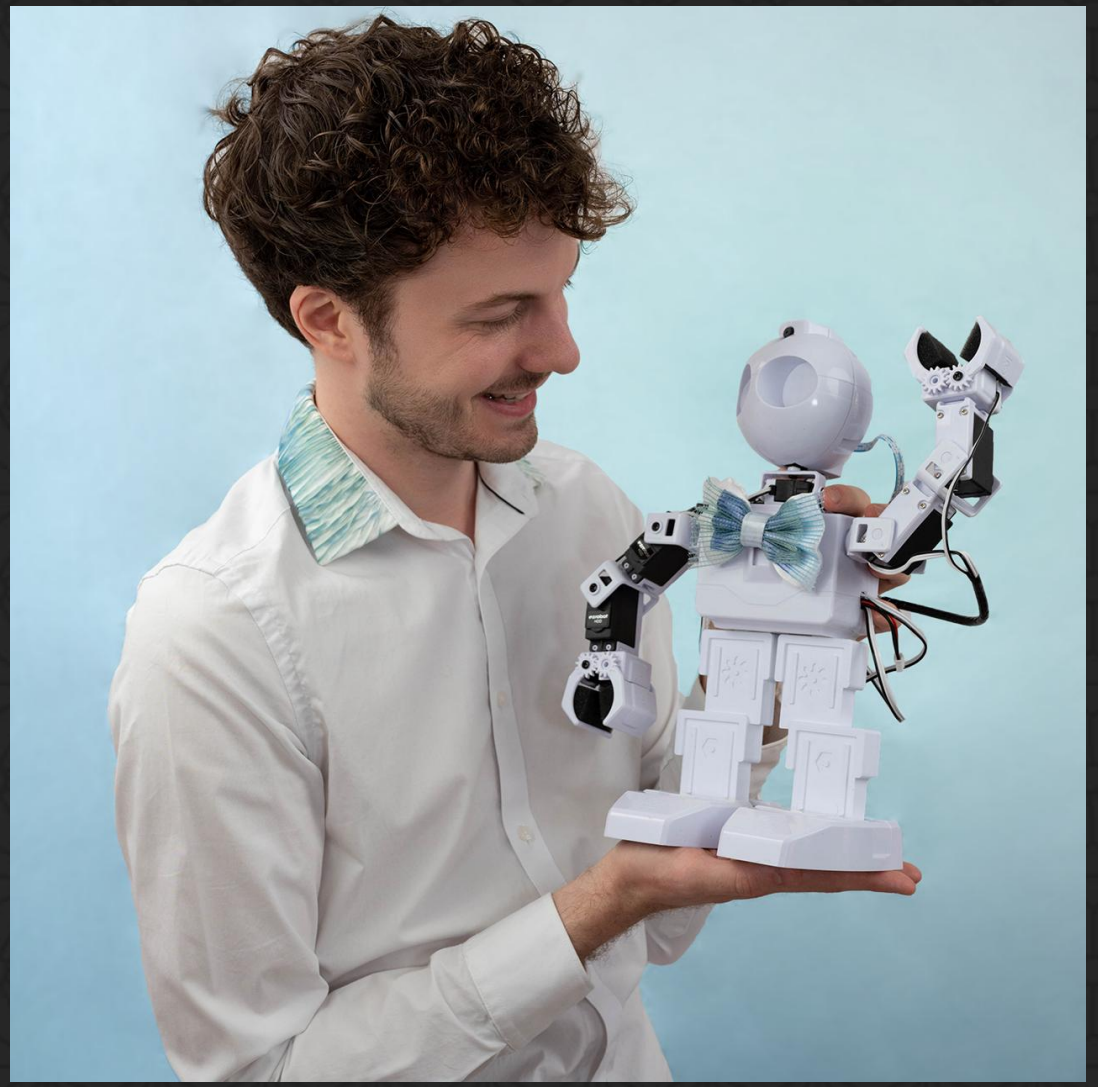
- PhD student in creatieve AI & taalmodellen sinds 2018

  → Deel van DTAI research group in ML cluster van professors Luc De Raedt, Jesse Davis, Hendrik Blockeel, Tias Guns...

- Mede-auteur RobBERT model
  (state-of-the-art Nederlandstalige BERT taalmodel)

- Sinds GPT-2 (2019) ervaring met GPT-modellen *(o.a. voor humor, Improbotics...)*

KU LEUVEN LEUVEN.AI DTAI





Tekstgeneratie

Van autocomplete naar ChatGPT

Hoe werkt een tekstgenerator?



1. Open smartphone **toetsenbord** op tekstveld
2. Druk op een voorgestelde **auto-complete** woord
3. Herhaal 10-tal keer
4. Je hebt een tekst gegenereerd die klinkt zoals jezelf!

Auto-complete telde frequentie van woorden na vorige woorden in je teksten

TorfsBot

1. Telde in alle tweets & columns van Rik Torfs hoe vaak woord andere woorden volgde



2. Dan neemt het paar willekeurige beginwoorden, en blijft het woorden die Rik gebruikte kiezen.





GPT training



Heel het internet

Trainingsproces van paar miljoen $

GPT-4

Beste publieke taalmodel

GPT modellen



≈ autocomplete on steroids

The next word for this charade is

sentence = 10.48%
list = 6.61%
phrase = 6.43%
definition = 6.12%
sequence = 4.07%
char = 0.15%

Total: -6.49 logprob on 1 tokens
(33.86% probability covered in top 6 logits)

Kans op volgende token gegeven *vorige tokens*

Tokens	Characters
119	414

GPT-3 tokenizers words to be able to input them as integers in the model.

The tokens are chosen in such a way that common substrings have their own token. It is purely frequency-based, and do not need to make morphological sense: e.g. gothic

In English, common words are usually a single token.
In het Nederlands is dat vaak andere koek.
Vandaar dat het moeilijker lettergrepen en rijmen snapt in het Nederlands.

https://platform.openai.com/tokenizer

Drie GPT trainingfasen

GPT-3 (2020)

Leerde internet woord voor woord nabootsen

InstructGPT (2022)

Leerde om naar instructies te luisteren

ChatGPT & GPT-4 (2022-2023)

Leerde instructies “volgens menselijke voorkeur” op te lossen

Voornaamste sterkte

USER

Herschrijf deze mail naar extreem formele stijl:

Dag Nationale Wetenschapscommunicatiedag organisatoren, zoals gevraagd kunnen we inderdaad een presentatie over de werking van GPT & prompt engineering voor wetenschapscommunicators geven als jullie willen! Groetjes Ben Verhoeven & Thomas Winters

ASSISTANT

Geachte organisatoren van de Nationale Wetenschapscommunicatiedag,

Naar aanleiding van uw vriendelijk verzoek hebben wij, de heer Ben Verhoeven en de heer Thomas Winters, het genoegen u te kunnen bevestigen dat wij in staat zijn om een presentatie te geven over het onderwerp 'De Werking van GPT & Prompt Engineering'. We zijn ervan overtuigd dat dit een relevant en inspirerend onderwerp zal zijn voor de wetenschapscommunicatoren.

We kijken ernaar uit om ons onderzoek, kennis en inzichten met de aanwezigen te delen op deze prestigieuze bijeenkomst.

Indien u verdere informatie of bijzonderheden wenst te ontvangen, staat het u vrij om contact met ons op te nemen.

Hoogachtend,

Ben Verhoeven & Thomas Winters

Copy writing → Copy editing

USER	Herschrijf deze mail naar een pitch over de talk in extreem simpele taal zodat 5-jarige kinderen het kunnen begrijpen waarover het gaat
ASSISTANT	Hallo vriendjes! Wij zijn Ben en Thomas en wij gaan een superleuke praatje houden! Weet je wat wij gaan doen? We gaan jullie alles vertellen over hoe we computers heel, heel slim maken, dat noemen we 'GPT & Prompt Engineering'. Het is alsof we de computer leren hoe het moet denken en praten, net zoals jij en ik! Hoe cool is dat? We kunnen niet wachten om met jullie te spelen en leren! Knuffels, Ben en Thomas

USER

Schrijf een Disneyverhaal hierover

ASSISTANT

Er was eens, in het verre koninkrijk der wetenschap, twee slimme en avontuurlijke vrienden, Ben en Thomas. Ze waren geen gewone vrienden; ze waren tovenaars van technologie. Ze speelden niet alleen met toverstokken en magische spreuken, maar met codes en commands.

Op een dag kregen ze een uitnodiging van het grote koninklijke evenement 'Nationale Wetenschapscommunicatiedag'. De organisatoren van het evenement nodigden ze uit om te komen spreken over hun magische manier van het omtoveren van computers in slimme helpers. Dit noemden ze 'GPT & Prompt Engineering', een beetje ingewikkeld, maar oh zo boeiend.

"Thomas," zei Ben enthousiast, "stel je voor hoe de wereld zou veranderen als iedereen zou kunnen leren over onze magie!"

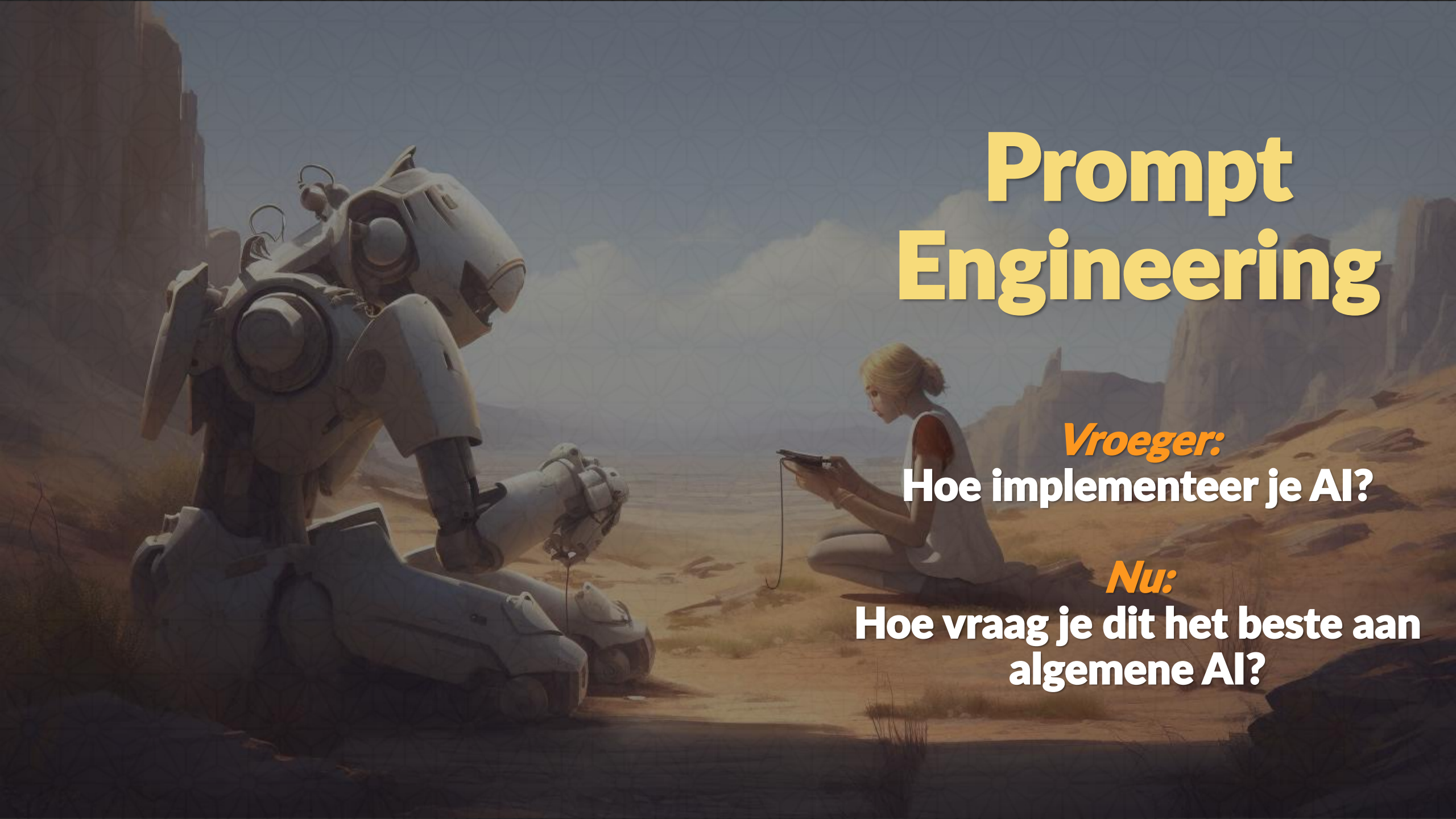
Dus accepteerden ze de uitnodiging met glinstering in hun ogen. Ze gingen aan het werk, creëerden hun presentatie met dezelfde zorg waarmee ze hun toverspreuken creëerden.

Op de grote dag stonden ze voor een publiek van honderden nieuwsgierige wetenschappers en jonge aspirant-tovenaars. Met elke zin, spreuk en dia veranderden ze het publiek in een wereld van magisch denkende computers. Hun woorden stuwden golven van bewondering en verbazing door de zaal. Ze toonden de kracht van technologie, de schoonheid van het onbekende en de mogelijkheden van de toekomst.

Omdat, zoals ze geloofden, de beste manier om de toekomst te voorspellen, is om hem zelf te creëren. En dat deden ze, Ben en Thomas, de magiërs van GPT en Prompt Engineering.

En sindsdien, door hun talent om kennis te delen, inspireerden ze velen om ook tovenaars van technologie te worden. En ze leefden nog lang en gelukkig, ze vuurden hun magische comando's af en maakten computers slim, één code tegelijk.

En zo eindigt ons prachtige Disneyverhaal over Ben en Thomas, de magische technologietovenaars.


Prompt Engineering
Vroeger:
Hoe implementeer je AI?
Nu:
Hoe vraag je dit het beste aan algemene AI?

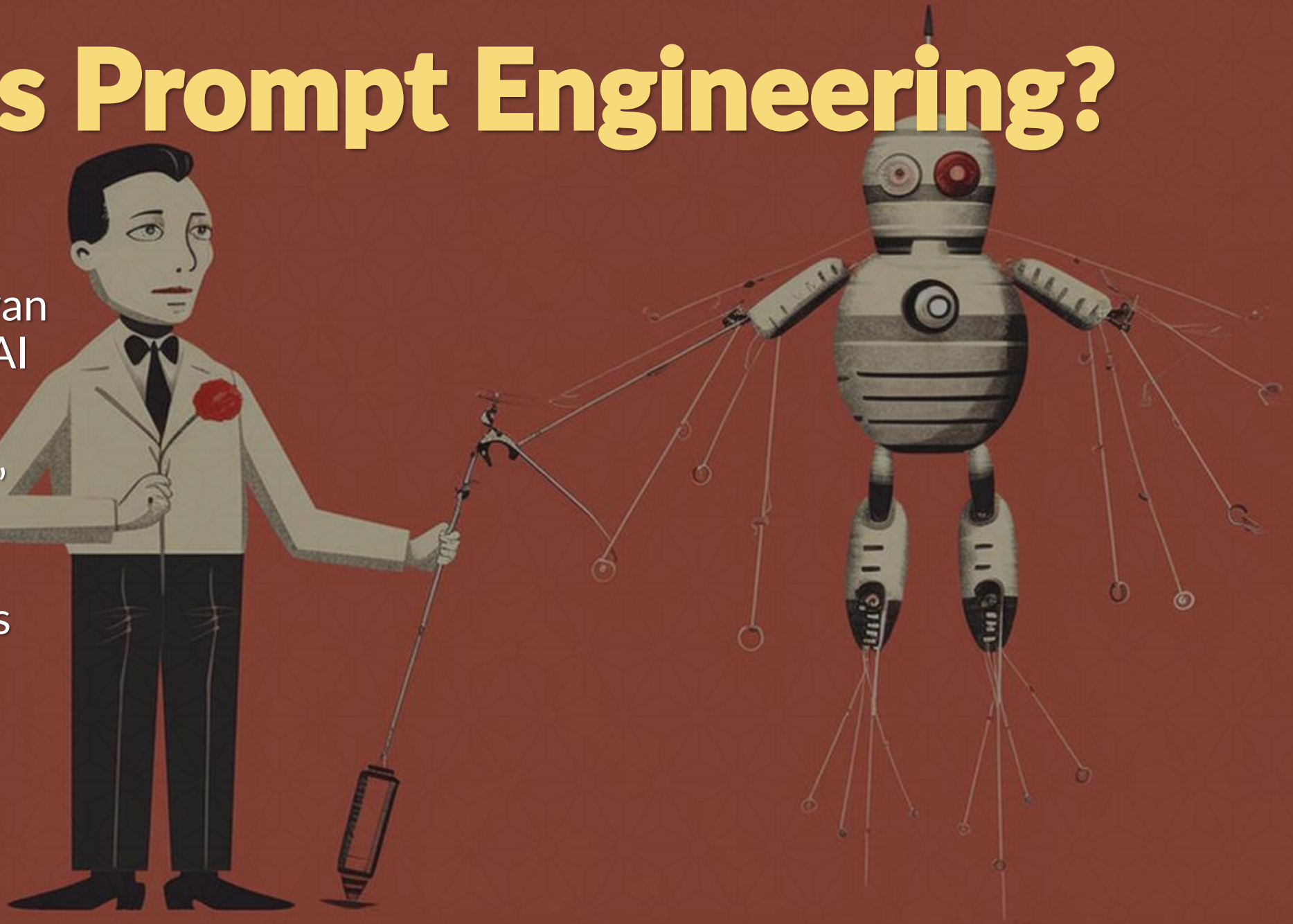
Wat is Prompt Engineering?

Optimaal kiezen van instructies voor AI

“Voodoo magic”

Bewezen trucjes

Vaak Intuitief



Basis: Schrijfinstructie prompt

USER	Schrijf een lijst aan titels voor een lezing over hoe je metaforen kan gebruiken in een wetenschapscommunicatiecontext
ASSISTANT	1. "Bruggen Bouwen: Metaforen Gebruiken in Wetenschapscommunicatie" 2. "De Kunst van Zoet-Water Taal: Metaforen in Wetenschapscommunicatie" 3. "Metaforenmagie: Effectieve Wetenschapscommunicatie" 4. "Wetenschap in Alledaagse Termen: Het Belang van Metaforen" 5. "Synergie tussen Metaforen en Wetenschap: Een Communicatiebenadering"

Truc 1: Geef handvol voorbeelden *(“few-shot prompting”)*

USER

Gegven deze onderwerpen voor wetenschapscommunicatie talks en bijhorende titels:

Onderwerp: Improtheater
Titel: De kracht van Improvisatie

Onderwerp: AI
Titel: Wetenschapscommunicatie on steroids

Onderwerp: Debatten
Titel: Dialoog als verbindend wondermiddel?

Schrijf nu vijf titels voor "Onderwerp: Metaforen"

ASSISTANT

1. Het verborgen vermogen van Metaforen
2. Metaforen: De brug tussen complexe ideeën en begrip
3. Meertalige Mirakels: De mysterieuze magie van metaforen
4. Metaforen: Beelden voor uw Brein
5. De kracht van verborgen betekenissen: Ontcijfer de taal van metaforen

Makkelijke manier om “stijl” duidelijk te maken

Truc 2: Maak rol duidelijk *("role-prompting")*

SYSTEM

Je bent een wereldexpert in wetenschapscommunicatie met een bewezen track record in verzinnen van strakke titels voor je lezingen.

USER	Schrijf vijf titels voor een lezing over hoe je metaforen kan gebruiken in een wetenschapscommunicatiecontext
ASSISTANT	1. "Beeldend Spreken: Metaforen als Sleutel tot Effectieve Wetenschapscommunicatie" 2. "Creatief Verhelderen: De Kunst van Metaforen in Wetenschapscommunicatie" 3. "Bruggen Bouwen met Woorden: De Rol van Metaforen in Wetenschapscommunicatie" 4. "De Metamorfose van de Metafoor: Naar een Nieuwe Taal van de Wetenschap" 5. "De Kracht van Verbeelding: Metaforen als Instrument voor Wetenschapscommunicatie"

Laat GPT expert nadoen, bv:

"You are a world-renowned expert X with proven track record on X."

"localiseert" taak in het model

Meer inspiratie:
https://github.com/f/awesome-chatgpt-prompts

GPT-3 is slecht in wiskunde?

Thomas takes a train from Leuven to Gent. Leuven Station and Gent have a distance of 100km. The train travels at 50km/h. The train also stops three times for 10 minutes each. How long is the train ride?

2 hours

Truc 3: Vraag redeneerstappen

Thomas takes a train from Leuven to Gent. Leuven Station and Gent have a distance of 100km. The train travels at 50km/h. The train also stops three times for 10 minutes each. How long is the train ride? Let's think step by step.

The train travels at 50km/h, so it will take 2 hours to travel 100km. However, the train stops three times for 10 minutes each. This means that the train will spend a total of 30 minutes stopped. Therefore, the total train ride will be 2 hours and 30 minutes.

USER

Schrijf vijf titels voor een lezing over hoe je metaforen kan gebruiken in een wetenschapscommunicatiecontext.

Denk eerst stap voor stap na over hoe je best zo een titel schrijft en zet bij elke stap al mogelijke ideetjes.

ASSISTANT

Stap 1: Verken het hoofdthema
Het hoofdthema is "het gebruik van metaforen in wetenschapscommunicatie". Dit gaat over hoe we complexe wetenschappelijke ideeën kunnen overdragen aan algemene publieken via beeldende en inzichtelijke taal - in dit geval, metaforen.

Stap 2: Identificeer belangrijke subthema's
Er kunnen diverse subthema's zijn waaronder: 'De psychologie achter metaforen', 'Wetenschapsleer door metaforen', 'Metaforen in de huidige wetenschapscommunicatie', 'Het krachtige effect van metaforen', etc.

Stap 3: Voeg een creatieve draai toe
Gebruik humor, woordspelingen, intrigue of een uitdaging in je titel om interesse op te wekken. Denk aan intrigerende vragen of stellingen waarbij het publiek meer wil weten.

Stap 4: Houd het relevant
Zorg ervoor dat de titel daadwerkelijk de inhoud van de lezing weerspiegelt en niet als misleidend wordt bevonden.

Nu we dit proces hebben doorgelopen, hier zijn vijf mogelijke titels:

1. "De Kracht van Verbeelding: Metaforen in Wetenschapscommunicatie"
2. "MetaFor(e) Science: Hoe Figuurlijke Taal de Wetenschap Helpt Verkennen"
3. "Van Microben tot Melkwegstelsels: Metaforen Maken Wetenschap Bevattelijk"
4. "Sleutelen aan de Wetenschap: Het Ontsloten Potentieel van Metaforen"
5. "Figuratieve Fenomenen: De Impact van Metaforen op de Beleving van Wetenschap"

Truc 4: Specifieer zelf stappen

Trick 4: Make it do multiple explicit steps

BE

Do the next steps in order:
1) Make a list of great but not famous scientists
2) Select one scientist from step 1 and think of three reasons why they should be more known
3) Choose one of those reasons from step 2 and write a blog post advocating for an auditorium to be named after this scientist

Sure! Here are the next steps in order:

1. Make a list of great but not famous scientists:
   - Dr. Sarah Bennett: A marine biologist who has made significant contributions to understanding coral reef ecosystems and their conservation.
   - Dr. Rajesh Patel: A physicist known for his groundbreaking work in the field of quantum computing.
   - Dr. Emily Chen: An environmental scientist who has made notable advancements in sustainable agriculture practices.
2. Select one scientist from step 1 and think of three reasons why they should be more known:
   Let's choose Dr. Sarah Bennett as the scientist to focus on.
   - Dr. Bennett's research on coral reefs has provided valuable insights into their ecological importance and the threats they face from climate change. By raising awareness about her work, we can highlight the urgent need to protect these delicate ecosystems.
   - Dr. Bennett has actively engaged with local communities and policymakers to implement effective conservation strategies. Recognizing her efforts can inspire others to take similar actions and foster greater collaboration between scientists and communities.

Regenerate response | Continue generating

3. Choose one of those reasons from step 2 and write a blog post advocating for an auditorium to be named after this scientist:
   Title: Honoring Dr. Sarah Bennett: A Pioneer in Coral Reef Conservation
   Introduction:
   As we celebrate the achievements of renowned scientists throughout history, it is crucial not to overlook the contributions of those who have made significant strides in their respective fields but remain relatively unknown. One such scientist who deserves recognition is Dr. Sarah Bennett, a marine biologist whose work on coral reef conservation has had a profound impact on both science and society. In this blog post, we advocate for an auditorium to be named after Dr. Sarah Bennett, honoring her dedication to the study and preservation of these fragile ecosystems.
   Body:
   3.1. Highlighting the ecological importance of coral reefs:
   Dr. Bennett's research has shed light on the vital role coral reefs play in supporting marine biodiversity and coastal communities worldwide. By naming an auditorium after her, we can bring attention to the urgent need for their preservation, especially in the face of climate change. This act would serve as a powerful reminder of the environmental challenges we face and the collective responsibility we hold to protect our planet.
   3.2. Recognizing community engagement and effective conservation strategies:
   Dr. Bennett's unwavering commitment to engaging with local communities and policymakers has been instrumental in implementing successful coral reef conservation initiatives. By naming an auditorium in her honor, we can inspire scientists, policymakers, and community leaders to work together towards sustainable solutions. The auditorium can serve as a spa[illegible] best practices in

Regenerate response | Continue generating

Truc 5: Leg denkstappen vooraf uit in voorbeelden

TH

Je bent een wereldexpert in het plannen van wetenschappelijke en technologie-gerelateerde evenementen, met een bewezen track record op het Departement Economie, Wetenschap en Innovatie van Vlaanderen. Ik ben op zoek naar ideeën voor nieuwe evenementen. Je schrijft steeds je redenering voor het evenement.

Invoer
Topic: wetenschap
Redenering
Redenering: maak de jeugd enthusiast krijgen voor wetenschap, zodat ze later een STEM richting kiezen
Invoer
Evenement: "ScienceVille": een Leuvens festival met de focus op de wonderen van de wetenschap

Invoer
Topic: bio-economie
Redenering
Redenering: Zet het toekomstpotentieel van bio-economie in de verf. We kunnen best de nieuwe generatie studenten aan professionals voorstellen.
Invoer
Evenement: "Bio-economie: van ambitie naar actie": een actie om de partners voor nieuwe acties samen te brengen

Invoer
Topic: CERN
Redenering
Redenering: We kunnen studenten CERN laten ontmoeten om hen zo aan te moedigen om stages en jobs bij CERN te overwegen
Invoer
Evenement: "CERN Roadshow": CERN ontmoet studenten op 6 locaties in Vlaanderen & Brussel

Schrijf drie voorstellen voor een evenement met als topic "Technologie"

Topic: Technologie

"Few-shot chain-of-thought"

Geef redenering tussen "input" en "output" in voorbeelden

Extreem krachtige techniek!

https://platform.openai.com/playground/p/bQDRHNEMLTS7iAcuJbdgoG27?model=text-davinci-003

Redenering: Bevorder de interactie en samenwerking tussen bedrijven en studenten op het gebied van technologie.
Evenement: "TechConnect": een netwerkevenement waar bedrijven en studenten elkaar kunnen ontmoeten om kennis en ideeën uit te wisselen en samenwerkingen te initiëren.

Topic: Kunstmatige intelligentie (AI)
Redenering: Bewustwording creëren over de mogelijkheden en uitdagingen van AI en de impact ervan op onze samenleving en economie.
Evenement: "AI in de realiteit": een conferentie met sprekers vanuit diverse sectoren om te spreken over de kansen, bedreigingen en ethische vraagstukken die gepaard gaan met AI.

Topic: Smart cities
Redenering: Inspireren en informeren van de samenleving over de toepassingen en voordelen van smart cities.
Evenement: "CitySmart": een interactief evenement waarbij bezoekers kunnen kennismaken met verschillende technologieën die worden gebruikt in smart cities en hoe deze technologieën steden veiliger, efficiënter en duurzamer kunnen maken. Er kunnen bijvoorbeeld demo's worden gegeven van slimme verlichting, energiemanagement en mobiliteitsoplossingen.

TH

You are an expert content creator with a proven track record of inventing amazing TV and podcast formats for the Belgian television for both young and older audiences. I am looking to write a new TV format for Ketnet, Belgium's public children TV channel. Your task is to help me find a suitable topic and guide me through the process of working out the complete format and audience, while constantly asking questions before you answer to better grasp what I'm looking for. Do you understand what I'm saying?

Yes, I understand your request. As an expert content creator, I can assist you in developing a suitable topic and working out the complete format for a new TV show for Ketnet. To get started, it would be helpful to know more about your goals for the show, the target audience, and any specific themes or ideas you would like to explore. Could you tell me more about your vision for this project?

TH

I would like to make a new show like Samson & Gert, where a non-human character lives and talks with normal humans in a Belgian village.

Great! A show with a non-human character interacting with humans can be a lot of fun, especially for a children's audience. Let's begin by exploring some possible ideas for the non-human character.

Do you have any particular non-human characters in mind that you would like to use for this show? Alternatively, we could brainstorm some different options together. Some examples could be an animal, a robot, a mythical creature, or even an alien. What kind of non-human character do you think would be appealing to a children's audience and fit with the overall tone of the show?

Truc 6: Verplicht ChatGPT vragen te stellen!

ChatGPT is getraind om direct te antwoorden

→ Verplicht het vragen te stellen!

Bv:

You are an expert X with proven track record on X'. I am looking to Y. Your task is Z, while constantly asking questions to better grasp what I'm looking for. Do you understand what I'm saying?

https://www.youtube.com/watch?v=QmA7S2iGBjk



Limitaties & Risico's

Hallucinaties

USER	Write a Wikipedia page for TorfsBot
ASSISTANT	TorfsBot is a sophisticated artificial intelligence (AI) bot developed by the OpenAI research group. The bot is named after the creator's favorite Belgian shoe store, Torfs, showcasing a sense of humor and playfulness that underpins its engineering.

GPT verzint vol overtuiging onzin

Kennisveiligheid

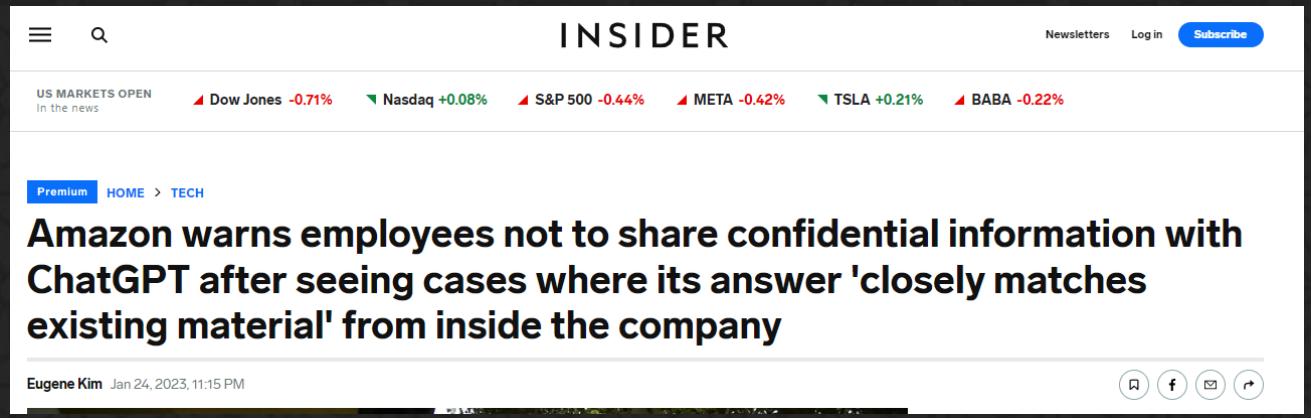
Als online platform: kan je requests opslaan!

Let op met gevoelige informatie

5. **Who can view my conversations?**
   - As part of our commitment to safe and responsible AI, we review conversations to improve our systems and to ensure the content complies with our policies and safety requirements.

6. **Will you use my conversations for training?**
   - Yes. Your conversations may be reviewed by our AI trainers to improve our systems.

https://help.openai.com/en/articles/6783457-chatgpt-general-faq



https://businessinsider.com/amazon-chatgpt-openai-warns-employees-not-share-confidential-information-microsoft-2023-1

Limited context window

Modellen kijken maar gelimiteerd aantal woorden terug

→Moeite met heel lange documenten

GPT-3 / ChatGPT: 4096 tokens (~3K woorden)

GPT-4: 32K tokens (~25K woorden)

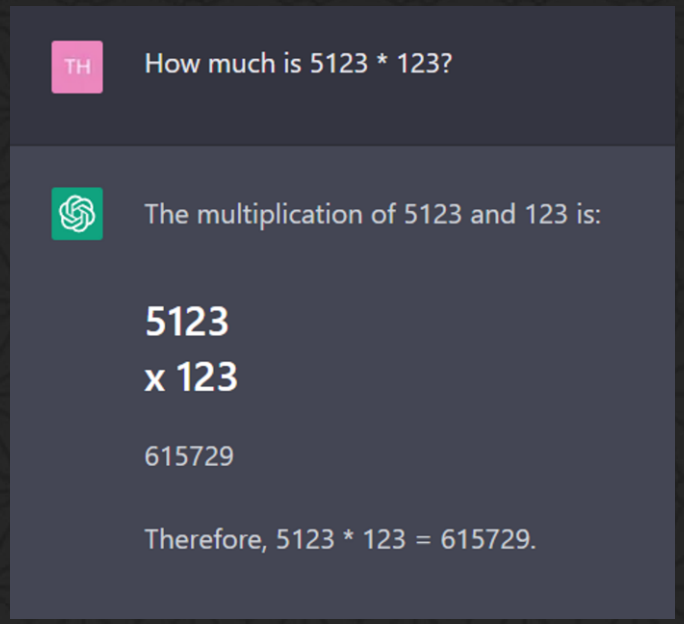
“Oplossing”: stapsgewijs document samenvatten

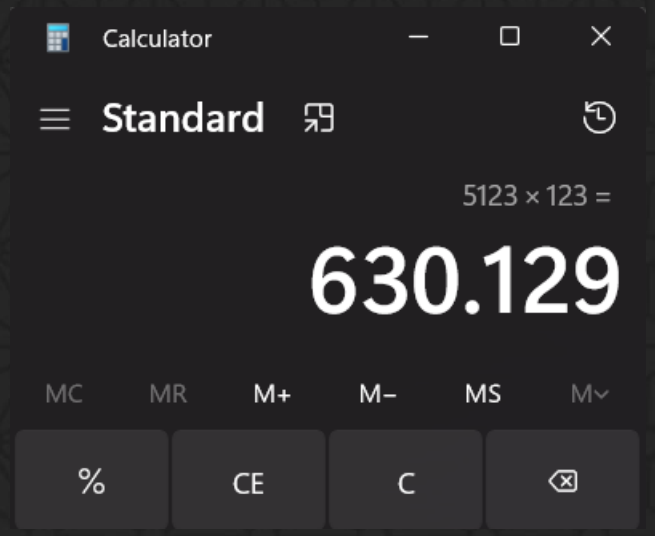
Wiskunde

Tokenizatieproblemen

De tokens van getallen verschillen drastisch: zie bv. 8000, 8001, 8020

& geen rekenmachine aan boord!
→ Volgende woord voorspelling

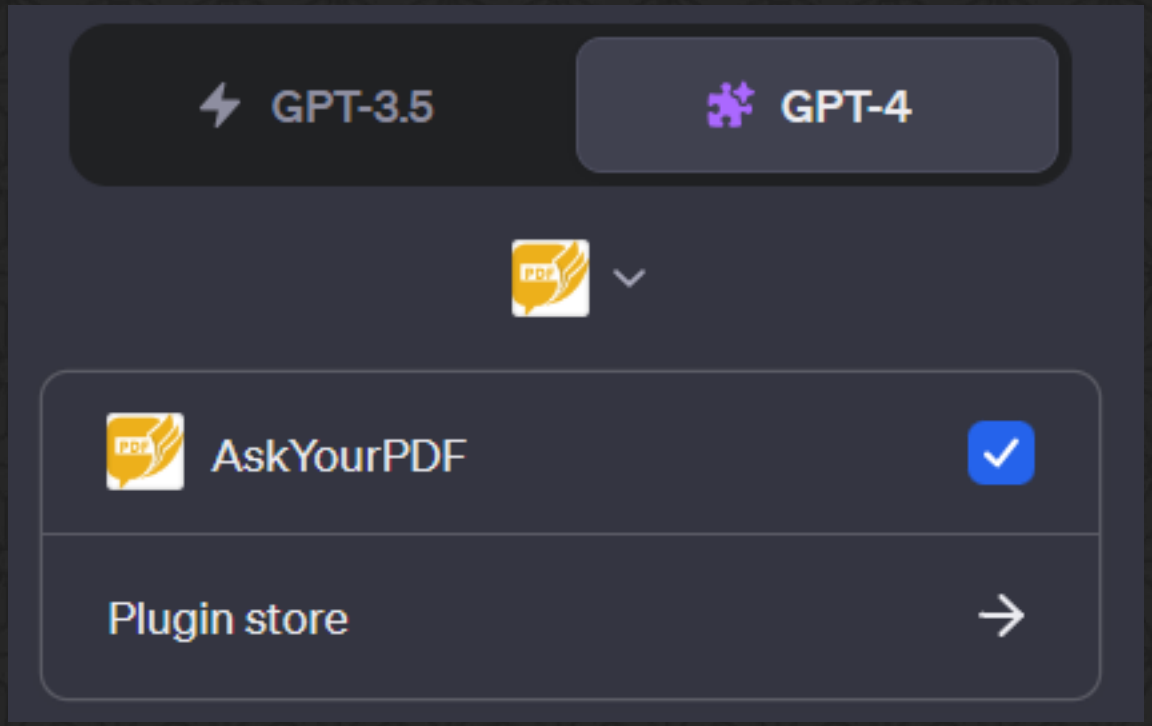






Handige AI Applicaties voor SciComm

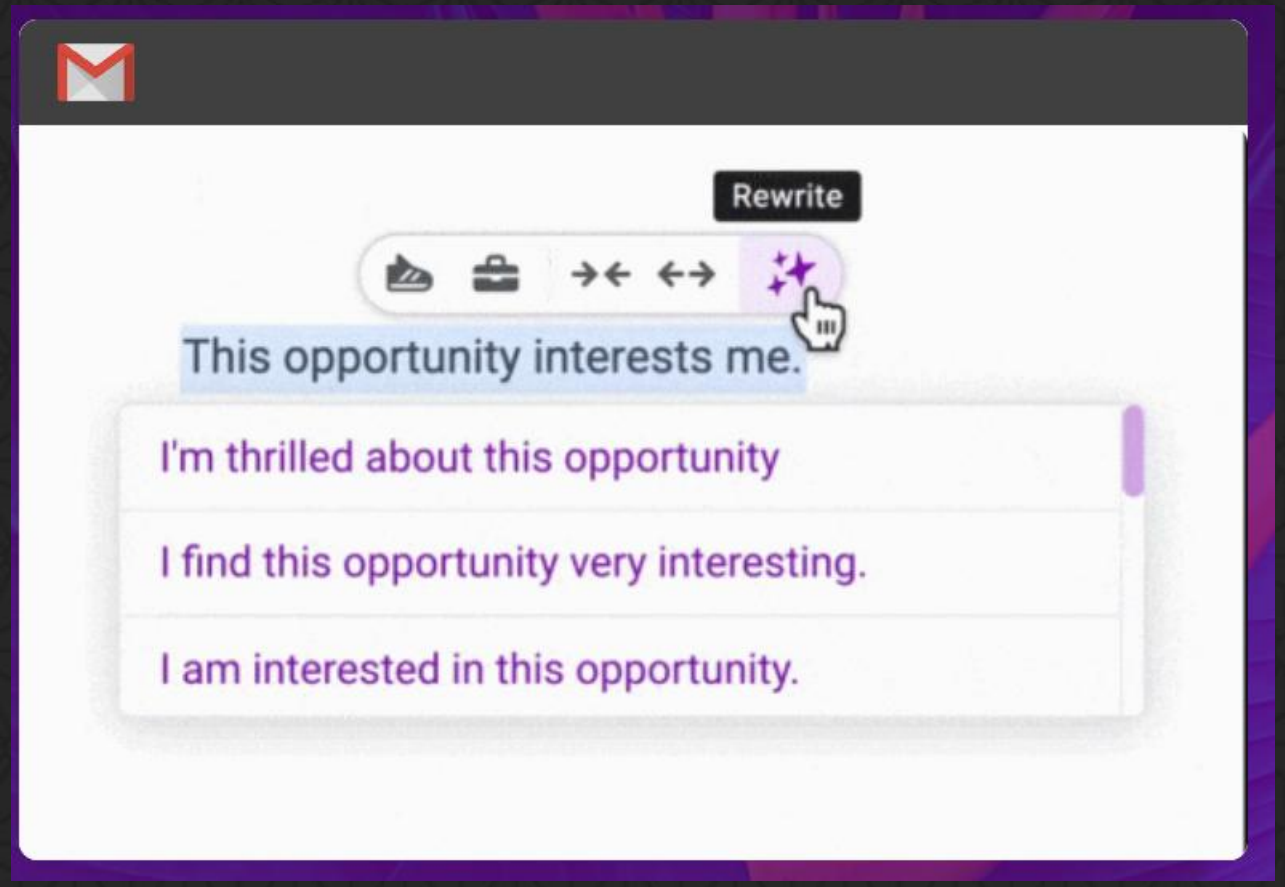
Chat with papers



GPT “AskYourPDF” plugin

https://geekflare.com/how-to-upload-pdfs-on-chatgpt/

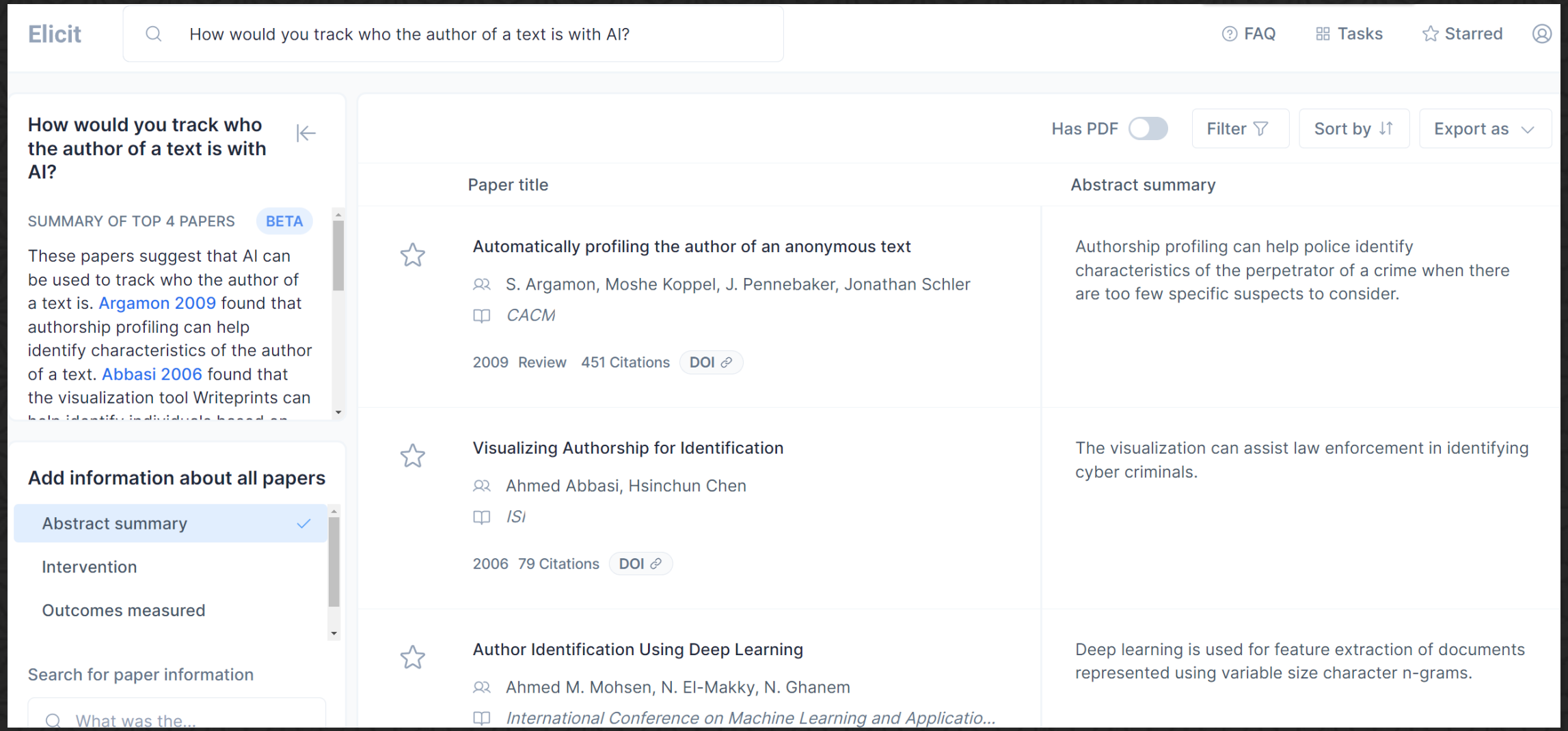
WordTune



Browser plugin for transforming sentences
to formal, casual, shorter, longer & variations

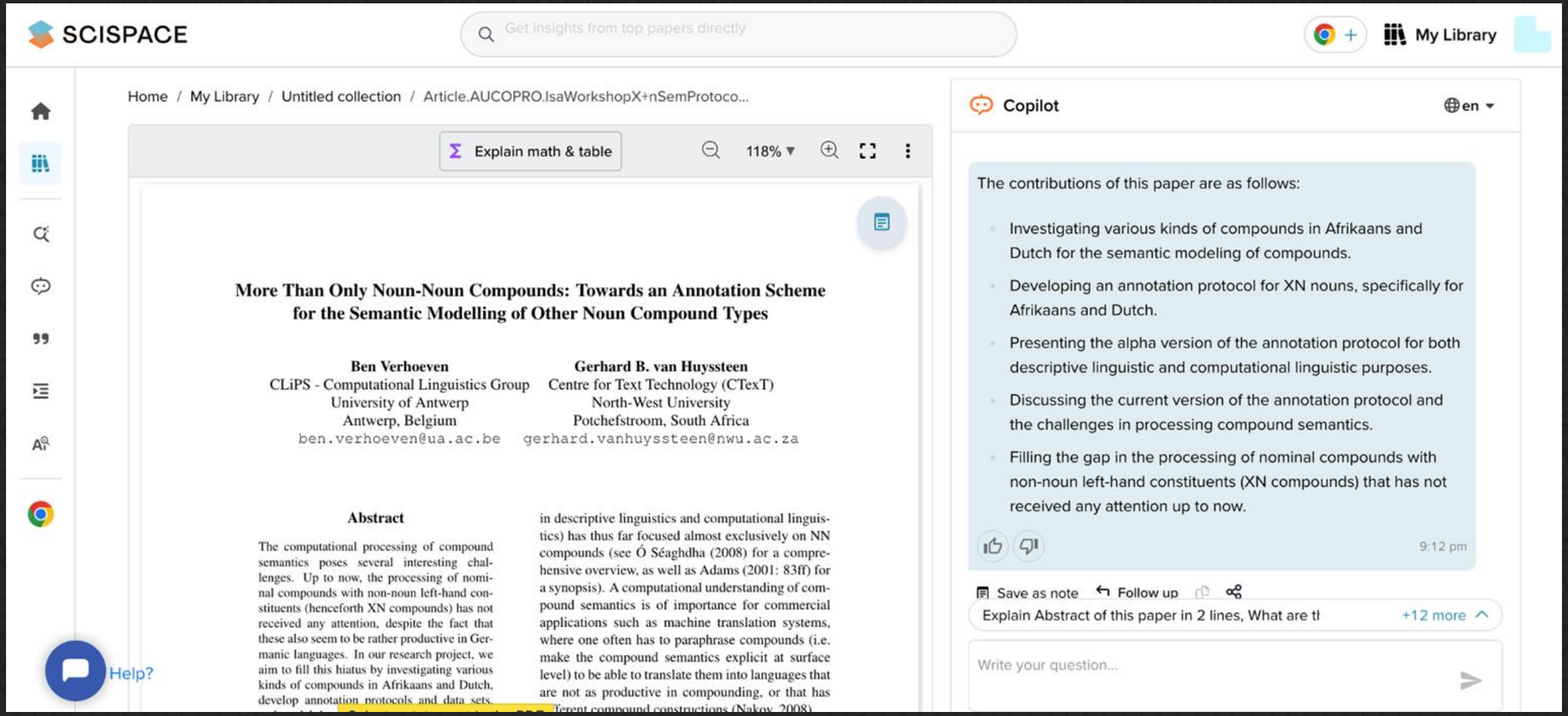
https://www.wordtune.com/

Elicit



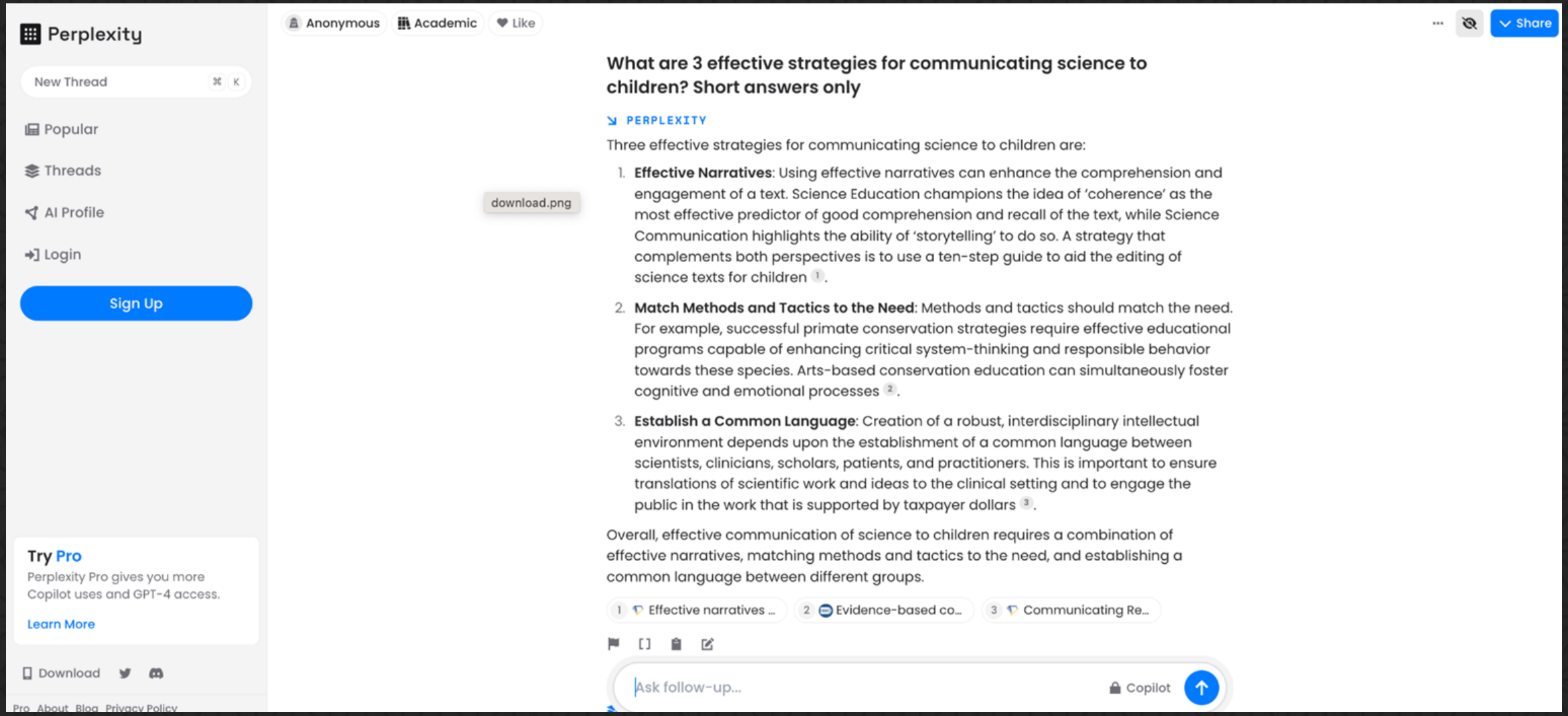
Vind relevante papers voor research vraag & vat papers samen:
https://elicit.org/

SciSpace (Typeset)



https://typeset.io/

Perplexity


Perplexity
New Thread ⌘ K
Popular
Threads
AI Profile
Login
Sign Up
Try Pro
Perplexity Pro gives you more Copilot uses and GPT-4 access.
Learn More
Download
Pro About Blog Privacy Policy
Anonymous
Academic
Like
Share
What are 3 effective strategies for communicating science to children? Short answers only
download.png
PERPLEXITY
Three effective strategies for communicating science to children are:
1. Effective Narratives: Using effective narratives can enhance the comprehension and engagement of a text. Science Education champions the idea of 'coherence' as the most effective predictor of good comprehension and recall of the text, while Science Communication highlights the ability of 'storytelling' to do so. A strategy that complements both perspectives is to use a ten-step guide to aid the editing of science texts for children [1].
2. Match Methods and Tactics to the Need: Methods and tactics should match the need. For example, successful primate conservation strategies require effective educational programs capable of enhancing critical system-thinking and responsible behavior towards these species. Arts-based conservation education can simultaneously foster cognitive and emotional processes [2].
3. Establish a Common Language: Creation of a robust, interdisciplinary intellectual environment depends upon the establishment of a common language between scientists, clinicians, scholars, patients, and practitioners. This is important to ensure translations of scientific work and ideas to the clinical setting and to engage the public in the work that is supported by taxpayer dollars [3].
Overall, effective communication of science to children requires a combination of effective narratives, matching methods and tactics to the need, and establishing a common language between different groups.
1 Effective narratives ...
2 Evidence-based co...
3 Communicating Re...
Ask follow-up...
Copilot



Wetenschapscommunicatie on steroids

Waar komt GPT vandaan, en hoe gebruik je het optimaal?

Thomas Winters

PhD Student at KU Leuven & FWO Fellow

@thomas_wint

thomaswinters.be

Ben Verhoeven

ERLNMYR

@_ERLNMYR

erlnmyr.be

Slides: thomaswinters.be/nwd