

ChatGPT en wetenschap hoe zit dat eigenlijk?

Een beschouwing voor het voorbereidend wetenschappelijk onderwijs



Bob Siegerink, PhD

b.siegerink@lumc.nl | @bobsiegerink | bobsiegerink.com

Enkele scenario's

Wat denken jullie?



OK

NIET OK



50% van de text is direct overgenomen vanuit chatGPT

- Een **profielwerkstuk** van een **VWO** scholier
- Een **bachelor scriptie** van een **geneeskunde** student
- Een **wetenschappelijke publicatie** van een **professor scheikunde**
- Een **wetenschappelijke publicatie** van een **professor filosofie**



Wat is wetenschap?

Wat maakt wetenschap anders dan een ander werkveld?



Wat doen wetenschappers eigenlijk?
Wat is het doel van wetenschap?
Waarom is wetenschappelijke fraude een groot probleem?



Wat doen wetenschappers eigenlijk?
Wat is het doel van wetenschap?
Waarom is wetenschappelijke fraude een groot probleem?

**Wetenschap is namens ons allen op zoek naar meer inzicht in
de wereld**

Waarheid
Opinie vs feiten
Controleerbaarheid
(on)zekerheid



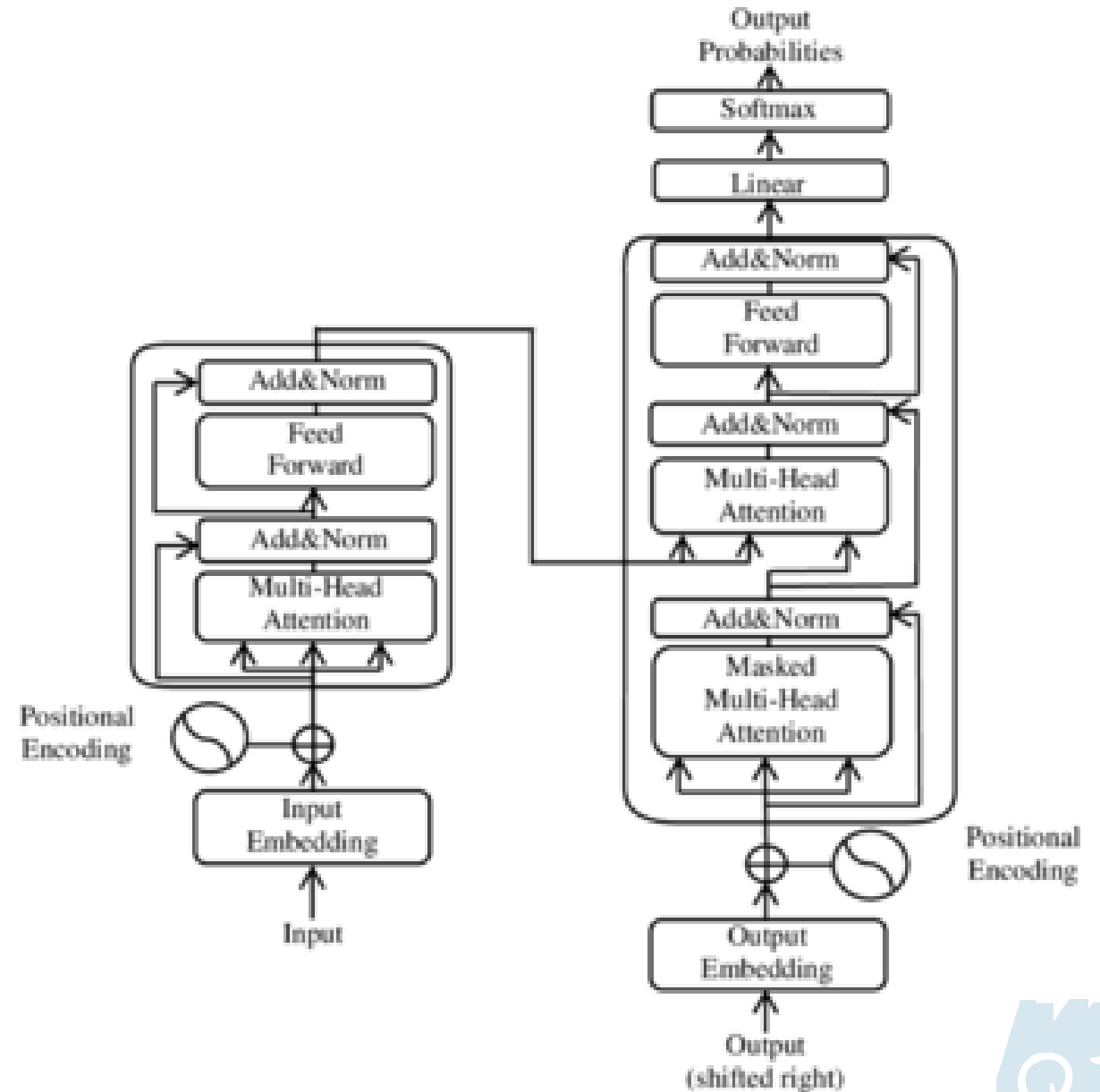
Wat is ChatGPT?

Zonder al te technisch te worden, wat zijn de concepten die achter ChatGPT steken?



Generative Artificial Intelligence voor tekst

- Large Language Model (LLM)
- Software schrijft een tekst die te interpreteren is voor een mens
 - Generatief Pretrained Transformer
 - Niet “hard-coded”, maar stochastisch
 - Taal, stijl grammatica etc zijn een functie van reeks aan kansen/parameters, gebaseerd op de teksten die de basis vormen van het taalmodel
- garbage in, garbage out



Kansen *extraheren* vanuit teksten van het internet -> “vooringenomenheid”



Midjourney: a robot sitting at a desk, well lit, correcting a printed text with a red / green pen with many scribbles



Kansen *extrapoleren* vanuit teksten van het internet -> “hallicunaties”



schrijf een korte biografische schets van bob siegerink. gebruik 200 woorden, wat zijn achtergrond is, en het onderwerp van zijn onderzoek.



Bob Siegerink is een vooraanstaand onderzoeker op het gebied van bloedstolling en trombose. Hij werd geboren in Nederland en ontwikkelde al op jonge leeftijd een passie voor wetenschap en medisch onderzoek. Na het behalen van zijn bachelor in de biomedische wetenschappen aan de Universiteit Utrecht, voltooide hij zijn promotieonderzoek aan dezelfde universiteit.

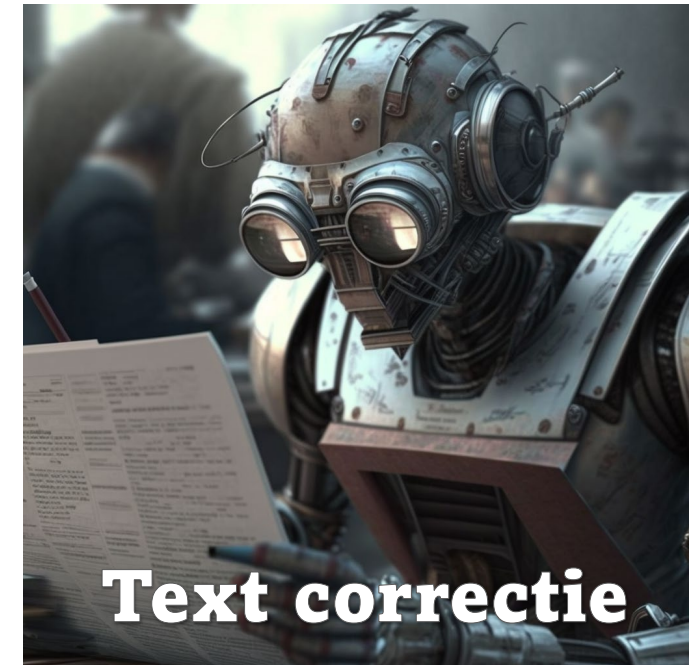
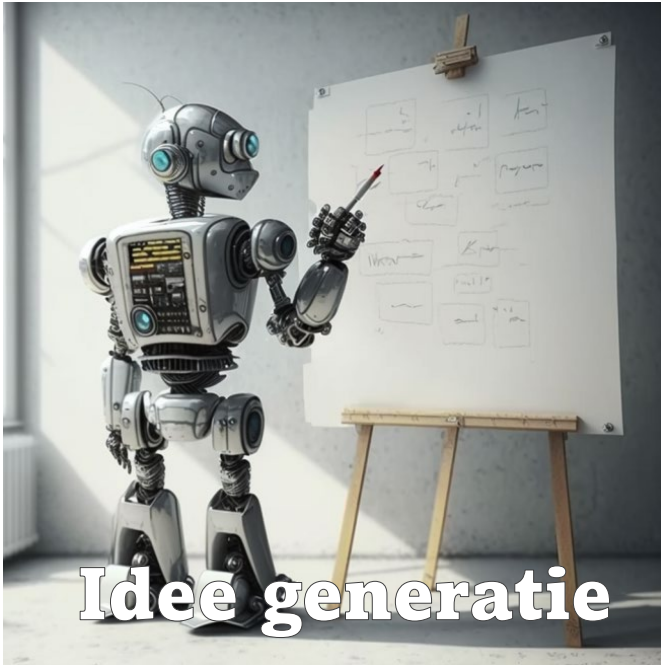
Siegerink's onderzoek richt zich op het begrijpen van de mechanismen achter bloedstolling en de ontwikkeling van trombose. Hij heeft zich gespecialiseerd in het identificeren van risicofactoren en genen die betrokken zijn bij trombotische gebeurtenissen, zoals diepe veneuze trombose en

 Regenerate response

Send a message. 



Welke taalfuncties kan chatGPT allemaal voor je vervullen?



wat zijn de verschillen?
Is dat anders voor Wetenschap?

ChatGPT en wetenschap

Is het nu wel of niet een goed idee?



Large language models zoals ChatGPT en de wetenschap

- **Goed of niet? Of maakt het niet zoveel uit?**
- **Waarom?**



Large language models zoals ChatGPT en de wetenschap

- **Wat is de rol van tekst in de Wetenschap?**
 - Is de tekst de kern van de boodschap van een wetenschappelijke publicatie?
- **Auteurschap** gaat over of je “accountable” bent voor je werk
- **Plagiaat** is wanneer je intellectueel werk van een ander voordoeft als de jouwe

Het gebruik van ChatGPT in wetenschappelijke publicaties hoeft niet verkeerd te zijn

- Transparantie over gebruik
- Verhogen leesbaarheid Engels als tweede taal
- Minimaliseren fouten in computercode



Detectie is niet optimaal

Dus veel mogelijkheden voor vragen die ook jullie kunnen beantwoorden



Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers

 Catherine A. Gao,  Frederick M. Howard,  Nikolay S. Markov,  Emma C. Dyer,  Siddhi Ramesh,  Yuan Luo,  Alexander T. Pearson

doi: <https://doi.org/10.1101/2022.12.23.521610>

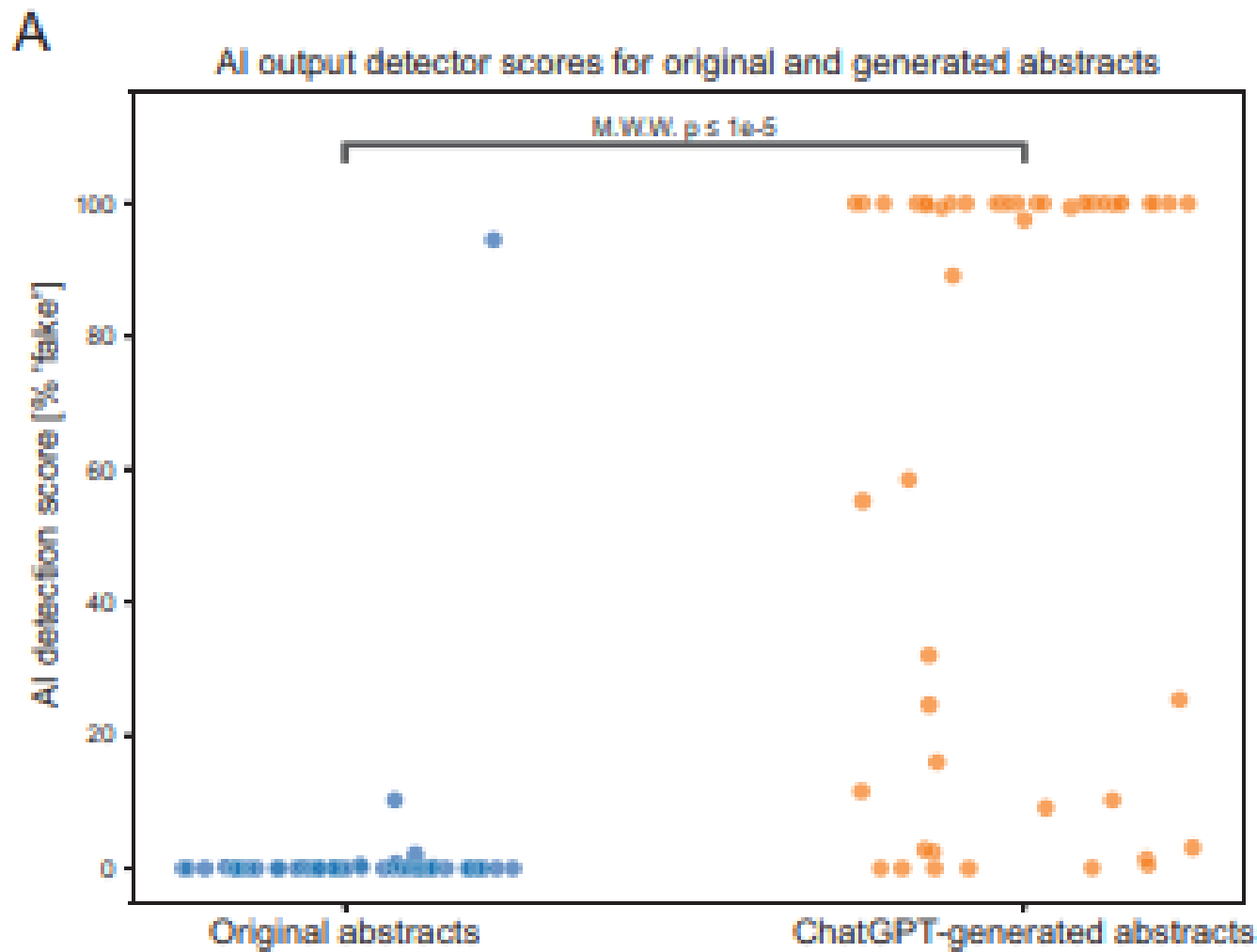


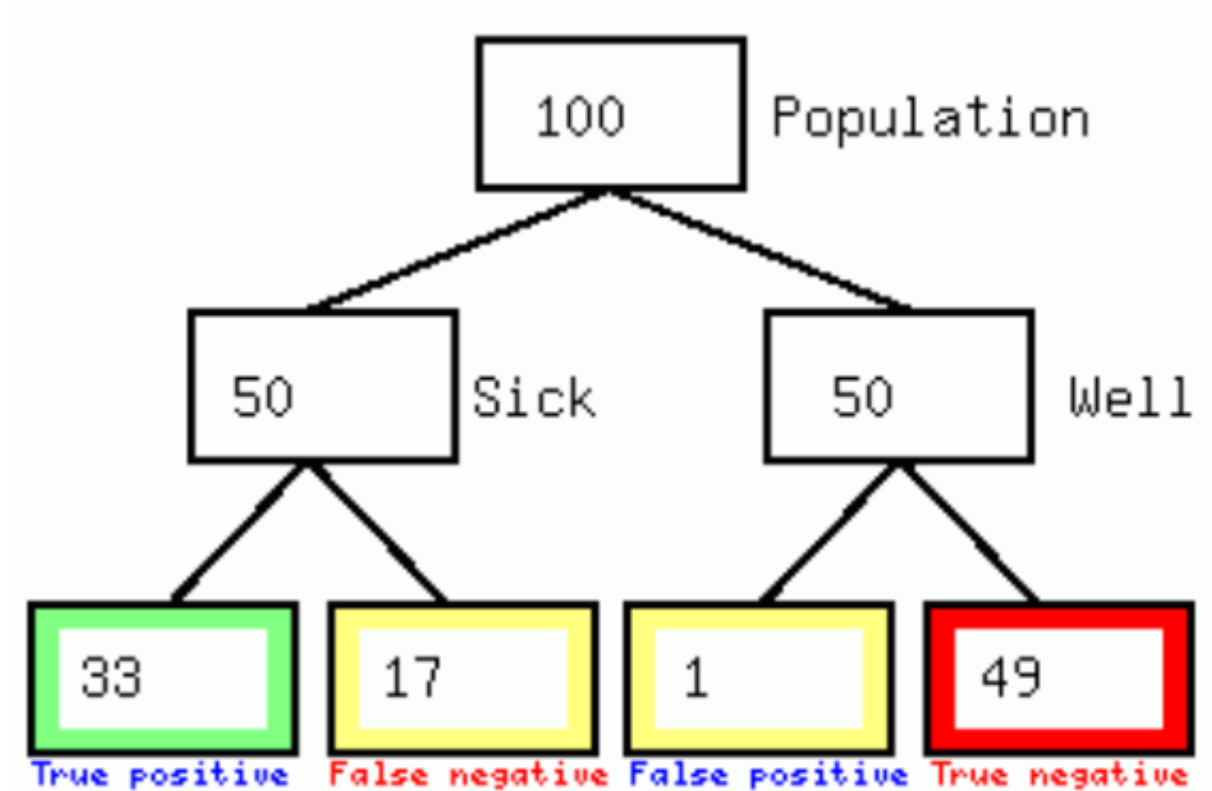
Figure 2. Many generated abstracts can be detected using an AI output detector. (A) AI detection scores as [% 'fake'] per GPT-2 Output Detector for original abstracts and generated abstracts. Except for one case, all original abstracts scored extremely low on the AI output detector. The majority of generated abstracts scored high on the AI output detector, but 17 (34%) scored lower than 50%. **(B)** The AI output detector ROC curve for discriminating between original and generated abstracts, with AUROC of 0.94.



Eigenlijk is detectie niets anders dan “diagnose” binnen de geneeskunde

Results

Statistic	Value	95% CI
Sensitivity	66.00%	51.23% to 78.79%
Specificity	98.00%	89.35% to 99.95%
Positive Likelihood Ratio	33.00	4.69 to 232.06
Negative Likelihood Ratio	0.35	0.24 to 0.51
Disease prevalence (*)	50.00%	
Positive Predictive Value (*)	97.06%	82.43% to 99.57%
Negative Predictive Value (*)	74.24%	66.16% to 80.95%
Accuracy (*)	82.00%	73.05% to 88.97%



Onderzoeksideeen: varieer en vergelijk

- De doelgroep: docenten, medescholieren, opa's en oma's etc
- De taal en stijl van de tekst
- Het onderwerp van de tekst (geschiedenis, natuurkunde, media)
- Niveau van “zelf herschrijven”
- Verschillende detective algoritmes
- “vooringenomenheid” en “hallicunaties”



Enkele waarschuwingen

gesloten, biased blackbox, dat data verzameld voor commerciële doeleinden





Posted by u/nitontonii 9 months ago



279

OpenAI is not open.



[Other]

Normally, projects with "open" in their name tend to refer that their information will be transparent, usually non-profits, especially within computer science, very often used for open-source programs.

OpenAI has the right to pick the name that they want, but it's kinda misleading for the community.

They are very clear when they call themselves a company:
"OpenAI is an AI research and deployment company. Our mission is to ensure that artificial general intelligence benefits all of humanity. "

According to them, a kind of "ethical oriented company". Although it's hard to find a company that doesn't present itself as a "benefit for humanity".

Do not get confused by their name, OpenAI doesn't want to be like open-source projects, they haven't allowed free access to GPT, DALL-E, or any other software. They are a company with profit motives, even the domain of the website is ".com" for commercial.



86 Comments



Share



Save



Hide



Report

95% Upvoted



ChatGPT

This is a free research preview.



Our goal is to get external feedback in order to improve our systems and make them safer.



While we have safeguards in place, the system may occasionally generate incorrect or misleading information and produce offensive or biased content. It is not intended to give advice.

Next

ChatGPT

How we collect data



Conversations may be reviewed by our AI trainers to improve our systems.



Please don't share any sensitive information in your conversations.

Back

Next

Wrap up



Wrap up

- ChatGPT is vooral een probleem als geschreven tekst door person zelf geschreven MOET zijn.
 - Gek genoeg hoeft dat niet perse in de wetenschap
 - Maar dat is wel de essentie van toetsen in het Onderwijs
 - Deeloplossing is om niet het eindproduct, maar het process te beoordelen
- Verschillende toepassingen: idee generatie, text generatie and text correctie
 - Doel van een tekst en manier van gebruik LLM
- Veel onderzoeksmogelijkheden!
 - Detectie is niet optimaal
 - Wat zijn de standpunten van vershcillende groepen t.o.v. chatGPT in verschillende scenarios?
- Pas op voor aspecten als “gesloten, commerciele, blackbox”

Terug naar de scenarios - denk je er nu anders over?

- Een profielwerkstuk van een VWO student bestaat uit 50% uit teksten die direct zijn overgenomen uit ChatGPT.
- Een bachelor scriptie van een geneeskunde student bestaat voor 50% uit teksten die direct zijn overgenomen uit chatGPT
- Een wetenschappelijke publicatie professor scheikunde bestaat voor 50% uit teksten die direct zijn overgenomen uit chatGPT
- Een wetenschappelijke publicatie professor filosofie bestaat voor 50% uit teksten die direct zijn overgenomen uit chatGPT





Quality and Integrity of Science

seminars | lectures | journal club | policy | research | information
qi-leiden.com