



Een (menselijke) chat over ChatGPT

Thomas van Osch

Machine Learning Advisor bij SURF

thomas.vanosch@surf.nl

16 januari 2023

| Wat is ChatGPT?

Op 30 november gelanceerd door OpenAI



Gebaseerd op InstructGPT (ook wel GPT-3.5 genoemd)



Getraind op een enorme hoeveelheid data



ChatGPT



Interactief taalmodel dat met natural language processing tekst genereert



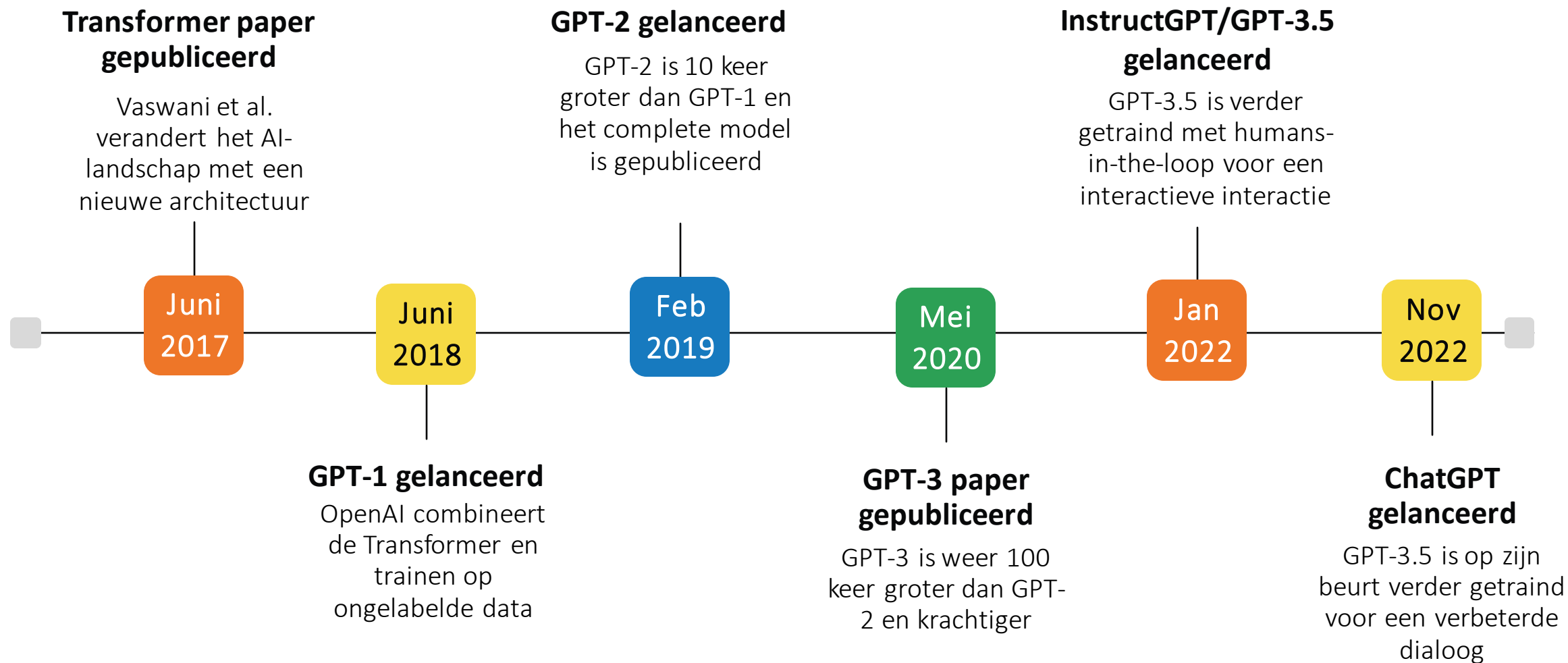
Goed in het genereren van context-relevante tekst



Momenteel publiekelijk toegankelijk

Timeline

Op 30 november gelanceerd
door OpenAI

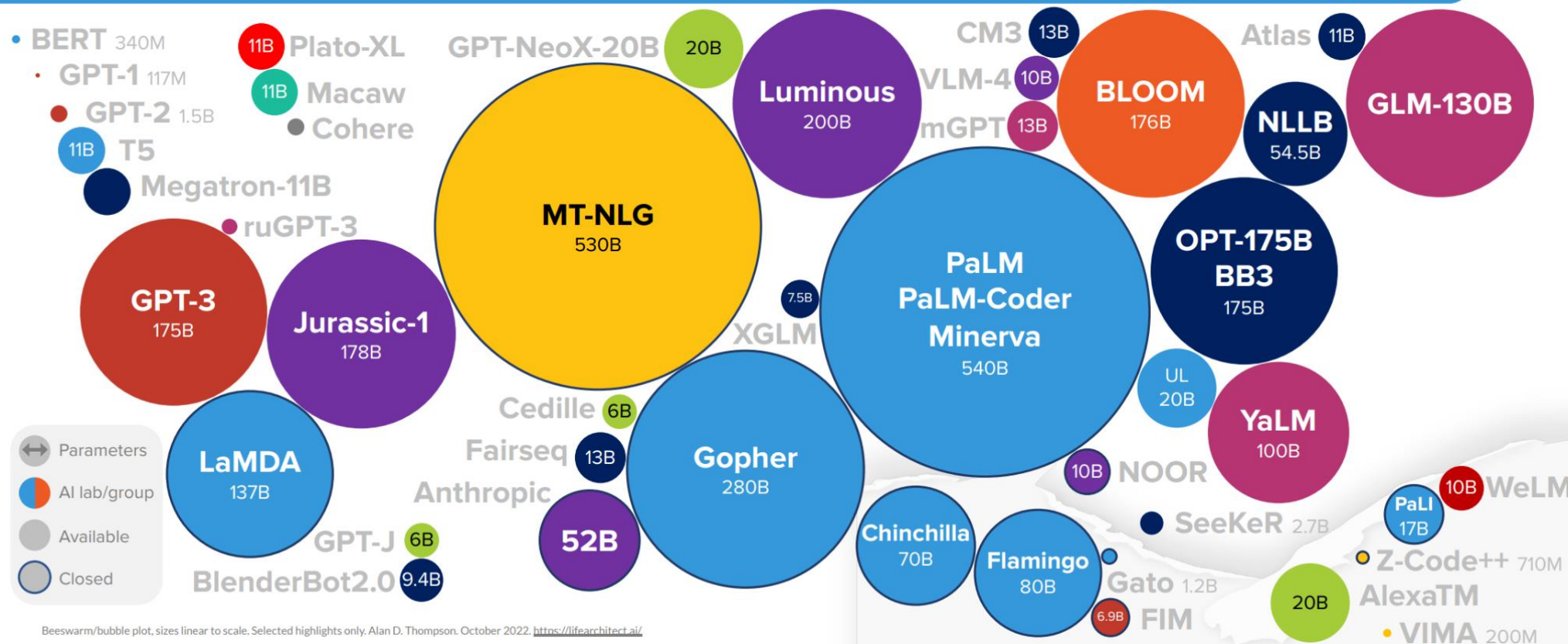


Wat is ChatGPT?



Getraind op een enorme hoeveelheid data

LANGUAGE MODEL SIZES TO OCT/2022

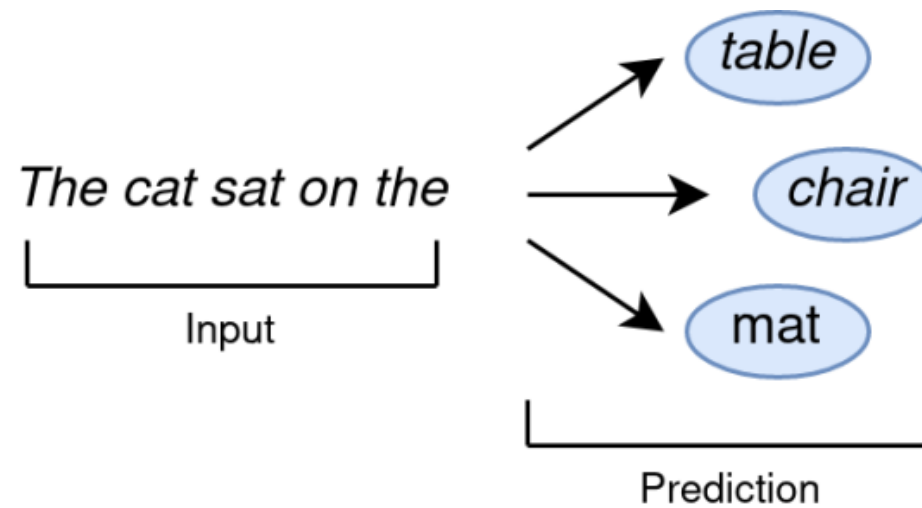


| Wat is ChatGPT

Gebaseerd op InstructGPT
(ook wel GPT-3.5 genoemd)



- ChatGPT is een chatbot-achtige taalmodel (large language model)



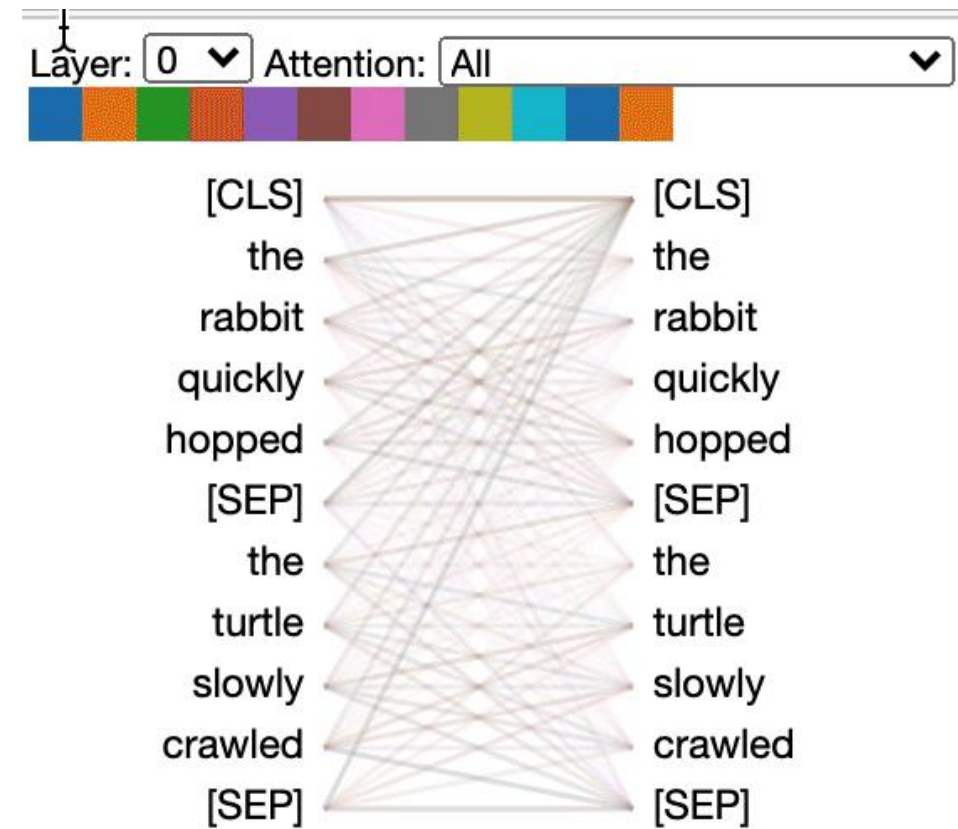
De basis van language modeling is het voorspellen van een nieuw woord in een gegeven zin of woordgroep.

| Wat is ChatGPT

Gebaseerd op InstructGPT
(ook wel GPT-3.5 genoemd)



- ChatGPT is een chatbot-achtige taalmodel (large language model)
- Variant van GPT-3
 - Transformer-architectuur
 - Gigantische hoeveelheid (internet) data



Bron: <https://github.com/jessevig/bertviz>

| Wat is ChatGPT

Gebaseerd op InstructGPT
(ook wel GPT-3.5 genoemd)



- ChatGPT is een chatbot-achtige taalmodel (large language model)
- Variant van GPT-3
 - Transformer-architectuur
 - Gigantische hoeveelheid (internet) data
- Geoptimaliseerd voor conversatie
 - Training met feedback van mensen



Om met een analogie RLHF uit te leggen: stel je hebt een robot genaamd Rob. Rob wil graag een taal leren en heeft een zogenaamde language model ter beschikking. Hiermee kan hij zinnen genereren en begrijpen zoals "ik wordt later beroemd". Een mens gaat vervolgens feedback geven over wat het model van Rob gegenereerd heeft en zal Rob eventueel corrigeren. Deze feedback gaat Rob vervolgens doorgeven aan zijn eigen model die weer de parameters updatet. Met vele iteraties zorgt RLHF er voor dat de interactie van de language model via Rob op een natuurlijke en menselijke manier gaat.

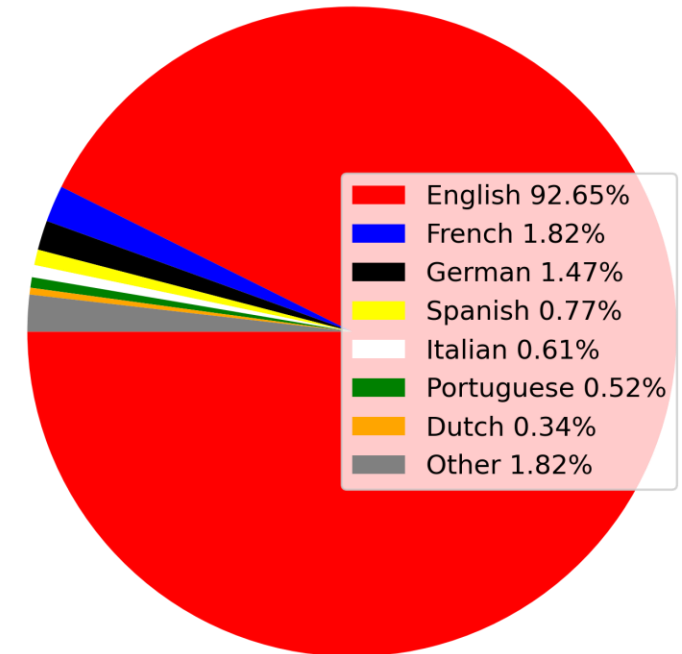
| Wat kun je met ChatGPT?

- Antwoorden genereren op (complexe) vragen
- Brainstormen/pointers laten geven over trefwoorden of een onderwerp
- Automatisch samenvatten
- Korte essays laten schrijven
- Taalfouten corrigeren



En dit in verschillende talen.
Ook programmeertalen!

Languages in GPT-3 training data



| ChatGPT kanttekeningen

- Geen begrip van semantische of epistemische kennis
- Feitelijke correctheid
- Geen vervangende gesprekspartner
- Ethische kwesties
 - Vooringenomenheid (bias)
 - Intransparant
- Frauduleus gebruik



Hoe zal Nederland het doen op het WK van 2026



Het is moeilijk om te voorspellen hoe Nederland zal presteren op het WK van 2026, omdat er nog veel kan veranderen in de komende jaren in termen van spelers, coaches en de algemene voorbereiding van het team.  

De Nederlandse voetbalploeg heeft een rijke geschiedenis met veel succes, inclusief een wereldkampioenschap in 1988 en een eindnotering als tweede in 2010. De afgelopen jaren heeft het team echter wat ups en downs gehad.

Misvattingen

- "Plagiaatdetectiesoftware ziet het verschil wel"
 - GPTZero, huggingface
- "Mensen kunnen zelf wel gegenereerde tekst onderscheiden"
- "Watermerk zal gegenereerde content identificeren"

Input text below:

ChatGPT is een taal model dat is gebouwd met de GPT (Generative Pre-training Transformer) architectuur. GPT is ontworpen om een breed scala aan taaltaken uit te voeren, zoals het voorspellen van tekst, het genereren van tekst, het begrijpen van tekst, het classificeren van tekst, en nog veel meer.

ChatGPT is specifiek getraind op conversatie-dataset, dus het is speciaal geschikt voor conversatie-gerelateerde taken. Hieronder zijn enkele voorbeelden van wat ChatGPT kan doen:

Chatbot development: ChatGPT kan worden gebruikt om chatbots te ontwikkelen die natuurlijke en menselijke conversaties kunnen hebben.

Generate automated dialog : ChatGPT kan worden gebruikt om geautomatiseerde dialogen te genereren voor virtuele assistenten, zoals Siri of Alexa.

Language Translation: Met fine-tuning op specifieke taaltaken, ChatGPT kan worden gebruikt voor machinale vertaling.

Text summarization: ChatGPT kan worden gebruikt om samenvattingen te maken van lange stukken tekst.

Language Correction: ChatGPT kan worden gebruikt voor taalcorrigerende taken zoals grammaticafouten corrigeren, woordkeuze verbeteren.

Het is belangrijk om te onthouden dat GPT-modellen, zoals ChatGPT, worden aangestuurd door hun trainingsgegevens, dus hoe goed een bepaalde taak wordt uitgevoerd, is afhankelijk van hoe goed de modellen zijn getraind op de specifieke taak.

Bovendien, ChatGPT is een "gegenereerd" model, wat betekent dat het gebruik maakt van de informatie in zijn trainingsset om een output te geven, dus de output kan beperkt zijn tot de informatie die beschikbaar is in zijn trainingsset, en kan niet altijd juist zijn of geschikt voor bepaalde situaties.

1681/50006

GPTZero has finished analyzing your text!

Get GPTZero Result

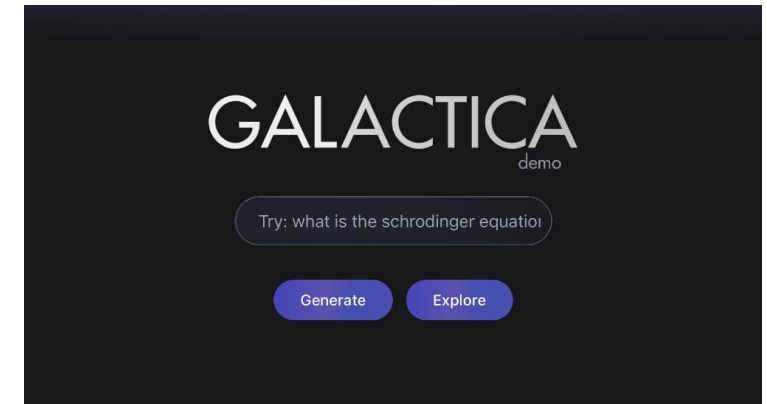
Your GPTZero score corresponds to the likelihood of the text being AI generated:

291.1564871336374

Your text is likely human generated!

Misvattingen

- "Plagiaatdetectiesoftware ziet het verschil wel"
 - GPTZero, huggingface
- "Mensen kunnen zelf wel gegenereerde tekst onderscheiden"
- "Watermerk zal gegenereerde content identificeren"
- "ChatGPT heeft geen weet van iets na September 2021"
- "ChatGPT produceert geen bronnen of citaten"
 - Galactica



REAGEREN



**Eerste reactie is het gebruik
van dergelijke modellen
verminderen en/of
verbieden. Maar hoe? En
voor hoe lang?**

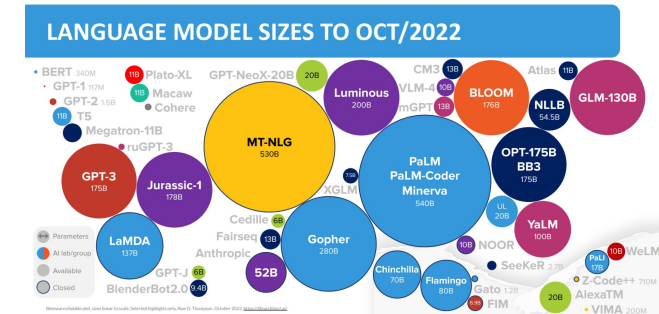
ADAPTEREN



**Hoe accepteren we dergelijke
modellen en zetten we ze in
als hulpmiddelen in plaats
van ze te bestrijden.**

Hoe verder?

- OpenAI & co blijven betere language models ontwikkelen (GPT-4, Google)
 - Maar ook onderzoek naar detectie (Turnitin, OpenAI)
- Stof van toetsing aanpassen (specifiek, persoonlijk, recent)
- Type toetsing aanpassen (mondeling, op locatie, proctoring)
- ChatGPT opnemen in het curriculum



| Wat is de rol van SURF

Technische support

SURF levert technische expertise over AI en ChatGPT aan haar leden

Platform voor discussies

SURF biedt een platform aan voor onderlinge kennisuitwisseling



Informatiedagen

SURF host informatiedagen over ChatGPT, maar ook andere onderwerpen.

Leren van de experts

SURF leert graag van haar leden om zelf ook up-to-date te blijven met de ontwikkelingen op het gebied van AI

| Discussiepunten

- Waar is de balans tussen bestrijden en benutten
- In hoeverre is het gebruik van ChatGPT toegestaan
 - Brainstorm, controleren, literatuuronderzoek
- Hoe kunnen we ChatGPT inzetten als 'tool'
- Zal ChatGPT impact hebben op de vaardigheden van studenten
 - Schrijfvaardigheid
 - Creativiteit
- Of is het toch een overschatte hype
 - Spellingscontrole (jaren 80)
 - Google/Wikipedia (jaren 00)





**Voor meer
informatie:**

✉ AI@surf.nl

🌐 www.surf.nl

SURF