



INSTITUTE FOR LANGUAGE SCIENCES 28-05-2024

De inzet en bruikbaarheid van AI en leesbaarheidstools

Antal van den Bosch, met Ted Sanders

Heerlijk Duidelijkdag, Den Bosch, 21 mei 2024

1

Automatisch bepalen van leesbaarheid van teksten: geschiedenis

Leesgemak [bewerken | brontekst bewerken]

De Amerikaanse schrijfadviseur **Rudolf Flesch** presenteerde in *The art of readable writing* in 1949 een formule waarbij het leesgemak gelijk is aan de volgende formule:

$$206,835 - 1,015 \left(\frac{\text{totaal aantal woorden}}{\text{totaal aantal zinnen}} \right) - 84,6 \left(\frac{\text{totaal aantal lettergrepen}}{\text{totaal aantal woorden}} \right)^{11}$$

De landbouwsocioloog W.H. Douma paste de formule aan naar het Nederlands. De Flesch-Doumaformule stelt het leesgemak gelijk aan de volgende formule:

$$206,835 - 0,93 \left(\frac{\text{totaal aantal woorden}}{\text{totaal aantal zinnen}} \right) - 77 \left(\frac{\text{totaal aantal lettergrepen}}{\text{totaal aantal woorden}} \right)$$

De score verwijst naar de opleidingsniveau:

- 0-30: academici
- 30-50: studenten
- 50-60: hoger middelbaar onderwijs
- 60-70: lager middelbaar onderwijs
- 70-80: groep 8 basisschool
- 80-90: groep 7 basisschool
- 90-100: groep 6 basisschool

Utrecht University Wikipedia, "Leesbaarheid" 28-05-2024 2

2

Dus: geen lange zinnen en geen moeilijke woorden?

"Sla een willekeurige pagina op en het zal u direct opvallen: in [xxx] staan geen lange zinnen of moeilijke woorden. Het taalgebruik is voor 100% afgestemd op de vmbo-leerling. Zelfs de meest complexe eindtermen van de overheid zijn in [yyy] omgezet in klare taal."

Website Educatieve uitgever



Universiteit Utrecht Faculty of Humanities

3

Fragment (1)

Niemand wist toen iets over hygiëne. Overal liepen luizen en vlooien. Mensen wasten zich nauwelijks. De koning dacht dat wassen ongezond was. De pruiken van mensen zaten vol luizen en vlooien. In de zomer was het erg heet in het paleis....




Universiteit Utrecht Faculty of Humanities

4

Fragment (2)

Niemand wist toen iets over hygiëne en dat veroorzaakte grote problemen. **Ten eerste** liepen overal luizen en vlooien, **omdat** mensen zich nauwelijks wasten. **De reden was** dat de koning vond dat wassen ongezond was. **Ten tweede** was het in de zomer erg warm in het paleis.



Universiteit Utrecht Faculty of Humanities

5

We deden onderzoek

We vergeleken in diverse experimenten steeds twee tekstversies: met en zonder signalen van coherentie

- Honderden scholieren (15-16 jaar; diverse niveaus)



Jacqueline Evers-Vermeul



Gerdineke van Sijfhout



Jentine Land



Ted Sanders



Universiteit Utrecht Faculty of Humanities

6

We deden onderzoek

- We toetsten tekstbegrip
- Inferentievragen:
 - Waarom wisten mensen zich nauwelijks?
- Sorteervragen:
 - zet de volgende trefwoorden in de juiste categorie
 - Vul onderstaand schema in



7

Resultaten:

Teksten zonder coherentie-signalen worden

- ...door lezers slechter begrepen; leerlingen scoren lager op tekstbegrip
- dan op een tekst die wél coherentie-signalen bevat: **Maar, dus, aan de andere kant.**
- En dat was zo voor alle onderwijsniveau's: Vmbo én HAVO- en VWO-lezers
- Onderzoek met Jentine Land (2009), Gerdineke van Silfhout (2015), samen met Stichting Lezen. In diverse wetenschappelijke tijdschriften.



8

Samenvattend: teksten met coherentie-signalen

- zijn gemakkelijker te lezen, omdat de signalen de relatie expliciet maken.
- worden beter begrepen omdat ze tot een coherente representatie leiden. Zo scoren lezers beter op tekstbegrip
- Dus het is écht niet zo simpel als: korte woorden en korte zinnen lossen het op.



9

Veel onderzoek gedaan. Wat is onze rol?

- Wetenschappelijk onderzoek en advies over begrijpelijkheid
- Mede vanwege landelijk NWO-programma Begrijpelijke Taal & Effectieve communicatie
- NWO met vele maatschappelijke partners
- Netwerk Begrijpelijke Overheid



10

Zo deden we veel onderzoek

- Experimenteel onderzoek waarin we tekstversies vergelijken
- Welke wordt het beste begrepen door de beoogde lezers?
- Kan ook gaan om vindbaarheid en bruikbaarheid
- veelal onderzoek samen met maatschappelijke partners
- in maatschappelijk belangrijke en relevante contexten



11

Teksten en financiële producten: pensioen en hypotheek

The image shows two newspaper headlines. The left one reads 'Hypotheekakte onbegrijpelijk' with a sub-headline 'Dit is de laatste kans om te weten, in vijf teksten, hoe de hypotheekakte werkt'. The right one reads 'Bijsluiter complexe financiële producten zelf te complex' with a sub-headline 'Dit is de laatste kans om te weten, in vijf teksten, hoe de bijsluiter werkt'. Both headlines include social media icons and a QR code.



12

Andere voorbeelden

- Monitor Begrijpelijke overheid (Pander Maat, samen met HAN, voor BZK)
- Juni 2023: een advies over begrijpelijke overheid-burger-communicatie (Van Cramwinckel, Pander Maat & Sanders); nu een wetsvoorstel....



UNIVERSITEIT
NEDERLANDSE TALEN EN VERBODINGEN

Klare taal

Wolters Kluwer

Universiteit Utrecht
13
Faculty of Humanities

13

Maar kunnen we begrijpelijkheid ook voorspellen?

- Begripsonderzoek doen is veel werk, gebonden aan die ene context.
- Kunnen we zo langzamerhand niet gewoon aan de tekst zien hoe moeilijk hij is?
- De wetenschap schrijdt immers voort: corpora, computationeel, psycholinguïstisch

UNIVERSITEIT
NEDERLANDSE TALEN EN VERBODINGEN

Universiteit Utrecht
14
Faculty of Humanities

14



LiNT

Een nieuwe leesbaarheidstoel voor het Nederlands

UNIVERSITEIT
NEDERLANDSE TALEN EN VERBODINGEN

Universiteit Utrecht
Faculty of Humanities

15

Wat is LiNT ?

- Leesbaarheidsinstrument voor Nederlandse Teksten
- Gebouwd op basis van T-Scan: analyse-tool
- T-Scan werd gebouwd dankzij onderzoeksprogramma Begrijpelijke Taal.
- Partners: NWO, Tilburg/Nijmegen, CITO en de Taalunie
- Universiteit Utrecht (cöördinatie)

UNIVERSITEIT
NEDERLANDSE TALEN EN VERBODINGEN

Universiteit Utrecht
16
Faculty of Humanities

16

Leesniveaus en leesbaarheidsonderzoek

- Je voert een tekst in LiNT en je krijgt terug:
 - Het lees- of moeilijkheidsniveau van je tekst
 - Een score op tekstenmerken, die je ook kunt vergelijken
- Een leesniveau is een automatisch vastgesteld moeilijkheidsniveau voor een tekst, dat voorspelt hoeveel lezers moeite zullen hebben om een tekst te begrijpen

UNIVERSITEIT
NEDERLANDSE TALEN EN VERBODINGEN

Universiteit Utrecht
17
Faculty of Humanities

17

Wat betekent een LiNT-niveau voor volwassenen?

Tekstniveau	LiNT-score	Hoeveel volwassen lezers kunnen dit niveau niet goed aan?
1	0 - 36	14%
2	36 - 51	30%
3	51 - 61.5	52%
4	61.5 - 100	80%

UNIVERSITEIT
NEDERLANDSE TALEN EN VERBODINGEN

Universiteit Utrecht
18
Faculty of Humanities

18

LiNT-niveaus

Hoe moeilijk is uw tekst?

Voor hoeveel lezers levert uw tekst problemen op?

Kleuterverhaal	1
Brief schuldhulp, nieuw	33
Brief schuldhulp, oud	51
Bestuursrecht, nieuw	55
Bestuursrecht, oud	94

Universiteit Utrecht 19 Faculty of Humanities

19

Voorbeeld: ambtelijke brief

Geachte heer Post,

In deze brief geef ik antwoord op uw schriftelijke reactie naar aanleiding van mijn antwoord van 5 januari 2011.

U vraagt zich af waarom uw ex-vrouw een aanslag afvalstoffenheffing heeft ontvangen voor haar nieuwe adres. De verordening op de heffing en de invoering van afvalstoffenheffing en reingeregelingen 2010 lot en de basis van een aanslaglijst afvalstoffenheffing. Hierin is geregeld dat de gebruiker van een woonruimte de afvalstoffenheffing verschuldigd is tot het begin van het belastingjaar. Indien de belastingplicht in de loop van het jaar aanvangt dan is de afvalstoffenheffing verschuldigd naar tijdsereendebiet. Op grond hiervan is de aanslag afvalstoffenheffing van uw ex-vrouw terecht berekend op basis van 5 maanden. Zij is namelijk gebruiker van een woonruimte vanaf 20 juli 2010.

Daarnaast is, ten aanzien van de meestaf en tarieven, in de genoemde verordening geregeld dat indien een woning op 1 januari van het belastingjaar wordt bezocht door een persoon er een lager tarief van toepassing is. Omdat u op 1 januari 2010 niet alleenwond was komt u niet voor de lagere tarief in aanmerking.

Kortom, de belastingplicht voor de afvalstoffenheffing wordt bepaald aan de hand van het aantal maanden dat men op een bepaald adres woonachtig is. Verwijgens wordt er voor de hoogte van de aanslag gekeken naar de situatie op de peildatum van 1 januari van het belastingjaar.

Ten aanzien van de beroepsmogelijkheid verwijs ik u naar de inhoud van mijn brief van 5 januari 2011.

Voor vragen kunt u telefonisch contact opnemen met bovengenoemde medewerker. Ook kunt u terecht in de gemeentewinkel in het stadhuis. De riging is aan de Platanenstraat.

De gemeentelijke opening:
- maandag tot en met donderdag van 09.00 uur tot 15.00 uur;
- dinsdag is er ook een avondopening van 17.30 uur tot 19.00 uur;
- vrijdag van 09.00 uur tot 13.00 uur.

Ik vertrouw erop u hiermee voldoende te hebben geïnlicht.

Hoogachtend,
De wettigambtenaar van de gemeente X.

Universiteit Utrecht 20 Faculty of Humanities

20

Voorbeeld: herziene brief

Geachte meneer Post,

U heeft mij een brief geschreven. Daarin vraagt u waarom uw ex-vrouw over 2010 huisvuilbelasting moet betalen. Ook vraagt u om een lagere huisvuilbelasting voor uzelf. In deze brief geef ik antwoord op uw vragen.

Waarom moet uw ex-vrouw over 2010 huisvuilbelasting betalen?
Omdat zij sinds 20 juli 2010 op haar nieuwe adres woont. Zij heeft daar in 2010 dus vijf hele maanden gewoond. Over die vijf maanden moet zij huisvuilbelasting betalen. Onze regels zijn hierover duidelijk. We kijken naar het aantal maanden dat iemand ergens woont. Zo rekenen we uit hoeveel hij of zij moet betalen.

Kan uw huisvuilbelasting over 2010 omlaag?
Nee, dat kan niet. We alken woont, betaalt inderdaad minder huisvuilbelasting dan bijvoorbeeld een gezin. Maar daarbij kijken we naar de situatie op 1 januari. Op 1 januari 2010 woonde u nog niet alleen in uw huis. Daarom betaalt u over 2010 geen lagere huisvuilbelasting.

Hoe kunt u bezwaar maken?
Bent u het niet eens met deze besluiten? Maak dan bezwaar. In mijn brief van 5 januari 2011 lees u hoe dat moet.

Waar kunt u meer informatie krijgen?
Heeft u vragen? Bel mij dan. Mijn gegevens staan bovenaan deze brief. U kunt voor informatie ook de gemeentewinkel in het stadhuis bezoeken. Het adres is Platanenstraat 10. De gemeentewinkel is op de volgende dagen en tijden open:
- maandag, dinsdag, woensdag en donderdag: van 09.00 tot 15.00 uur;
- dinsdag ook van 17.30 tot 19.00 uur;
- vrijdag: van 09.00 tot 13.00 uur.

Met vriendelijke groet,
J. Maas
Medewerker gemeente X

Universiteit Utrecht Faculty of Humanities

21

Tekst vergelijken met andere teksten

Genres Uw teksten

Universiteit Utrecht Faculty of Humanities

22

Vergelijking: voorbeeld bijvoeglijke bepalingen

Bijvoeglijke bepalingen

Bijvoeglijke bepalingen kunnen een deeltje behoorlijk informatiedicht maken. Hoeveel bijvoeglijke bepalingen heeft uw tekst per deeltje? We geven ook de aantallen van een eenvoudig genre (oddabeeldichten) en een complex genre (wetenschappelijke artikelen). [Meer informatie](#)

Brief schuldhulp nieuw	~0.8
Roedelichten	~0.9
Bestuursrechtsgenueel nieuw	~1.5
Brief schuldhulp oud	~1.6
Wetenschap	~2.1
Bestuursrechtsgenueel oud	~2.6

0.5 1.0 1.5 2.0 2.5 3.0

< makkelijker | moeilijker >

Universiteit Utrecht 23 Faculty of Humanities

23

Feedback aan schrijver: drie groepen tekstenkenmerken

- Hoe moeilijk zijn uw woorden?
 - Woordfrequentie
 - Woordconcreetheid
 - Woordherhaling
- Hoe persoonlijk is uw tekst?
 - Mensen
 - Persoonlijke voornaamwoorden
 - Aansprekingen
 - Lijdende vormen
- Hoe moeilijk zijn uw zinnen?
 - Zinslengte
 - Bijzinnen
 - Lengte van deelzinnen
 - Afhankelijkheidslengtes
 - Bijvoeglijke bepalingen
 - Opsommingen

Universiteit Utrecht 24 Faculty of Humanities

24

Wat kan LiNT *niet*? En wat wel?

- LiNT herschrijft je tekst niet
- LiNT garandeert niet dat je tekst werkt, want het negeert:
 - de inhoud (grotendeels)
 - de opbouw
 - de vormgeving
 - de toon van de formulering (grotendeels)
- LiNT geeft lampjes die knippen bij complex formuleren



25

Hoe hebben we LiNT gebruikt?

Als je er een paar teksten in stopt

1. Verzamel feedback op je eigen teksten
2. Geef anderen feedback op teksten

Als je er veel teksten in stopt

3. Bepaal met welke teksten iets moet gebeuren:
 - herschrijven; testen; minimaal beter bekijken
4. Peil je tekstniveaus, om te zien:
 - of herschrijven het taalniveau verlaagt
 - hoe ver je organisatie is met begrijpelijkheid.

26

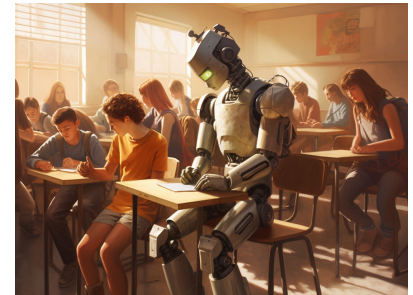
De vier topvoorspellers

- **Woordfrequentie**
 - Hoe vaak komt een woord voor in een corpus Nederlands
- **Woordconcreetheid**
 - De proportie zelfstandige naamwoorden uit de tekst die verwijst naar zintuiglijk waarneembare zaken, plaatsen en tijden
- **Afhankelijkheidslengte: wat bij elkaar hoort staat bij elkaar**
 - Per zin bepalen we door ontleding de langste afhankelijkheid ('tangconstructie') en nemen daarvan het gemiddelde per tekst
- **Lengte van deelzinnen: informatie-dichtheid**
 - We tellen niet per zin, maar per deelzin: elk stukje tekst met een vervoegd werkwoord
 - We tellen alleen de inhoudswoorden, niet de functionwoorden (lidwoorden, voorzetsels e.d.); want die belasten de lezer minder

27

De stap naar Generatieve AI

- Hoe werkt Generatieve AI zoals ChatGPT?
- Kan Generatieve AI teksten herschrijven, zodat de leesbaarheid beter wordt,
- Maar zo dat de boodschap wel dezelfde blijft?
- Ja, maar...



28

Wat is Generatieve AI niet?

- *Artificial General Intelligence* - intelligentie die die van de mens benadert of overschaduwet. **Van geen kanten.**

Als gevolg van hoe de technologie werkt kan Generatieve AI niet:

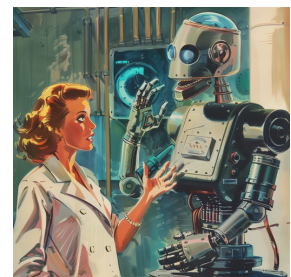
- redeneren (het genereert slechts plausibel klinkende teksten) en kan niet rekenen;
- feiten controleren aan de hand van andere feiten;
- buiten de trainingsdata 'redeneren', het heeft geen kennis van de wereld.



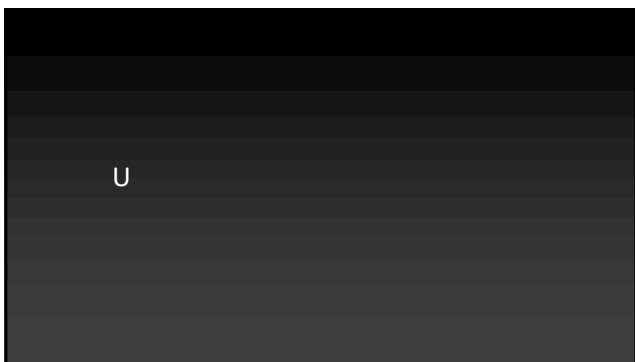
29

Wat is Generatieve AI wel? LLMs:

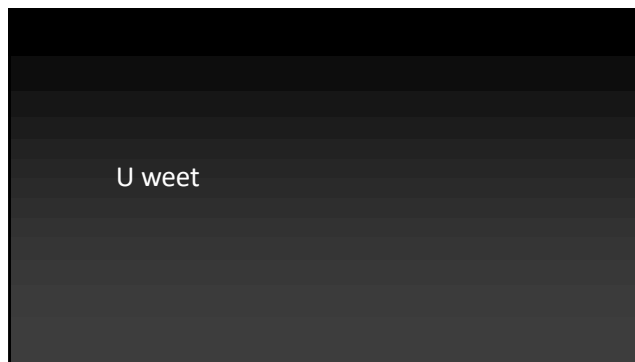
- LLMs zijn generators van teksten die een **plausibel** en **behulpzaam** vervolg zijn van de input van de gebruiker, de prompt.
- Dat doen ze door in de basis heel goed te zijn in het **voorspellen van het volgende woord**.
- En door uit mogelijke plausibele antwoorden de meest **behulpzame** te kiezen.
- (Ze **hallucineren** dus niet – dat zijn ze voortdurend aan het doen!)



30



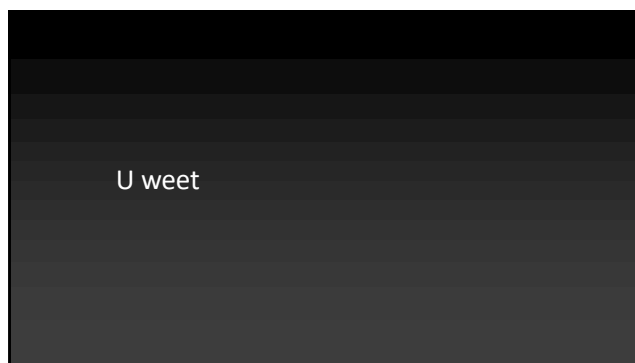
31



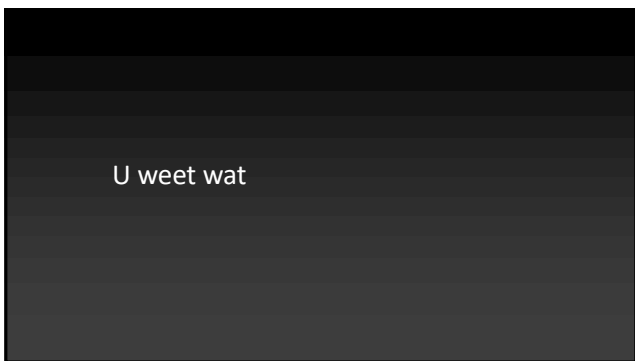
32



33



34



35



36

U weet wat

37

U weet wat ik

38

U weet wat ik ga

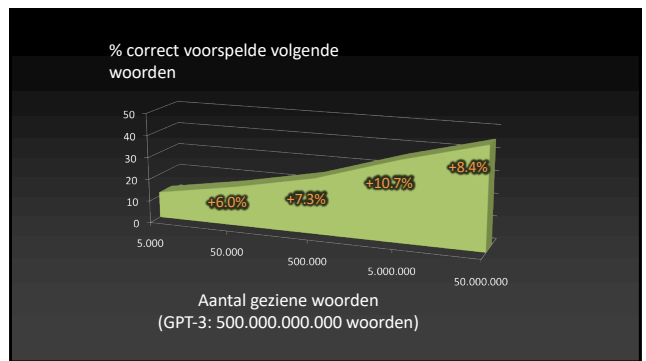
39

U weet wat ik ga zeggen.

40



41



42

Generatieve AI: prompt engineering

Verschillende richtlijnen, zoals

- BRIGHT, Christine Merie Fox (rechts)
- Mollick, Ethan R. and Mollick, Lilach, Assigning AI: Seven Approaches for Students, with Prompts (2023).

The BRIGHT Framework		
	Definition	Example
B	Boundaries Set limits on the format, length, or any other constraints you want to impose on the AI's output.	Provide bullet points only, no need for full sentences.
R	Role Identify the role you want the AI tool to take while generating content.	If you're looking for scholarly feedback, you might specify the role as a "journal reviewer."
I	Intention Explain why you need the task to be completed; this helps in aligning the AI's output with your objective.	So I can review the notes for my upcoming exam.
G	Goal Clearly state what you want the AI tool to accomplish.	Create an outline from my course notes.
H	Hone After receiving the AI's output, provide feedback to improve future interactions.	The bullet points are too vague; I need more specific details.
T	Tone Specify the tone you want the AI tool to adopt in its output. This could be formal, informal, academic, conversational, etc.	Write in a formal academic tone suitable for a research paper.

49

Generatieve AI: andere slimmigheden

- Plugins voor het inlezen van documenten, spraakherkenning, ...
- **Retrieval augmented generation:** combinatie van LLM met zoekmachine / database
- Steeds groter geheugen voor prompts; van honderden naar honderdduizenden woorden
- Multimodale output: plaatjes, spraak

<https://aws.amazon.com/blogs/machine-learning/question-answering-using-retrieval-augmented-generation-with-foundation-models-in-amazon-sagemaker-jumpstart/>

50

Generatieve AI en leesbaarheidstools

- Combinatie van automatisch gegenereerde teksten en LINT (reward-functie in reinforcement learning)
- Functies van T-Scan en LINT laten leren door LLMs, inclusief automatische en tekstafhankelijke uitleg van de LINT-score
- Maar: menselijke expertise blijft nodig voor het doorzien van de hele context (schrijver, doel, doelgroep, ...)

51

Wat staat ons op korte termijn te wachten?

Positieve insteek:

- Nieuwe architecturen: beter, duurzamer
- Integratie modaliteiten: audio, beeld, spraak, video
- Betere *alignment*, betere *guardrails*
- Koppeling aan betrouwbare kennisbronnen (RAG)
- Meer keuze voor gebruikers
- Goede Nederlandse modellen, digitaal soeverein (GPT-NL)
- Veilige personalisering

52

Wat staat ons te wachten?

Negatieve/realistische insteek:

- Amerikaanse big tech bestendigt marktdominantie
- Geen oplossing voor bias; opkomst synthetische data; geen inclusiviteit
- Grotere ecologische voetafdruk (Stargate, 2028, Microsoft/OpenAI, 100 miljard dollar)
- *Misalignment*, *dual use*, misbruik

53

Met dank aan

nature machine intelligence

Explore content | About the journal | Publish with us | Subscribe

nature > nature machine intelligence > comment > article

Comment | Published: 28 July 2023

Foundation models and the privatization of public knowledge

Fabian Ferrari ¹, José van Dijk ⁴, Antal van den Bosch

Nature Machine Intelligence 5, 818–820 (2023) | [Cite this article](#)

747 Accesses | 9 Altmetric | [Metrics](#)

Ted Sanders, Henk Pander Maat

Florian Kunneman, Yupei Du

Fabian Ferrari, José van Dijk

To protect the integrity of knowledge production, the training procedures of foundation models such as GPT-4 need to be made accessible to regulators and researchers. Foundation models must become open and public, and those are not the same thing.

54

