

Dit is een transcript bij: [Aflevering 1: Piek Vossen – Hoe werkt ChatGPT eigenlijk? - Nooit meer schrijven? | Podcast on Spotify](#)

In lijn met het thema van de podcast – kunstmatige intelligentie en schrijfvaardigheid – hebben we een automatisch transcript laten maken door Amberscript, die we minimaal hebben aangepast om fouten/onduidelijkheden eruit te halen.

00:00:00

Gea: Welkom bij "Nooit meer schrijven?", een podcast over kunstmatige intelligentie en schrijfvaardigheid. Leuk dat je luistert. Mijn naam is Gea Dreschler. Ik ben universitair docent Engelse taalkunde en wetenschappelijk directeur van het Academic Language Programme of terwijl het ALP, onderdeel van faculteit geesteswetenschappen. In elke aflevering van deze podcast praat ik met iemand die op één of andere manier met de thema ChatGPT en AI tools bezig is. De gesprekken gaan over hoe deze tools nou eigenlijk werken, wat ze kunnen (en wat niet) en vooral ook wat we ermee moeten in het onderwijs. Verbieden, anders toetsen, studenten maar niet meer laten schrijven? Samen met de gasten bekijk ik het onderwerp van alle kanten.

00:00:54

Gea: Vandaag te gast is Piek Vossen, hoogleraar Computatieve Lexicologie bij de faculteit geesteswetenschappen hier aan de VU. Piek gaat ons vooral vertellen over de technische kant van ChatGPT; Dus wat zijn de taalmodellen die achter ChatGPT zitten? Waarom is ChatGPT nou opeens zo populair? En werken AI detectors eigenlijk wel?

00:01:17

Gea: Welkom Piek, leuk dat je er bent! Ik stel elke gast eerst een paar vragen over hun eigen schrijfproces of schrijfvaardigheid. Ben jij een Mozart of een Beethoven, dus bedenk je de hele compositie en schrijf hem dan één keer op, of schrijf je in eerste versie en blijf je eindeloos schaven aan de tekst tot je eens een keer iets hebt, wat een eindversie is?

00:01:40

Piek: Ja, dat hangt er vanaf met welk doel ik schrijf, denk ik. Dus wetenschappelijke artikelen is wel wat meer gepland, gestructureerd. Maar als ik gewoon een mail schrijf, of een brief of een gedicht of iets anders of een verhaaltje, dan nou, dan is het gewoon: het ene woord roept het andere op, het ene beeld roept het andere op. Ja, dan heb ik niet van tevoren een idee.

00:02:04

Gea: Wat is de beste plek om te schrijven?

00:02:07

Piek: Ja, ik liefst eigenlijk toch op een laptop, soms trouwens doe ik wel weer gewoon pen en papier.

00:02:12

Gea: Luister je muziek als je aan het schrijven bent, of nee?

00:02:16

Piek: Nee, nee.

00:02:17

Gea: Geen muziek dus, absolute stilte?

00:02:18

Piek: Nee, dat doe ik niet meer, dat deed ik vroeger wel. Ja.

00:02:21

Gea: En welke tools gebruik je zelf als je aan het schrijven bent?

00:02:25

Piek: Maar ja, ik denk dat de spelling en grammen checkers die zijn. Die kun je bijna niet uitzetten dus ja, die gebruik ik wel veel. Ik gebruik geen tools die teksten genereren, of, zo, ik denk dat het de helft van de tijd ook verkeerde suggesties zijn, die ik dan weer moet verbeteren omdat ze iets veranderd hebben, of een verkeerd woord. Of dan, dan lees ik het terug en denk ik mijn god hebben ze weer wat anders ingevuld, wat ik niet geschreven heb.

00:02:50

Gea: Je wilt zelf zelf controle houden over je tekst.

00:02:53

Piek: Ja, ik heb het. Ik heb niet echt iets nodig om woorden of stukken tekst voor mij te verzinnen. Wel, als het, zeker als ik in het Engels schrijf, als ik niet precies weet hoe je iets in het Engels zegt, dan is het wel nuttig. Maar dan gaat het echt meer om bijvoorbeeld het juiste werkwoord, voorzetswoord, werkwoordcombinaties/constructies en niet zozeer lexicaal eerlijk gezegd, ja, en ook als je over je werk schrijft, dan ken je alle termen, dan weet je eigenlijk alle woorden wel. Dus het gaat echt om formuleringen dan, en dan in het Engels. Ja, daar ben ik geen native speaker, dus dan dan gebruik ik wel dat soort suggesties en zo.

00:03:30

Gea: En we gaan het hier uitgebreid over hebben straks. Maar: Al in het onderwijs, moeten we dat verbieden of omarmen?

00:03:38

Piek: Nou, ja, verbieden, gaat het je niet lukken denk ik. Ik zou niet weten hoe je dat voor elkaar gaat krijgen. Nee, kunt niet studenten permanent observeren, dat wil je ook helemaal niet. Volgens mij is dat geen prettige sfeer, zo moeten we niet samen in het onderwijs staan. Bovendien is het gewoon een onderdeel van de wereld. Dus je moet ook accepteren dat het er is en samen zorgen voor: Hoe kunnen we het inzetten in het leerproces? Want dat is het he, een leerproces, niet alleen maar toetsen: Leren is niet alleen toetsen, maar toetsen is wel ergens voor nodig, kennelijk. Maar sommige toetsvormen, die gaan niet meer. ie kun je vergeten, denk ik in de toekomst.

00:04:21

Gea: Gaan we straks nog verder over over praten, misschien voor mensen die hier niet elke dag mee bezig zijn; Wat is Computational taalkunde of lexicologie moet ik zeggen, Sorry.

00:04:32

Piek: Ja, nog moeilijker, Lexicologie ja dus Lexicologie dat is dan de de wetenschap die kennis van woorden probeert op te bouwen. En ja, als je die kennis in een computer stopt, dat zou je al computationele lexicologie kunnen noemen. Maar dan zou dat dus een database zijn van woorden, dus dan moet je eerst weten: wat zijn de woorden van de taal? Dat weet dan niemand. Soms zijn uitdrukkingen ook woorden, ook combinaties als "sterke koffie": is dat een woord of moet dat zijn eh, "Krachtige koffie", "bittere koffie" had ook gekund, maar wij gebruiken "sterke koffie" nu eenmaal dus dat soort collocatie achtige zaken zijn misschien ook wel woorden. Dus wel of geen spatie tussen zetten, dat is allemaal conventie. En als je zo doorgaat dan dan hebben we het over miljoenen miljoenen mogelijke woorden. En als je een taal wil leren, dan dan ken je die combinaties ook. Iemand die een taal leert, die maakt typische fouten, dat-ie op zich wel de woorden kennen, maar niet de juiste combinatie van woorden. Nou ja, als je dat wilt registreren, al dat is al een enorme taak en het liefst wil je dat dan ook nog eens doen door te koppelen aan hoe die woorden ook eigenlijk gebruikt worden. Geschreven teksten, maar ook gesproken taal in interacties tussen mensen, al die data die verzamelen we natuurlijk ook. Dus teksten, die vind je op het internet en interacties, die kun je opnemen en dan kun je zeg maar dit zou zouden de woorden kunnen zijn van de taal. Maar hier zie je hoe de mensen die woorden gebruiken, en deze mensen gebruiken ze zo, en die deze mensen gebruiken ze zo. Die worden niet, of heel veel of heel weinig. Iedereen kent heel veel woorden, maar gebruik maar een klein deel daarvan. Hoe zit dat dan precies? Nou ja, zo ga je eigenlijk door en door en door, en dan opeens gegeven moment ga je naar het punt dat je zegt. Maar wacht effe, je kunt ook die computer die woorden zelf laten gebruiken. Dus dan ga je een systeem maken dat het menselijk gedrag, van hoe mensen woorden gebruiken, gaat imiteren of modeleren nou ja, dan komen we een beetje in de richting van de huidige AI achtige toepassingen waar je dan te maken hebt met een computer die geleerd heeft om woorden te gebruiken, zoals wij ze ook gebruiken.

00:06:37

Gea: Wanneer is die die ontwikkeling eigenlijk begonnen? Binnen de taalkunde want op zich lexicologen hebben we natuurlijk al veel langer. Dat gaat veel langer terug. Maar die die slag naar meer AI toepassing, is dat echt iets van de laatste vijf jaar, tien jaar?

00:06:53

Piek: Nee, nee, nee, toen ik ik weet nog dat ik masterstudent was bij de UvA bij Hugo Brandt Corstius en die die kwam toen met enorme dozen aanzetten en daarin stonden de drukproeven van het woordenboek "hedendaags Nederlands". Volgens mij had ie dat van Van Dale gekregen. En hij had ook een digitale versie ervan. Dus toen zeiden we: Hé, maar wacht even, nu kunnen we in die digitale versie op een grote schaal allerlei patronen proberen te ontdekken en dergelijke dus een beetje. Dat was ik denk begin jaren 80 of zo. Op dat moment waren ze ook in Cambridge bezig en ook bij de UvA bezig met Longman Dictionary of Contemporary English dat

was één van de eerste woordenboeken die digitaal beschikbaar waren, in het Engels dan en die voor onderzoek gebruikt konden worden. In Amerika waren ze daar ook mee bezig. Dus een aantal groepen die begonnen dat te doen en begonnen dat van elkaar te imiteren, en toen had je in één keer, waar je vroeger taalkundig ging, dan over een paar voorbeelden, typisch artikel waarin iemand zijn contractieve analyse over één woordje, zus of zo gebruikt. Nou ja, dan heb je vijf stukjes tekst, misschien tien woorden. In een artikel die bestudeerd worden, konden we er nu 100000 bekijken en analyseren. Dus dat ging in één keer, van een schaal van een handvol voorbeelden naar ja, bijna meer woorden dan een mens überhaupt kent. En daarnaast had je dus corpuslinguïstiek, dat was al wat ouder van de jaren 50, volgens mij, ja, dat ging als eerste teksten verzamelen en eigenlijk, ik denk in de jaren 80, is dan die combinatie ook ontstaan en is het ook empirisch geworden. Eigenlijk dus die hele empirische kant die is gekomen met corpuslinguïstiek maar ook met die Computationale lexicologie in de jaren 80, waarin je nu in staat bent om grote hoeveelheden teksten, grote hoeveelheden woorden te gaan analyseren op grote schaal. Ja, en dan blijkt dus dat die theorieën eigenlijk niet kloppen, helemaal.

00:08:54

Gea: Ja, en dat is natuurlijk heel anders dan onderzoekers die misschien in hun eentje op basis van nou ja, handschriften of voorbeelden of briefjes met voorbeelden van woorden een woordenboek proberen te maken. Nu, met zoveel data kun je natuurlijk veel meer zeggen over wat er echt gebruikt wordt.

00:09:09

Piek: Ja, en daarna krijg je dus woordenboeken waarin ze de informatie in het woordenboek illustreren aan de hand van echte voorbeelden uit het corpus maar dat is nog steeds niet gebaseerd op een empirische analyse. En eigenlijk is dat ook nooit gekomen omdat die lexicografen graven, dat zijn geen lexicologen en en ook wel niet alle lexicologentoch niet genoeg. De tools begrijpen hoe je uit de data ook meteen je modellen moet afleiden en daar zijn die wereld een beetje uit elkaar gaan lopen. Eigenlijk is dat hele machine learning tijdperk, dat een beetje in de jaren 90 echt gestart is en nog steeds doorloopt, dat zijn geen linguïsten, dat zijn echt computer science achtige wiskundigen, die al die machine learning modellen gemaakt hebben, die kijken toch op een andere manier naar taal. Die hebben een eigen pad gevolgd en die hebben uiteindelijk hele krachtige modellen gemaakt waarin het taal gedrag van mensen geïmiteerd wordt door een computer. En dat kun je doen door wat ze noemen "supervised learning". Dus dan zijn er toch veel mensen die eraan zeggen van ja, maar als je naar deze woorden kijkt of deze zinnestjes kijkt, daar hoor deze interpretatie bij. Nou, dat kan een sentiment zijn, maar dat kan ook allerlei andere informatie zijn. En dan kun je een computer leren om dezelfde associatie te maken. En dan geef je een willekeurige tekst en dan kan die interpretatie eraan plakken. Waarom en hoe die dat doet, dat is een heel ander verhaal. Of dat hetzelfde is als wij dat doen, is ook een ander verhaal; Waarschijnlijk heel anders. En dat is een heel eigen leven gaan leggen, en dat werd eigenlijk succesvoller dan alles wat ervoor gedaan is. Dus als wij een grammatica zouden schrijven in de klassieke vorm, met regeltjes: Dit en dat en dat zijn de woordvolgordes van het Nederland, dit zijn de regels die erbij horen. Hier heb je een, een woordenboek of lexicon met de woorden met hun eigenschappen. Nou, ga maar proberen om een willekeurige tekst daarmee te analyseren, dan doet die tot op bepaalde hoogte goed en voor de rest fout. En als

je een machine learning programma dat laat afleiden op een indirecte manier, doet die het veel beter dan welke grammatica kan ook. Dus ja, dan sta je dan dus met je theorie, nou echt, ik ben opgeleid in de klassieke linguïstische manier van denken. En ja, je moet gewoon toegeven dat die hele stomme computer science achter benadering van een taal uiteindelijk beter werkt, beter in staat is om een tekst te analyseren dan dan al die theorie. En dan heb je verloren.

00:11:42

Gea: Ja, dan hebben we toch heel veel werk gedaan, wat misschien in dit geval een computer inderdaad beter kan. Je hebt nu al iets uitgelegd over machine learning maar hoe? Als we naar ChatGPT kijken, want daar gaat het nu natuurlijk de hele tijd over: Hoe werkt dat? Wat zit daarachter?

00:12:00

Piek: Ja, ChatGPT is eigenlijk een een product van een weer een andere vorm van machine learning, wat ik net heb uitgelegd en dat is dan, zoals het heet, unsupervised learning. Bij supervised learning, dan geef je een voorbeeld als het ware en plakt er een label aan. En dan gaat die computer dat label associëren met de eigenschappen van je voorbeeld. En wat die eigenschappen zijn, daar kun je van alles en nog wat bedenken en dan hebben ze ook... Ze kunnen het niet zo gek bedenken, dan hebben ze al bedacht van oh misschien kunnen we dat ook als een feature van dat voorbeeld representeren. Dus dat is nou ja, iets van tien, 20 jaar ontwikkeling geweest waarin ze uitgevonden hebben van: Hoe kan ik die tekst op de goede manier representeren, zodat die computer kan generaliseren over die tekst en dat kan toepassen op een andere tekst. Bij een unsupervised learning zeg je helemaal niet hoe je een tekst moet interpreteren. Nee, je gaat niet een paar en dan heb ik het over misschien 10000 of 100000 voorbeelden, wat tot nu toe de aantallen waren. Nee, je geeft een computer miljoenen, miljarden teksten. En het enige wat je als opdracht opgeeft is van: Nou, als jij een reeks woorden in een bepaalde volgorde krijgt en ik zou af en toe een een woordje maskeren, "masken" heet dat dan, dan vraag ik aan jou, computer, wat denk jij dat het goede woord is dat daar gestaan heeft? Nou, dan hoeft ik, ik hoeft niks te doen, want uit al die data pak ik gewoon alle alle zinnen die ik maar kan vinden, en ik plak gewoon af en toe wat woorden af en ik weet wat het goede woord was en een random woord is het foute woord. Nou ja, wat dat unsupervised learning gaat doen, is dat je de andere woorden die in die tekst staan, die gaan ze "representeren". En wat dat is, komt zo, dat weer een ander verhaal. Die gaan ze een representatie geven waarmee ze het juiste woord kunnen voorspellen. Hè, dus ik werd vanmorgen wakker en ik maakte een.

00:13:58

Gea: Boterham?

00:14:00

Piek: Ja, of een kopje koffie of of een smoothie, of eigenlijk zijn we cliché, dieren, hè, dus we vullen dat lijstje heel makkelijk, heel snel in. En die computer gaat dat ook leren. Door gewoon alleen maar te kijken naar: Wat hebben al die Nederlanders allemaal ingevuld op dat dingetje wat daar afgeplakt is? In het Engels doet ie dat ook, en in het Frans en het Duits en welke taal dan ook. En wij zijn allemaal ongeveer dezelfde cliché dieren over de hele wereld heen. En dus we gaan allemaal dezelfde soort dingen invullen. Wat gebeurt er dan? Opstaan, had ik gezegd. Ik sta

vanmorgen op en eerste wat ik deed, ik "maakte een", "maken", het is allemaal woorden. Die krijgen nu een soort representatie, die ze in staat stelt om in dit zinnetje het volgende woord te voorspellen, of een woordje terug, links en rechts. Maar die representatie gaat nu het woord vervangen. Dus dat betekent dat de representatie van maken of opstaan even gaat lijken op andere woorden die we ook op die plek zouden gebruiken, die min of meer hetzelfde betekenen om weer zo'n woord te voorspellen. En in het eerste voorbeeld zei ik van ja in de machine learning, dan pakken ze hun zin en dan gaan ze die woorden als eigenschappen zien. En dan moeten ze een associatie leggen. En als ik een nieuwe tekst hebt, dan gaat die kijken of die dezelfde eigenschappen heeft. En als dat zo is, dan gaat die hetzelfde label voorspellen. Maar nu, in unsupervised learning, heeft-ie geleerd om te extraheren van die woorden en één of andere hele bizarre representatie te leren. Die staat voor alle woorden die een min of meer vergelijkbare betekenis of gebruik hebben, qua betekenis, maar ook qua gramaticale linguïstische eigenschappen op die plek, in die zin. Dus, kan aan de hand van die representatie dat model, als ik nu weer supervised learning ga doen. Er staat een reeks woorden die vervanging door die representatie en die generaliseren over alles wat daar maar kan staan. Dus in plaats van dat ik nu een associatie heb met mijn label, met één specifiek voorbeeld, heb ik als het ware die representatie gegeneraliseerd naar alle mogelijke teksten die je kan hebben, die min of meer dit representeren en dan wordt plotseling dat leren veel, veel, veel efficiënter.

00:16:16

Gea: En kun je dan eigenlijk zeggen dat, zo'n programma en ChatGPT, dat snapt niet wat die doet? Dat is allemaal op basis van frequentie en voorspellingen over combinaties die het meeste voorkomen, en op basis daarvan een verwachting van wat er volgt op dit woord of wat er in deze volgorde komt? Is het allemaal een soort van plat eigenlijk qua structuur, dus niet hiërarchisch of een soort zins analyse, maar echt op basis van woorden die je ziet?

00:16:45

Piek: Deze modellen, dat zijn dan, zeg maar transform modellen en GPT is daar één van en je hebt veel meer. Bird is een andere. Je hebt allerlei varianten van die modellen, hoor, en een OpenAI was eigenlijk niet de eerste die dit deden en zijn ook niet de enige die dit doen. Die die zijn op zich best wel, ze zijn op zich, ze zijn groot, laat ik het zo zeggen, niet-complex.

00:17:09

Gea: Je zei net al dat die die taal modellen, die waar ChatGPT op gebouwd is, zijn eigenlijk niet nieuw. Er zijn misschien onderdelen van ChatGPT dus het jasje door me heen met de chatbot wat wel nieuw is. Maar verbaast het jou hoeveel aandacht er nu is voor ChatGPT en wat het kan? Of waar komt dat eigenlijk door? Waarom zijn we er zo door AI gecharmeerd?

00:17:29

Piek: Nou ja, ik denk, de reden is omdat OpenAI een soort laag heeft gebouwd. Het model dat er onder zit, is GPT. En dat is zo'n transformer model we het net over hadden.

00:17:43

Gea: Ik ben de afkorting alweer vergeten, wel transformer maar ja.

Deze spraak-naar-tekst is gecreëerd met www.amberscript.com

00:17:47

Piek: Generative Pre-trained Transformer. Generative betekent dat die in dat wat ik zei dat "masken", hè, dus dat een stukje wordt afgepakt en dat moet hij voorspellen. Als het generative is, dan wordt alleen maar het volgende woord, het vervolg afgeplakt. En als je Bird, dat is bi-directional, dan kun je aan links en rechts dingen afplakken daardoor heeft Bird eigenlijk een betere representatie van de zin, omdat soms informatie aan het einde ook belangrijk is om iets te representeren. Maar ja, voor die generatieve achtige toepassingen is GPT heel erg goed, dus die weet echt gewoon van. Als jij dit stukje in mij geeft, weet ik precies hoe het doorgaat.

00:18:29

Gea: Hij kan het afmaken en hij hoeft niet het einde van de zin te hebben.

00:18:32

Piek: En dan doen iesteeds één woordje aanvullen en met dat woordje het volgende en de volgende: Zo kunnen ze een hele lange tekst genereren. En de generatieve eigenschap van dit model, dat waar het dan heel erg goed in is, dat hebben ze uitgebuit door allerlei gedrag daarboven op te bouwen. Want onderin dat model kan die eigenlijk alles genereren, de meest flauwekul, onzin, rotzooi en en alle lelijkheid die je maar kan bedenken. Het model kan alles genereren, dus gegenereerd ook dat, dus daar heb ik geen controle over, maar bovenop hebben ze dus zomaar wat zij noemen intenties gebouwd, dat zijn die chat-achtige functies. En ze hebben zich gericht op samenvatten, vertalen, vraag-antwoord het en een aantal van die... Op zich helemaal niet zo heel veel. Als je zegt van ik heb 30 AI in mijn hand, hoe zwaar is dat? Dan, zal die iets genereren.

00:19:22

Gea: Zegt ie iets.

00:19:24

Piek: Altijd.

00:19:24

Gea: Of het dan waar is, is natuurlijk maar weer de vraag.

00:19:26

Piek: Ja en hoe absurder de input is, hoe absurder de output. Maar vaak vragen mensen heel veel standaard, simpele, basale dingen.

00:19:35

Gea: Maar hij zegt nu ook wel af en toe, of hij/zij/het... "Dit weet ik niet. Ik ben een model, dus ik kan dit niet..."

00:19:42

Gea:Ik weet niet wat voor weer het wordt vandaag, want ik heb daar geen grip op."

00:19:45

Piek: Ja, dat is die dat gedrag dat ze erbovenop creëren, en dat doen ze eigenlijk weer door supervised training, dus om om te laten zien als je, als je dit binnenkrijgt moet je zo reageren. Onderliggend heb je GPT en die genereert van alles, en ze laten gewoon 20 varianten van GPT zien, die leggen ze voor, en dan zeggen ze: ja, dat is de goede, dat is minder goed en dat is nog minder goed en dit is helemaal fout, enzovoort, enzovoort. Dus het systeem weet niet precies waarom nou de één beter

is dan de andere, maar het gaat dat gedrag, gaat-ie uiteindelijk nadoen en daar zit ook in van als je dit en dit gevraagd wordt, moet je geen antwoord geven.

00:20:22

Gea: Daar zit maar er zit eigenlijk dus nog best wel heel veel mensenwerk in, ook die die training.

00:20:26

Piek: Ja, dat hebben ze dus heel efficiënt gemaakt. Dat is de claim van OpenAI, dat dat hun bijdrage is, want de GPT zelf hebben ze niet bedacht. Ze hebben het wel ge-engineerd, dat wel heel goed. Maar dat hebben anderen eigenlijk al bedacht maar zij zeggen dat ze op een hele efficiënte manier dat gedrag bovenopmaken en dan snap je ook, omdat dat gedrag zo menselijk is en zo makkelijk dat je in je eigen woorden iets kan vragen aan een systeem en dan een antwoord terugkrijgt dat heeft het tot een hype gemaakt. Dus ze hebben eigenlijk... Het is gewoon heel laagdrempelig geworden om die complexe technologie te gebruiken waarvoor je eerst misschien een masteropleiding had moeten volgen, of zo. En nu kan iedereen het gebruiken, want je hoeft niet meer naar direct naar GPT te gaan. Nee, je stelt gewoon simpel je vragen aan dat ding dat er bovenop zit, en daardoor kan nu het grote publiek gebruik maken van deze technologie met alle goede en slechte eigenschappen.

00:21:21

Gea: Ja, want vind jij dat we daar eigenlijk... Hebben wij daar als het grote publiek iets aan? Of is het eigenlijk een soort van trucje hebben? Dingetje, kijk eens wat leuk en over een maand zijn we het weer vergeten.

00:21:31

Piek: Nou ja, sommige dingen zijn hartstikke handig en nuttig, dus ik zou het zeker bijvoorbeeld samenvattingen laten maken van allerlei dingen. Secundaire dit en dit en dit, vat het even voor mij samen. Ik denk dat dat heel goed kan. Dus als de informatie, dat noemen ze de prompt, dat is de input die je geeft en als je geen input geeft, krijg je ook geen output. Dus je moet iets geven en dat zijn je instructies. Nou, als, als als je het systeem vraagt om iets met die informatie te doen en hij heeft geen andere informatie nodig dan dat wat er staat, dan kan die vaak heel zinnig dat interpreteren en daar iets aan koppelen. Maar als je er iets van buiten bij moet gaan halen, ja, dan weet je niet waar die vandaan haalt. Dat zit ergens verstopt, in alles wat die allemaal gezien en gelezen heeft, en dan gaat die wat hallucineren zoals dat heet.

00:22:25

Gea: Die term heb ik we nou wel vaker voorbij horen komen, want wat houdt dat hallucineren in?

00:22:29

Piek: Nou, je moet je voorstellen dat zoals het dat onderliggende model werkt, en dat is gewoon moeilijk om daar controle op te hebben, en dat is wat zij proberen met die functies. Je hebt er een reeks woorden. Het volgende woord, wat die daar gaat genereren is eigenlijk een representatie van het volgende woord. En dan gaat die kijken naar het hebben we weer het woord lexicon/vocabulary dat zijn alle woorden die die in zijn representatie ooit heeft gezien. Die heeft die namelijk een soort lijst van gemaakt. En ieder woord heeft ook een representatie. En hij gaat gewoon kijken hoe

hoog scoort ieder woord als mogelijk wordt op wat ik verwacht, wat daar moet komen. Dus het woord dat daar uiteindelijk gekozen wordt door het systeem daar, dan kan die denken van: ja, dit wordt pas echt het allerbeste van allemaal. Ja, het heeft helaas in ons algemene wereldmodel helemaal niks te maken met waar wij specifiek in onze input om gevraagd hadden. Maar op grond van alles wat die eerder gelezen heeft, denk ik toch echt wel dat dit het beste woord is dat daar past. En dan gaat ie dat er gewoon bijzetten.

00:23:37

Gea: En verklaart dat ook dat die heel veel kan produceren, kan leveren, waarvan je dan denkt: Ja, het klinkt, het klinkt perfect, het ziet er heel overtuigend uit, maar maar inhoudelijk... Dus een verzonnen biografie van iemand of een verzonnen verhaal over waar een bepaalde achternaam vandaan komt, of.

00:23:57

Piek: Ja, maar dat komt er omdat het zeg maar bij dat een unsupervised learning heeft, hij die generalisatie gemaakt van wat die gelezen heeft. Dus er is geen index van de feiten die die gelezen heeft. Het is niet zo van oh, ja, ik heb dat gelezen en daar en daar wonen zoveel mensen en die doen dit en dat en dat. En met de kennis zelf, feitelijke kennis, die wij dan toch nog op één of andere manier een beetje onthouden. Als we dingen leren en lezen hij het heeft, heeft daar heeft daar alleen maar een soort conceptuele laag uitgehaald, maar de feitelijkheden direct heeft die niet geïndexeerd. Dus hij gaat eigenlijk iets genereren dat conceptueel daarbij past op grond van alles wat die gelezen heeft.

00:24:41

Piek: Hè, als die gewoon heel veel verhalen zou lezen over dat mensen met een linkerbeen eentram instappen als eerste geen idee of dat zo is, weet ik veel, dan denkt ie van ja, dat dat ja, dat heb ik toch eigenlijk duidelijk afgeleid, al die dingen, dus ik denk dat dat dat bij hierbij wordt. Dan gaat ie dat dan zomaar bijzetten.

00:24:58

Gea: Ja, dus dus in die zin geen filter op de input en en wat hij levert.

00:25:04

Piek: Nee, de enige filter is de aansturing van prompt, zoals ik noemde. Zo heet dat dan technisch tegenwoordig de prompt, en daarom heb je nu ook een nieuw beroep, de prompt engineer, dat is degene die probeert om zoveel mogelijk controle te houden over wat die nou precies gaat genereren door die prompt op een hele slimme, hele expliciete manier. Heel duidelijk moet je zijn tegen dat systeem als een klein kind. Dit moet je doen en dat moet je niet doen en je moet het zo doen, want anders kan die iets zomaar erbij verzinnen.

00:25:38

Gea: Ja, dus dit is al een heel duidelijk voorbeeld, denk ik, van van de beperkingen van van het systeem.

00:25:44

Piek: Ja, maar het hangt van het doel af. Hè, als jij geïnteresseerd bent in het creatieve, ja dan fantastisch, laten het helemaal lekker verzinnen maakt niet uit.

00:25:53

Gea: Misschien voor een wetenschappelijk artikel niet.

00:25:56

Piek: Nee, ik zou het geen wetenschappelijke titel laten schrijven, ook geen journalistiek artikel en eigenlijk... studenten, toch ook niet je essay.

00:26:03

Gea: Nee-ja, dat is een mooie brug, namelijk het... We hebben natuurlijk ook als docenten dit allemaal voorbij zien komen, uitgetoetst en misschien ook al wel studenten essays gezien, waarbij we hebben gedacht: Hm, heeft de student dit zelf geschreven. Wat moeten we hiermee in in het onderwijs? Maar misschien ook wel, wat kunnen we ermee?

00:26:22

Piek: Ja, ik denk dan eigenlijk: wat kunnen we ermee. Ja, wat er niet meer moet is studenten essays laten schrijven en zo maar zeggen van oké. Ik ga nu het essay op zich evalueren, maar.

00:26:32

Gea: Moet dat mensen überhaupt nog leren schrijven?

00:26:35

Piek: Nou ja, ze misschien leren schrijven met deze tools aan de hand van deze tools dus hoe schrijf ik een artikel dat ik, waar ik zelf achter sta, terwijl ik deze tools gebruik? Het is een beetje de vraag van wat het, wat het doel was van het van het artikel. Kijk, als het doel is van toetsen of jij in staat bent om een wetenschappelijk artikel te schrijven, echt in termen van kwaliteit en stijl van schrijven, dan ja, dat is misschien iets dat we niet meer doen, net zo goed als dat we ook geen handschriften meer toetsen.

00:27:09

Gea: Goed, er zijn mensen die teruggaan naar handgeschreven tentamens.

00:27:13

Piek: Ja, dat is een manier, maar ik durf te worden dat dat daar ook een ChatGPT gaat komen, die handschriften gaat genereren, dat kan hij waarschijnlijk nu al. Dat zou ik niet eens van opkijken eigenlijk gezegd. Maar ja en mondelingen toetsen toch, ja. Maar je mag van mij best wel een artikel laten genereren. Maar het lijkt mij heel moeilijk als het gaat om de inhoud van wat achter het artikel zit, A: of ChatGPT in staat, is om dat juist te doen en B: vind ik dat je misschien meer moet doen dan alleen maar dat inleveren en nakijken. Misschien moet je een mondelinge toets hebben.

00:27:47

Gea: Wat jij zegt is: je moet heel goed nadenken over wat is nou eigenlijk het doel van de opdracht die je studenten geeft? En hoe toets je dat vervolgens? Maar maar een soort standaard opdracht schrijf een essay of of schrijf een samenvatting, dat gaan we niet meer doen. Is het eigenlijk, want er zijn van die ChatGPT detector of AI detectors, werken die? Kunnen we op één of manier achterkomen dat een tekst van een AI is?

00:28:15

Piek: Ja, maar dat is... ChatGPT/OpenAI heeft zelf zo'n service, om te detecteren of iets door AI gegenereerd is, en zij gaan die data weer gebruiken om te verzorgen dat het nog moeilijker te onderscheiden is, nog natuurlijker wordt.

00:28:31

Gea: Kun jij het zien als mens?

00:28:33

Piek: Nou, d'r zijn sommige dingen waarvan ik weet dat, daar kun je naar kun je op checken op dit moment en één van de dingen is godzijdank is de meeste data Engels en zit er maar heel weinig Nederlands bijvoorbeeld, en wij weten dat het Engels-achtige, grammaticale structuren gebruikt voor Nederland, die wij net niet zo gebruiken.

00:28:54

Gea: Ja.

00:28:55

Piek: Maar ja, dan moet je dan maar toevallig tegenkomen.

00:28:57

Gea: Ja, wat vind je eigenlijk sowieso van de output want, is het... schrijft ie goed?

00:29:06

Piek: Ik vind het een hele vervelende stijl die die hanteert. Een beetje van dat hyper correcte, altijd vriendelijke, behulpzame AI... dat gaat je toch een beetje tegenstaan hoe meer je het gebruikt, eerlijk gezegd.

00:29:22

Gea: Hij is ook niet heel beknopt.

00:29:24

Piek: De beknoptheid is inderdaad een probleem, en zelfs als we heel er duidelijk zeggen dat het beknopt moet zijn; je mag niet meer dan zoveel zinnen of woorden gebruiken, dan gaat ie daar toch nog af en toe de fout in. Ja, en het is wel zo, vind ik, dat het één van de meest indrukwekkende dingen is, wel, dat als je het weet, heel goed allerlei de juiste stijl te vinden bij bepaalde soorten... Laat ik zeggen, als je bijvoorbeeld bepaalde gemoedstoestand zou beschrijven, dan weet ie daar bijvoorbeeld de hele passende stil bij te vinden. Dus het stilistische talent is best wel groot en dat kun je aanprikken of aansturen om dat naar boven te krijgen.

00:30:10

Gea: Je zegt dus eigenlijk van: we moeten leren samenwerken met met de AI tools, verbieden heeft niet zoveel zin; We moeten er toch op één of andere manier mee leren werken. Ook de studenten: Moeten we ze een soort extra training gaan geven. Zo van, hoe herken je een AI?

00:30:24

Piek: Ik denk zeker studenten sowieso moeten weten wat de eventuele fouten en problemen zijn, maar ook wat de mogelijkheden zijn van de tool. De kwestie van of iets door een AI gegenereerd is of door een mens: Ik denk dat er een veel groter maatschappelijk probleem is dan dat dat een probleem is voor studenten. Eerlijk gezegd.

00:30:44

Gea: Ja, in zekere zin kunnen we natuurlijk nog wel een beetje sturen wat ze wel en niet gebruiken, of of wat ze ermee doen, waar ze het voor inzetten.

00:30:52

Piek: Ja, sommige studenten zullen geïnspireerd zijn om daarmee later te frauderen en maar ik neem aan de meeste niet.

00:31:00

Gea: Er zijn ook studenten die zeggen: Ik kan het veel beter dan ChatGPT, dus waarom zou ik zo'n programma vragen om het te doen?

00:31:06

Piek: Ja, goed, ik denk dat wij ook ethische discussies moeten hebben, maar dat geldt ook voor als je, als je bijvoorbeeld studenten leert hoe computervirus werken, dan weten de studenten ook hoe ze een computervirus moeten maken. Toch?

00:31:22

Gea: Even nog vooruitkijken: Over tien jaar, schrijven studenten dan nog werkstukken of doen we alleen mondelinge tentamens dat?

00:31:34

Piek: Ja, nou, dat is een beetje de output; de vorm van de output zal zeker anders zijn. Ja, ja, ik weet niet, misschien is het niet zo. Misschien is het wel eerder de docent vervangen door AI eerlijk gezegd, dus misschien toetst de AI wel de student, of de student conceptueel de kennis heeft en het begrip heeft.

00:31:57

Gea: Het schijnt dat dat je ook hele goeie rubric kan maken met AI.

00:32:02

Piek: Dat geloof ik zeker. Ja, ja, ja, en ik denk... Dat is natuurlijk het belangrijkste, wat je wil. Je wil dat studenten bijvoorbeeld de juiste mix hebben van zowel een stuk kennis die je die je moet hebben, want je kunt niet alleen maar alles opzoeken, je moet ook een stuk kennis zelf hebben en het begrip hebben, maar ook in staat zijn om juist scherpe analyses te kunnen maken, aan creatief oplossend vermogen hebben, voor als zich iets voordoet waarvoor geen antwoord is. Dus stel, niemand heeft het antwoord erop. Hoe vinden we dan het antwoord? En ik denk dat is wat je moet leren. Uiteindelijk. Dus het probleemoplossend vermogen is het allerbelangrijkste, dus niet waar vind ik het antwoord, maar hoe vind ik zelf het antwoord?

00:32:45

Gea: Ja, dus leerdoelen heel duidelijk maken, niet alleen maar het schrijven van de tekst, maar wat komt daar allemaal bij kijken.

00:32:52

Gea: En wat zijn de ontwikkelingen op het gebied van AI die jij de komende maanden vooral in de gaten gaat houden?

00:32:58

Piek: Nou, de komende weken moet ik zeggen: het gaat bijna per week. Dan krijg je weer, "er is weer een ander model", je kunt er niet tegenop lezen.

00:33:12

Gea: Misschien toch maar GPT vragen om samenvattingen te maken.

00:33:13

Piek: Nou ja vooral, je moet vooral die ChatGPT vragen iets over zichzelf te vertellen

want dan wordt die lyrisch over zichzelf. Daar hebben ze kennelijk ook op gefinetuned.

00:33:29

Gea: Blijft dit in dit tempo doorgaan, denk je?

00:33:31

Piek: Ik heb geen idee, want ik ik hoop van niet.

00:33:37

Gea: We zijn er nog niet van af, voorlopig.

00:33:40

Piek: Nee, nee.

00:33:42

Gea: Dank je wel voor dit gesprek.

00:33:44

Piek: Graag gedaan.

00:33:44

Gea: Ik heb deze podcast kunnen maken dankzij een subsidie van het BKO SKO alumni fonds van het centrum voor teaching en learning.

00:34:01

Gea: Verder hebben we aan deze podcast meegewerkt: Jens Branum, Abby Gambrel, Joost Canters, Megan van der Vorst en een heleboel AI tools.