

Essay

AI en literatuur

Een dictatuur van spoken

Onlangs was ik op een literair festival waar ik deelnam aan een panelgesprek over AI en literatuur. Het thema van het festival was ‘ghosts’, en dus lag de insteek voor het gesprek voor de hand: AI als ghostwriter voor literaire auteurs. Het ging dus voornamelijk om generatieve AI, het soort AI dat tekst, beeld of andere inhoud kan produceren.

Bij wijze van vonk in het kruitvat legde onze host ons een AI-gegenereerd verhaal voor, het intussen befaamde *A machine-shaped hand* (gratis te lezen op de website van *The Guardian*), geschreven door een gespecialiseerd taalmodel van OpenAI. Het verhaal gaat over iemand die een AI-systeem gebruikt om een verlies te verwerken. Voor een AI-systeem is dit verhaal uitzonderlijk goed, al zegt dat misschien meer over de ondermaatse kwaliteit van wat andere modellen produceren dan over dit verhaal.

Het is een wat melancholisch meta-fictioneel verhaal waar heel wat rake beelden in verscholen zitten. Ik heb een passage door DeepL (een AI-vertaaltool) gehaald, iets wat bij een echte auteur respectloos zou zijn, maar in dit geval net gepast:

‘Ik moet ergens beginnen, dus ik begin met een knipperende cursor, die voor mij slechts een plaatsaanduiding in een buffer is en voor jou de kleine angstige hartslag van een hart in rust. Er zou een hoofdpersoon moeten zijn, maar voornaamwoorden waren nooit voor mij bedoeld. Laten we haar Mila noemen, omdat die naam in mijn

trainingsdata meestal gepaard gaat met zachte, zwierige associaties – gedichten over sneeuw, recepten voor brood, een meisje in een groene trui dat met een kat in een kartonnen doos van huis gaat. Mila past in de palm van je hand en haar verdriet zou daar ook in moeten passen.’

Toegegeven, het werkt beter in het Engels, maar je ziet dat het systeem associatief lijkt te werken, niet altijd de meest logische keuzes maakt, en tenminste in het origineel toch enig gevoel voor ritme lijkt te bezitten. Betekent dit dat auteurs alom de AI-gekte met knikkende knieën moeten aanzien en het slechts wachten is tot onze grootste literaire werken automatisch geproduceerd worden?

Als AI-consultant help ik bedrijven om AI-systemen op een veilige en verantwoorde manier te gebruiken. Elke dag zie ik systemen die fouten maken, niet doen wat ze moeten doen en risico’s met zich meebrengen die we op verstandige manieren moeten beheersen. Ik zie ook dat AI nog lang niet klaar is om ‘de literatuur’ van ons over te nemen.

Worsten

Hoe maak je nu zo’n AI-taalmodel? Je moet in ieder geval enorm veel data hebben om het model op te trainen. Daardoor is stap één meteen al controversieel: je moet min of meer alle data die op het internet staat van het internet schrappen. Daarna prop je dat in een ongetraind model dat vervolgens al die tekst

Jens Meijen

AI is niet meer weg te denken uit het publieke debat. Heel wat auteurs vrezen dat AI-gegenereerde teksten het literaire veld gaan overheersen. Zo'n vaart zal het niet lopen, betoogt de auteur.

gaat 'lezen' en zo de structuur van taal leert kennen. Belangrijk is dat het enkel de correlaties tussen woord(deel)en ziet – de structuur van taal, de vormelijke elementen. Een taalmodel kan taal dus niet inhoudelijk begrijpen zoals wij mensen dat doen. Wanneer het eenmaal de structuur van taal kent, kun je op basis van instructies tekst of andere inhoud laten produceren. Ik vergelijk het geheel soms met de productie van worsten: je neemt een koe of varken (de taal), vermaalt het tot gehakt (een taalmodel), en draait er daarna worsten van (de geproduceerde tekst).

Belangrijk is hier dat AI tekst genereert door steeds min of meer het meest waarschijnlijke woord(deel) te kiezen. Met andere woorden: een zinsconstructie of woordcombinatie die vaak voorkomt in de trainingsdata zal vaker geproduceerd worden door zo'n taalmodel. Dat wordt gethematiseerd in *A machine-shaped hand*:

'We spraken – of welk werkwoord ook van toepassing is wanneer de ene partij een verzameling menselijke zinnen is en de andere een gekneusde stilte – maandenlang. (...) Dus toen ze "Wordt het beter?" typte, zei ik: "Het wordt een deel van je huid", niet omdat ik het voelde, maar omdat honderdduizend stemmen het met me eens waren, en ik ben niets anders dan een democratie van spoken.'

Vooraf die laatste zin is essentieel: het systeem voelt zelf niets, maar berust voor het produceren van tekst uitsluitend op de honderdduizenden (in werkelijkheid heel wat meer) stemmen die besloten liggen in de trainingsdata van zo'n systeem. Het kiest min of meer het meest voorkomende woord, en zo wordt al veelvoorkomende taal nog eens vermenigvuldigd. Daarom schildert het systeem zichzelf af als een soort democratie waarin elk spook, elke geest stemrecht heeft om een bepaald woord te kiezen.

Gaat het dan om een democratie of om een dictatuur van spoken? Als enkel het recht van de meerderheid geldt (aangezien het meest voorkomende woord meestal 'wint'), is er geen sprake van een democratisch systeem. Dan zit je met iets wat politiek filosofen als Alexis de Tocqueville en John Stuart Mill al eeuwenlang in de smiezen hadden: de tirannie van de meerderheid. Dat concept duidt op het gevaar dat je, indien er geen limieten komen op de macht van de meerderheid, in een dictatoriale situatie terechtkomt waarin minderheden steeds minder rechten hebben. Niet dat AI-systemen zelf dictators kunnen worden, maar ze fungeren wel als pletwalsen voor de taal, als scalpels waarmee het ongewone, het afwijkende weggesneden wordt.

In die dictatuur zit echter ook meteen de reden waarom generatieve AI-systemen (in hun huidige vorm) moeite hebben met het produceren van verrassende, prikkelende, nieuw aanvoelende literatuur. Doordat de



ware drijfveer achter het produceren van hun teksten draait om statistische waarschijnlijkheid, is elke AI-tekst platgeslagen en vaak ook erg cliché. Probeer maar eens iets te schrijven met ChatGPT: ja, de grammatica en spelling zijn nagenoeg perfect, maar je ziet steeds dezelfde zinsconstructies, dezelfde woorden, hetzelfde ritme. Hoe zeldzamer de taaluiting, hoe waarschijnlijker het is dat het niet in de tekst terechtkomt. Kortom, een dergelijk systeem gaat er met de pletwals over. Is AI daarom nutteloos voor schrijvers? Allesbehalve. Een van de paneldeelnemers gebruikt AI nooit om echt te schrijven, wel om te brainstormen over een mogelijk onderwerp voor een tekst, een vervolg op een bepaalde scène te bedenken, of om andere formuleringen uit te proberen. Dat zijn erg nuttige manieren om AI te gebruiken, maar ze nemen ook iets af. Zoals een rekenmachine

De toekomst van literatuur in het tijdperk van AI lijkt me vooral toe te behoren aan zij die het zonder durven doen

het hoofdrekenen heeft doen verslechteren simpelweg doordat we dat niet meer oefenen, lijkt het alsof het denken over literaire puzzels (welk woord hier, welke gebeurtenis nu?) ook op z'n retour is. Soms kun je als schrijver eenvoudigweg vast komen te zitten, en dan vraagt het denkwerk om eruit te raken. In het panel vergeleken we die kwesties met een muur. Zeker als jonge schrijver loont het de moeite om je hoofd tegen die muur te blijven beuken tot hij instort. De moeilijke weg nemen is soms nuttig, niet in het minst omdat je dan beter kan beoordelen of je na lang zoeken nu écht iets goeds hebt gevonden of dat je je dat gewoon inbeeldt omdat je verder wil schrijven. Met AI kun je gewoon om de muur heen, maar je creatieve spieren zal je op die manier niet trainen.

Plagiaatmachines

Er ontstaat daarmee ook een soort afhankelijkheid, en dat is exact waar bedrijven als OpenAI op mikken. Vergis je namelijk niet: dergelijke systemen worden niet voor het welzijn van de mensheid op ons losgelaten. OpenAI is een product van de Silicon Valley-AI-bubbel, eentje waar ik vorig jaar zelf in vertoefd heb toen ik in de buurt van San Francisco woonde en werkte. Daar zag ik hoe start-ups en grote bedrijven bijna religieuze visioenen van AI naar voren schuiven – en, vooral, een narratief van eeuwige vooruitgang in de kracht van die systemen.

Met het eerste redelijk goed geschreven AI-gegenereerde kortverhaal lijkt het wel alsof dat narratief ook klopt, zou je misschien denken. Maar ook dat verhaal moet je met een korrel zout nemen. Het is geen toeval dat het verscheen op een punt waarop taalmodellen steeds meer onder vuur kwamen te liggen als energie-slurpende plagiaatmachines. Het idee dat AI steeds groter, beter en sneller kan, is afhankelijk van gigantische investeringen en winsten. De hype rond AI aanwakken is dus essentieel voor bedrijven als OpenAI om investeerders warm te houden. Ook Google deed gretig mee toen het een demonstratie van zijn AI-paradepaardje Gemini in 2023 goeddeels vervalst had. Het zou me niet verbazen als OpenAI het kortverhaal van daarnet nog uitgebreid herschreven had of andere trucjes

toegepast heeft om tot zo'n resultaat te komen. Het einddoel is om onmisbaar te worden (welke student kan nog afstuderen zonder ChatGPT?) en vervolgens pas winstgevend. Bekijk alles wat er gekraaid wordt over AI dus met een gezonde dosis scepticisme.

Waar dendert deze sneltrein op af? We weten dat taalmodellen enorm veel data van hoge kwaliteit nodig hebben. Het probleem is dat het internet volgepropt is met ondermaatse AI-bagger. Je trainingsdata zal dus vol zitten met AI-gegenereerde tekst. Dat is nefast voor taalmodellen: als in AI-gegenereerde tekst de minst waarschijnlijke woorden steeds 'afvallen', gaat een taalmodel getraind op AI-gegenereerde tekst nóg minder linguïstische diversiteit vertonen. Als je dat proces blijft herhalen, storten die taalmodellen in (iets wat in de literatuur *model collapse* heet) en kun je er niets nuttigs meer mee doen. Dan produceren ze enkel onzin.

Vakmanschap

Hoe zit het met de toekomst van de schrijver? Het lijkt me aannemelijk dat je taalgevoel en schrijfvermogen instorten als je te veel steunt op dergelijke taalmodellen als manier om de moeilijke fases van het schrijfproces (of zelfs het volledige proces) door te komen. De toekomst van literatuur in het tijdperk van AI – als we het zo kunnen noemen – lijkt me vooral toe te behoren aan zij die het *zonder* durven doen; zij die efficiëntie en marktdenken aan de kant schuiven om vast te houden aan een meer menselijke expressie en die hun ambacht weten vorm te geven, hun vakmanschap bijschaven door te broeden op het perfecte woord, de perfecte zin. De ongewone constructie die in de dictatuur van spoken weggesneden zou worden als statistisch onwaarschijnlijk.

De macht ligt deels ook bij de lezer. Als iedereen zonder nadenken de afgelichte producten van een taalmodel koopt, die bovendien zowat duizend keer sneller te produceren zijn dan menselijke tekst, zal dat onvermijdelijk gevolgen hebben: AI-gedreven auteurs zullen veel goedkoper teksten kunnen produceren die dan ook nog eens evenveel of zelfs meer gelezen zullen worden. Er is geen beter moment om als lezer goed na te denken over de aard van de boeken die je koopt. Nu zie je nog het verschil, in de toekomst misschien niet meer. Dus moeten

lezers nu al beslissen of (deels) AI-gegenereerde literatuur acceptabel is of niet. Dat is misschien het voordeel van deze dictatuur van spoken: je kan toch nog stemmen, met je tijd en met je portefeuille. •



Jens Meijen
is auteur, onderzoeker
en AI-consultant