



Impact AI op wetenschappelijk onderzoek: groot en groeiend

Hoe disruptief is ChatGPT voor wetenschappelijk onderzoek?

In mei 2023 voorspelde Accenture¹ dat generatieve AI-tools zoals ChatGPT en Google Bard zo'n 40 procent van de werkuren in diverse sectoren kunnen automatiseren. 'Bijna alle banen worden beïnvloed – sommige zullen verdwijnen, de meeste functies zullen veranderen en er zullen veel nieuwe banen worden gecreëerd', aldus een conclusie van Accenture. De kracht van generatieve AI voor organisaties ligt echter niet eens zozeer in de tools zelf, maar in de mate waarin de organisaties in staat zijn de tools te integreren in de dagelijkse processen. Organisaties die nu stappen zetten om AI te integreren in hun bedrijfsprocessen én hun medewerkers trainen om samen te werken met de technologie, zullen een voorsprong hebben op concurrenten die minder creatief zijn.

& DOOR MAURICE MAGNÉE & JULES LANCEE

Generatieve AI heeft ook in toenemende mate impact op onderzoek en wetenschappelijke communicatie. Een survey² van Nature onder 1.600 onderzoekers toont dat ruim 25 procent van de onderzoekers AI gebruikt om manuscripten te helpen schrijven. Meer dan 15 procent gebruikt de technologie om subsidieaanvragen te schrijven.

Zoals veel onderzoekers weten, vereist een subsidieaanvraag, naast een plan van aanpak, ook een reeks minder aansprekende documenten, zoals impact statements, datamanagementparagrafen, rapportages en risicoanalyses. Recent zien we dan ook een scala aan AI tools ontstaan die inspelen op de behoefte deze werkzaamheden te automatiseren. Daarnaast kunnen de tools ook een faciliterende rol spelen in het peer review proces (waaronder een kwaliteitscontrole en statistiek check), al kunnen ze vooralsnog geen menselijke reviewers vervangen³.

Generatieve AI blijkt ook steeds beter in staat om complexe data zoals kwalitatieve gegevens

te verwerken en te analyseren. Software die bedoeld is voor het analyseren van interviews en literatuur, heeft bijvoorbeeld AI-assistenten⁴ ingebouwd voor het automatisch samenvatten en analyseren van deze data. Daarnaast kondigde OpenAI afgelopen maand aan gebruikers zelf de mogelijkheid te geven om AI-assistenten te creëren binnen de ChatGPT-omgeving. Bijvoorbeeld een eigen onderzoeksassistent die gespecialiseerd is in het type analyses die je uitvoert. Wetenschappers die deze tools gebruiken, zijn minder (kostbare!) tijd kwijt met voorheen tijdovende aspecten van het onderzoeksproces.

Handen in het haar

Onder uitgevers van wetenschappelijke journals is vooralsnog geen eenduidig beleid. Verantwoord gebruik van generatieve AI lijkt het toverwoord, al verschillen de uitgevers in de uitwerking van wat zij verstaan onder verantwoord gebruik. Science heeft in vernieuwd beleid aangegeven dat ChatGPT (of iedere andere AI tool) niet gebruikt mag worden voor het genereren van tekst, figuren en afbeeldingen. Nature staat geen AI toe in beeldgebruik, maar is genuanceerder over de voordelen en beperkingen in het gebruik van AI.

Nadelen die aan het gebruik kleven, zijn gebrek om de authenticiteit van de inhoud te verifiëren, geen erkenning van originele auteurs, schending van auteursrecht en het verspreiden van valse informatie. Juist eerlijkheid, transparantie, openheid, nauwgezetheid en objectiviteit zijn belangrijke waarden voor wetenschappelijk onderzoek.

Ook het gebrek aan reproduceerbaarheid is hierbij heel belangrijk. Een probleem bij het gebruik van door AI gegenereerde tekst is dat AI niet verantwoordelijk te houden is voor het werk. De tools kunnen op deze manier het vertrouwen in de wetenschap schenden.

Potentiële gevaren

De opmars van AI in het wetenschappelijke domein is even inspirerend als ontvullend. De mogelijkheden lijken eindeloos, de valkuilen ook. Een van de grootste zorgen betreft de nauwkeurigheid van door AI gegenereerde informatie. Hoewel systemen zoals ChatGPT zijn getraind op uitgebreide datasets, kunnen ze onnauwkeurigheden reproduceren of zelfs nieuwe onnauwkeurigheden creëren. De plausibele output die past bij het patroon van de invoer kan een rooskleurig

maar misleidend beeld schetsen van bijvoorbeeld onderzoeksresultaten.

Uit de eerder genoemde survey van Nature blijkt dat 29 procent van de onderzoekers ChatGPT gebruikt om literatuur te vinden. Dat is reden tot zorg. Deze chatbot staat er inmiddels om bekend wetenschappelijke artikelen en citaten te presenteren die niet bestaan. Onderzoekers die niet goed getraind zijn in het juist gebruik van de tools, kunnen dus uiteindelijk juist hun tijd ermee verdoen en onjuiste informatie verspreiden.

Verzonden onderzoek

Het potentieel voor het fabriceren van compleet verzonden onderzoek is een ander zorgwekkend punt. Het gemak waarmee gegevens kunnen worden gegenereerd, gemanipuleerd en gepresenteerd, kan de deur openen voor frauduleuze praktijken die de integriteit van het wetenschappelijk discours ondermijnen. In een tijdperk van misinformatie en verminderd vertrouwen in wetenschappelijke bevindingen is dit een serieus risico. Er zal geïnvesteerd moeten worden in tools om dergelijke fraude op te sporen.

Andere uitdagingen betreffen bias (vooroordelen) en privacy. AI-systemen staan er om bekend dat ze menselijke vooroordelen reproduceren, omdat die in trainingsgegevens van de technologie zitten. Zonder kritisch te zijn op de uitkomsten van generatieve AI (of tijdens de ontwikkeling ervan) kunnen vooroordelen onbewust tot uiting komen in bijvoorbeeld analyses van onderzoeksresultaten. Tenslotte kunnen er zorgen zijn over het gebruik van persoonlijke of gevoelige informatie in lerende systemen, zeker wanneer de werking van die systemen door de ontwikkelaar onzichtbaar wordt gehouden. Ook hier geldt dat onderzoekers zich er bewust van moeten zijn dat deze gevaren bestaan.

Hoe nu verder?

Om deze uitdagingen aan te pakken, zien we vanuit ons lectoraat een belangrijke rol weggelegd voor onderwijs. Onderwijsinstellingen en onderzoeksorganisaties moeten studenten en onderzoekers trainen in het effectief en ethisch gebruik van AI-tools. Met curriculum-aanpassingen om de technologische vooruitgang te weerspiegelen, kunnen jonge onderzoekers worden uitgerust met de vaardigheden om AI-tools kritisch te evalueren en te gebruiken. Met 'Future Campus⁵' en het Nationaal Groeifonds Programma 'Npuls⁶' hebben we belangrijke landelijke initiatieven om de digitale vaardigheden van studenten, docenten en onderzoekers in mbo, hbo en wo te vergroten. Hoewel de verantwoorde adoptie van AI onderdeel is van de programma's, zal echte verankering in visie en beleid van onderwijsinstellingen een stug vraagstuk zijn.

Ten tweede is transparantie cruciaal. Een editorial van het tijdschrift 'Accountability in Research'⁷ geeft enkele suggesties om verantwoord gebruik van AI te duiden: wees transparant over het gebruik bij het schrijven van de tekst en bij het genereren van ideeën, accepteer de volledige verant-

woordelijkheid voor de accuraatheid van de tekst, en geef aan wie het heeft gebruikt, het tijdstip en de datum, de prompts die zijn gebruikt om de tekst te genereren, de secties die de tekst bevatten en ideeën in de tekst die het gevolg zijn van het gebruik van de AI. Iemand van vlees en bloed moet verantwoordelijkheid nemen voor de tekst, daarom benadrukken we transparantie, eerlijkheid en volledige openbaarmaking van het gebruik van dit soort mogelijkheden' (Hosseini et al., 2023).

Ook softwareontwikkelaars moeten transparant zijn over de werking van hun tools. Dit omvat het delen van informatie over de datasets om AI-systemen te trainen en het documenteren van de algoritmen die ten grondslag liggen aan deze tools. Transparantie bevordert vertrouwen en stelt de wetenschappelijke gemeenschap in staat om de kwaliteit en de betrouwbaarheid van onderzoek te evalueren.

Effectieve wetgeving

Ten derde beginnen overheden AI-technologieën te reguleren, maar effectieve wetgeving laat nog op zich wachten. De 'AI-Act' van de Europese Unie (in onderhandelingsfase) eist transparantie, zoals onthullen welke inhoud AI-gegenereerd is. Digitale watermerken die de oorsprong van een tekst, afbeelding of video identificeren, kunnen een werkbaar reguleringsmechanisme zijn. De snelle ontwikkeling van AI-technologie kan er wel toe leiden dat regels al verouderd zijn tegen de tijd dat ze officieel beleid worden. Het beheersen van ontwikkelingen in AI vereist om deze reden een continu en lerend proces.

Tot slot is het van belang dat onafhankelijke onderzoekers het heft in handen nemen bij het verbeteren van de veiligheid van generatieve AI tools. AI is een miljardenindustrie: de belangen zijn te groot om te vertrouwen op zelfregulering van bedrijven. Samenwerking met onderzoekers vereist dat AI-ontwikkelaars zelf de algoritmes en training data beschikbaar moeten stellen aan een onafhankelijk wetenschappelijk instituut en dat zij inzicht krijgen in hoe de ontwikkelaars eventuele biases in de data hebben aangepakt. Een groep onderzoekers van de Universiteit van Amsterdam publiceerde recent hierover een reeks 'levende richtlijnen'⁸, met aandacht voor AI-geletterdheid, zodat onderzoekers, maar ook het bredere publiek veilig en ethisch generatieve AI tools kunnen gebruiken.

Kruispunt

In de snel evoluerende AI-wereld bevinden we ons op een kruispunt waar technologische innovatie en de wetenschappelijke traditie samenkomen. De inmenging van AI in wetenschappelijk onderzoek brengt ongekende mogelijkheden, maar onthult ook uitdagingen en gevaren die de kernwaarden van de wetenschap kunnen ondermijnen.

De oplossingen - onderwijs, transparantie, wetgeving en een actieve samenwerking tussen commercie en onderzoek - weerspiegelen a) reacties op de opkomst van AI en zijn b) een uitdrukking van de fundamentele wetenschappelijke houding, waarin

transparantie en objectiviteit centraal staan. Vanuit het Lectoraat Technologie voor Gezondheid testen we momenteel de eerste AI-assistenten voor wetenschappelijk onderzoek en brengen we de kansen en uitdagingen verder in kaart.

Disclaimer: AI is in deze tekst gebruikt voor het genereren van ideeën, vertalen en samenvatten van teksten en ondersteunen bij de opbouw van de verhaallijn (ChatGPT-3.5 voor MM en ChatGPT-4 voor JL in de periode oktober-november 2023; prompts kunnen worden opgevraagd bij de auteurs). De auteurs nemen volledige verantwoordelijkheid voor de juistheid van de inhoud. ■



Maurice Magnée is Lector Technologie voor Gezondheid aan de HAN University of Applied Sciences.



Jules Lancee helpt organisaties met de adoptie van generatieve AI.

Referenties

-
-
-
-
-
-
-
-



HAN University of Applied Sciences is ICT&health Innovation Partner.



HAN