Richtlijnen voor (Generatieve) Artificiële Intelligentie in onderzoek aan UAntwerpen

Bestuurscollege 18/11/2025

Onderzoek, Innovatie en Valorisatie Antwerpen
Dienst Onderzoeksaangelegenheden
Contact: Sarah Claes en Marianne De Voecht

Inhoudsopgave

1. Inleiding

Het gebruik en de ontwikkeling van (generatieve) Artificiële Intelligentie zorgt voor nieuwe perspectieven en mogelijkheden voor onderzoek, die we aan de UAntwerpen willen omarmen en integreren. Zo kunnen AI-tools onder meer ingezet worden voor brainstorming of als schrijfassistent om taalredactie te doen, vertalingen te maken en/of teksten samen te vatten, etc. Als creatieve assistent kunnen tools voor het genereren van afbeeldingen worden gebruikt voor het maken van visualisaties zoals diagrammen. Generatieve AI-tools kunnen auteurs ook ondersteunen bij het schrijven, valideren en of debuggen van code, bij data-analyse of als prereviewing-assistent om hun manuscript te beoordelen voordat het wordt ingediend.[1]

Het is echter belangrijk dat alle UAntwerpen onderzoekers (inclusief doctoraatsonderzoekers) vanuit de instelling voldoende ondersteund worden omcorrect en creatief gebruik te kunnen maken van tools die gebaseerd zijn op (generatieve) Artificiële Intelligentie. Daarbij is het van belang om op een wetenschappelijk integere manier met (generatieve) Artificiële Intelligentie om te gaan en steeds transparant te zijn over een eventueel gebruik in het onderzoek. Onderzoekers dienen ook stil te staan bij belangrijke ethische aandachtspunten en beperkingen van (generatieve) Artificiële Intelligentie. Tot slot is een grondige kennis van de mogelijkheden en beperkingen van generatieve AI-systemen essentieel. Denk bijvoorbeeld aan literatuurstudies waarin AI fictieve bronnen aanreikt, of aan feitelijke vergissingen wanneer een genAI systeem wordt ingezet als zoekmachine, terwijl de onderliggende trainingsdata mogelijk verouderd of onvolledig zijn. Alleen met een goed begrip van wat gen-AI systemen wel en niet kunnen, kunnen ze op een verantwoorde manier worden geïntegreerd in het onderzoeksproces.

2. Basisprincipes en wettelijk kader

De Universiteit Antwerpen ondersteunt het verantwoord gebruik van (generatieve) Artificiële Intelligentie in onderzoek mits je als onderzoeker hierbij steeds rekening houdt met een aantal basisprincipes en wettelijke bepalingen.

In het document Ethics guidelines for trustworthy AI van de Europese Commissie staan volgende basisprincipes geformuleerd:

Artificiële Intelligentie en het gebruik ervan moet steeds

1. wettig zijn, door te zorgen dat aan alle toepasselijke wet- en regelgeving wordt voldaan (zie punt 3. Wettelijk kader);
2. ethisch zijn, door te zorgen dat de ethische beginselen en waarden worden nageleefd. Betrouwbare AI baseert zich op volgende 4 ethische basisprincipes: menselijke autonomie, het voorkomen van leed, rechtvaardigheid, en verantwoording;
3. Robuust zijn, uit zowel technisch als sociaal oogpunt, aangezien AI-systemen ongewild schade kunnen aanrichten, zelfs al zijn de bedoelingen goed.

[1] De mogelijkheden voor het gebruik van AI-tools zijn overgenomen uit: How Can IJDS Authors, Reviewers, and Editors Use (and Misuse) Generative AI? | INFORMS Journal on Data Science en Arhiliuc, C. (1 maart 2024). *ChatGPT for research* [Powerpoint slides]. Faculteit Sociale Wetenschappen, Universiteit Antwerpen. Geraadpleegd op 29 maart 2024. ChatGPT for research - use cases (figshare.com).

Daarnaast werd in december 2023 de Europese AI Act goedgekeurd. Deze act past een *risk-based* benadering toe op de ontwikkeling van AI modellen[2]. De act benoemt vier ontwikkelingen die gecategoriseerd worden als '**onacceptabel risico'**[3]**:**

- Cognitieve gedragsmanipulatie van individuele personen of kwetsbare groepen.
- Sociale scoring waarbij mensen geclassificeerd worden op basis van gedrag, socio-economische status of persoonlijkheidskarakteristieken.
- Biometrische identificatie en categorisering van personen
- Biometrische identificatie in real-time of vanop afstand (bv. gezichtsherkenning)

Recent werd er ook een update voorzien van de Living Guidelines on the Responsible Use of Generative AI in Research. Deze lijsten eveneens enkele belangrijke verantwoordelijkheden van en aandachtspunten voor onderzoekers op wanneer zij generatieve AI gebruiken.

Kort samengevat wordt van jou als onderzoeker gevraagd bij het gebruik van AI steeds volgende vuistregel in het achterhoofd te houden: **hoe meer verantwoordelijkheid er bij het AI systeem wordt geplaatst, hoe meer menselijke controle er nadien nog is vereist.[4] De verantwoordelijkheid voor de correctheid en robuustheid van informatie ligt (nog) steeds bij jou als onderzoeker.**

3. Ethische uitdagingen en beperkingen van Artificiële Intelligentie

Indien je als onderzoeker bij de uitvoering van het onderzoek, het schrijven of editen van publicaties en projectaanvragen gebruik wenst te maken van AI-tools, dien je je bewust te zijn van enkele belangrijke ethische aandachtspunten en beperkingen van de huidige systemen. Net zoals de ontwikkelingen op het vlak van Artificiële Intelligentie zich in snel tempo opvolgen, zijn ook de huidige beperkingen ervan aan verandering onderhevig. In functie hiervan zullen deze richtlijnen ten gepaste tijde worden geüpdatet.

3.1.1 Bronvermelding

Bij door Artificiële Intelligentie gegenereerde teksten ontbreekt vaak de nodige bronvermelding of worden er door de AI-tool citaties bij verzonnen. Hierdoor is het niet altijd mogelijk om de inhoud van de teksten correct toe te schrijven aan de oorspronkelijke auteur en bestaat het risico op plagiaat en het schenden van intellectuele eigendomsrechten. Doordat het niet altijd mogelijk is de informatie te herleiden naar de oorspronkelijke bron, is het ook moeilijk om de correctheid van de gegevens te verifiëren. Het blijft daarom voor jou als onderzoeker belangrijk om de accuraatheid van de opgevraagde informatie te controleren en deze bij twijfel niet te gebruiken.

---

[2] Belangrijke kanttekening specifiek voor Onderzoek zoals vermeld in Recital 12c | EU Artificial Intelligence Act: "*This Regulation should support innovation, respect freedom of science, and should not undermine research and development activity. It is therefore necessary to exclude from its scope AI systems and models specifically developed and put into service for the sole purpose of scientific research and development. Moreover, it is necessary to ensure that the Regulation does not otherwise affect scientific research and development activity on AI systems or models prior to being placed on the market or put into service*."

[3] Er werd een compliance checker ontwikkeld om organisaties te helpen zich wegwijs te maken in de juridische verplichtingen van de AI act.

[4] https://research.kuleuven.be/en/integrity-ethics/integrity/practices/genAI#1

3.1.2 Gedateerde of foutieve informatie

De informatie die door generatieve AI systemen zoals Copilot of ChatGPT geproduceerd wordt, is niet altijd up to date, waardoor meer recente wereldwijde gebeurtenissen of specifieke ontwikkelingen binnen het onderzoeksveld mogelijks geen deel uitmaken van de aangeleverde informatie. Ook hier blijft het dus zaak voor jou als onderzoeker om de informatie te verifiëren en de stand van zaken binnen het onderzoek via academische literatuur en andere relevante bronnen in kaart te brengen.

Daarnaast kan het model, bij gebrek aan informatie, zelf aanvullingen creëren, waarbij namen, objecten en informatie verzonnen worden (i.e. 'hallucinaties').[5]

3.1.3 Reproduceerbaarheid

Het antwoord dat door generatieve AI-tools wordt gegenereerd op basis van een prompt, kan elke keer verschillen bij het opnieuw stellen van de vraag. Ook bij het opnieuw genereren van een eerder antwoord kan de output verschillen ten opzichte van de vorige versie. Dit maakt het moeilijk om correct te kunnen verwijzen naar AI gegenereerde content, maar ook omgekeerd om aan te tonen dat er sprake was van door AI gegenereerde informatie.

3.1.4 Bias

Hoewel er hiervoor de nodige aandacht begint te ontstaan bij producenten van (generatieve) Artificiële Intelligentie, kunnen zij op dit moment niet garanderen dat de gegenereerde content volledig vrij is van bias of schadelijke informatie. Dit hangt samen met het feit dat het niet altijd eenduidig vast te stellen is op welke broninformatie het AI-systeem zich heeft gebaseerd. Het is dus essentieel voor jou als onderzoeker om hier waakzaam voor te zijn en deze bias niet over te nemen in het eigen onderzoek.

3.1.5 Intellectuele Eigendomsrechten

Zoals hierboven aangegeven, kan door het ontbreken van correcte bronvermelding het risico bestaan op het schenden van intellectuele eigendomsrechten die rusten op de broninformatie gehanteerd door generatieve AI-tools. Ook bij de input van data in het AI systeem dient hier aandacht voor te zijn. Als je als onderzoeker IP-gevoelige informatie opneemt bij het aanmaken van prompts, kan deze potentieel later worden hergebruikt als bron in door de AI-tool gegenereerde antwoorden. Dit is bij uitstek het geval bij het gebruik van gratis online versies; deze slaan ingevoerde data op en gebruiken ze voor verdere ontwikkeling van de applicatie. Op die manier geef je misschien gevoelige informatie vrij met valorisatiepotentieel of informatie waar de intellectuele eigendomsrechten bij derden liggen, bijvoorbeeld als het gaat om een projectaanvraag waar je als reviewer optreedt. Omgekeerd bestaat ook het risico dat het idee van je paper, ook al heeft het niet direct valorisatiepotentieel, wel naar derden doorgegeven wordt, en mogelijks aan andere gebruikers wordt voorgesteld. Om deze reden is het gebruik van gratis versies in een professionele context dan ook sterk afgeraden tenzij eventueel om een tool te testen met generieke informatie.

3.1.6 Privacy

Net zoals door AI gegenereerde informatie IP-gevoelige data kan bevatten, kan er ook potentieel privacygevoelige informatie worden opgenomen in het antwoord op de prompt. Er wordt niet altijd

[5] Het verschil tussen menselijk schrijven en AI: lege doos vs. intentie* - bladspiegel (uantwerpen.be)

transparant gecommuniceerd door de eigenaars van AI-systemen over de oorsprong van de gebruikte leerdata noch over de opslagmodaliteiten hiervan. Het is dus niet altijd eenvoudig om vast te stellen of de gebruikte data in lijn met de Europese GDPR wetgeving werden verwerkt. Wees ook als onderzoeker zelf waakzaam welke persoonlijke informatie je opneemt in je prompts alsook bijlagen die je uploadt om potentiële privacy-schendingen te vermijden. Ook in dit geval wordt het gebruik van gratis versies in een professionele context sterk afgeraden tenzij eventueel om een tool te testen met generieke informatie.

3.1.7 Auteurschap

De Universiteit Antwerpen heeft institutionele richtlijnen voor auteurs van wetenschappelijke publicaties. In deze richtlijnen wordt geduid dat elke auteur die op een publicatie wordt vermeld, verantwoordelijkheid dient op te nemen voor wat betreft de inhoud en integriteit van de publicatie. Ondanks het genereren van inhoud, kunnen AI-tools nooit de verantwoordelijkheid opnemen hiervoor. Zij kunnen bijgevolg dan ook niet als auteur worden opgenomen op wetenschappelijke publicaties. Dit neemt niet weg dat het gebruik van AI-tools voor het uitwerken van een publicatie dient te worden erkend (zie onder).

3.1.8 Impact op het milieu

Het gebruik van artificiële intelligentie heeft ook ecologische en maatschappelijke implicaties. Grote AI-modellen vereisen aanzienlijke rekenkracht, wat gepaard gaat met een hoge energieconsumptie. Weeg tijdens de uitvoering van je onderzoek daarom steeds het gebruik van AI-toepassingen af tegen hun milieu-impact en kies waar mogelijk voor energie-efficiënte alternatieven of lokale infrastructuren.

3.1.9 Billijkheid

Toegang tot krachtige AI-tools, data-infrastructuren en expertise is wereldwijd ongelijk verdeeld. De sterkte van een AI-toepassing kan zelfs afhankelijk van de gehanteerde taal verschillen, alsook de kostprijs (indien er per token wordt gerekend). Dit risico op het versterken van bestaande ongelijkheden geldt zowel binnen als tussen onderzoeksinstellingen, disciplines en regio's. Verantwoord AI-gebruik houdt dus ook in dat je als onderzoeker bewust nadenkt over inclusie, toegankelijkheid en het vermijden van structurele uitsluiting.

4. Controle en erkenning van gebruik

Wanneer onderzoekers AI-tools gebruiken om enkele stappen in het onderzoeksproces te vereenvoudigen, zoals bijvoorbeeld codering, literatuurstudies, tekstnazicht, etc., is het belangrijk dat zij volgende vuistregel in het achterhoofd houden. Hoe meer verantwoordelijkheid er bij het AI-systeem wordt geplaatst, hoe meer menselijke controle er nadien nog is vereist.[6] De verantwoordelijkheid voor de correctheid en robuustheid van informatie ligt nog steeds bij de onderzoeker.

Naast verificatie van de aangeleverde informatie, is het belangrijk dat onderzoekers het gebruik van AI-tools erkennen in het onderzoek. Zoals eerder aangegeven, kunnen AI-tools niet als auteur vermeld

[6] https://research.kuleuven.be/en/integrity-ethics/integrity/practices/genAI#1

worden op een publicatie. Wel dient in de methodologiesectie van de publicatie, presentatie of doctoraatsproefschrift het gebruik van AI-tools te worden vermeld.

5. Te vermijden gebruik

Rekening houdende met bovenstaande en de algemene normen van de goede onderzoekspraktijk meent de Universiteit Antwerpen dat het volgende gebruik van (generatieve) Artificiële Intelligentie sterk vermeden moeten worden:

- Het creëren van concrete inhoud van publicaties of projectaanvragen zonder grondige factchecking en verdere inhoudelijke bewerking
- Peer review van publicaties of projectaanvragen van anderen[7]

Het is in deze belangrijk om ook steeds te kijken of het gekozen tijdschrift, uitgever en/of onderzoeksfinancier er specifieke voorwaarden op nahoudt inzake het gebruik van artificiële intelligentie en je eveneens met deze in lijn te stellen.

6. Ontwikkeling van Artificiële Intelligentie door UAntwerpen onderzoekers

Naast het gebruik van (generatieve) AI-tools in onderzoek, zijn er ook onderzoekers binnen de universiteit die zelf Artificiële Intelligentie toepassingen ontwikkelen. Naast het belang van menselijke supervisie en controle, is het essentieel dat onderzoekers de risico's voor verkeerd gebruik of *misuse* van deze toepassingen voldoende kunnen beperken.[8] Onder *misuse* verstaan we het mogelijk misbruik van onderzoeksresultaten voor onethische doeleinden. In uitzonderlijke gevallen kan dit ook van toepassing zijn op de ontwikkeling van Artificiële Intelligentie.[9]

De ontwikkeling van AI technologie staat op de lijst van kritieke technologieën die werd opgesteld door de Europese Commissie.[10] De EC raadt extra aandacht en een grondige risicobeoordeling aan voor deze kritieke technologieën.In geval van een verhoogd risico, of bij vereisten gesteld door de onderzoeksfinancier of een tijdschrift, staat de Ethische Commissie voor Misuse, Human Rights & Security (MiHRS) in voor het reviseren van ingediende aanvragen en leveren van advies. De MiHRS commissie kan bijstand leveren in het voorzien van mitigerende maatregelen indien er risico's op vlak van misuse, mensenrechten en/of kennisveiligheid worden vermoed.

Daarnaast werd binnen de Universiteit Antwerpen in 2023 het Antwerp Center on Responsible AI opgericht. De missie van het centrum is het bevorderen van de ontwikkeling, toepassing en verspreiding van verantwoorde AI in een breed scala aan domeinen, waaronder gezondheid, belastingen, verzekeringen, enz. Hier kunnen onderzoekers van de instelling terecht voor vragen

[7] Dit sluit nauw aan bij de richtlijnen van de Europese Commissie omtrent Horizon Europe en AI, zie AI has a place in research, but not in evaluation of Horizon Europe proposals, Commission says | Science|Business (sciencebusiness.net)
[8] Met uitzondering van de verboden ontwikkeling zoals vermeld in punt 2. Basisprincipes en wettelijk kader
[9] Voor meer info over ethisch verantwoord design kan je terecht op ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf (europa.eu)
[10] Recommendation on critical technology areas

omtrent de verantwoorde ontwikkeling van AI. Ook bij TextUA, het Antwerp Text Mining Center, kan je terecht voor training, consultancy of met vragen omtrent onder meer AI in onderzoek.

7. AI Officer en inventaris AI-toepassingen

Vanaf 1 oktober 2025 startte aan de Universiteit Antwerpen de AI Officer met specifieke opdracht om samen met de interne stakeholders het AI-beleid verder uit te werken. Daarnaast is eveneens de taak om te zorgen dat de Universiteit Antwerpen voldoet aan de vereisten van de AI Act waarvan boven al sprake. Dit houdt o.a. in dat er voorzien wordt in sensibilisering, informatiecampagnes en een opleidingsaanbod. Daarnaast zal ook een AI Inventaris opgebouwd worden, eveneens een verplichting in kader AI act.

Op de Pintra subsite Infocenter AI wordt informatie samengebracht om medewerkers te ondersteunen. Er is eveneens een mailadres infocenter.ai@uantwerpen.be.

8. Bronnen

Arhiliuc, C. (1 maart 2024). *ChatGPT for research* [Powerpoint slides]. Faculteit Sociale Wetenschappen, Universiteit Antwerpen. Geraadpleegd op 29 maart 2024. ChatGPT for research - use cases (figshare.com)

https://research.kuleuven.be/en/integrity-ethics/integrity/practices/genAI#1

AI has a place in research, but not in evaluation of Horizon Europe proposals, Commission says | Science|Business (sciencebusiness.net)

2023_RESEARCH_AI_Richtlijnen_ENG | Vrije Universiteit Brussel (vub.be)

ALLEA | All European Academies

Ethics guidelines for trustworthy AI | Shaping Europe's digital future (europa.eu)

EU AI Act: first regulation on artificial intelligence | News | European Parliament (europa.eu)

Generative Artificial Intelligence (AI) | Harvard University Information Technology

Guidance on the use of AI in research – @theU (utah.edu)

How Can IJDS Authors, Reviewers, and Editors Use (and Misuse) Generative AI? | INFORMS Journal on Data Science

Kenniscentrum Data & Maatschappij – Home (data-en-maatschappij.ai)

Living Guidelines on the Responsible Use of Generative AI in Research

Publication and authorship | Research Support (ox.ac.uk)

Recommendation on critical technology areas

The Act Texts | EU Artificial Intelligence Act