Verantwoord gebruik van Generatieve Artificiële Intelligentie aan AMA

Met de opkomst van tools die gebruik maken van generatieve artificiële intelligentie (GenAI) ontstaat een breed gamma van mogelijkheden voor creatie en verbetering van onder andere teksten, codes, audio- en videomateriaal of een combinatie ervan. AMA staat open voor het gebruik op een verantwoorde manier van deze technologie en geeft via deze beleidstekst info aan studenten, docenten en onderzoekers.

De duidelijke en uitgebreide richtlijnen rond het gebruik van GenAI die opgesteld werden door de collega's van de KU Leuven vormen de basis van deze tekst (KU Leuven, 2023a, 2023b). We danken hen voor de toestemming om ze te gebruiken.

De tekst zal worden bijgewerkt op basis van voortschrijdend inzicht in lijn met de snel evoluerende technologie.

Wat is Generatieve AI?

Generatieve Artificiële Intelligentie (GenAI) is een vorm van machinaal leren. Het is de verzamelnaam van algoritmes die in staat zijn om nieuwe content te creëren. Deze content kan verschillende vormen aannemen: tekst of code, maar ook beelden, video's en audio of een combinatie van dit alles.

GenAI genereert output als antwoord op een vraagstelling/prompt met behulp van generatieve modellen zoals Large Language Models (LLMs) en baseert zich hiervoor op grote datasets.

Enkele bekende voorbeelden zijn tekstgeneratoren zoals ChatGPT, de integratie van ChatGPT in Microsoft Bing en beeldgeneratoren zoals DALL-E en Midjourney.

Mag GenAI gebruikt worden aan de instelling?

De hogeschool kiest ervoor om GenAI verantwoord in te zetten. GenAI tools hebben immers hun weg gevonden naar het werkveld waar je als alumnus terecht zal komen. Het is belangrijk dat je als student begrijpt hoe GenAI werkt zodat je de academische standaarden kan handhaven en eigenaarschap kan behouden over alle geproduceerde werkstukken.

Deze beleidstekst geeft duidelijkheid over wat we verwachten met betrekking tot opdrachten, bachelor- of masterscripties of onderzoek.

Voor je begint …

- Deze beleidstekst is voor de hele hogeschool van toepassing. Waar nodig kan de docent bijkomende informatie geven.
- Er zijn **4 basisprincipes**:

Transparantie	over het gebruik van GenAI afhankelijk van het soort gebruik
Verificatie	van de correctheid van de gegenereerde output met aandacht voor correcte bronvermelding
Verantwoordelijkheid	voor het correct gebruik van GenAI (in eerste instantie als hulp en ondersteuning) en voor de output die je publiceert
Respect	voor persoonsgegevens en vertrouwelijke informatie door deze niet in te voeren op platformen die niet worden beheerd op eigen servers

- Overleg altijd vooraf met de docent of begeleider indien je voor de opdracht audio- of visueel materiaal zou willen genereren met GenAI. Motiveer in dat overleg waarom de manier waarop je GenAI wil gebruiken geen inbreuk is op het Onderwijs en Examenreglement (OER). Geef aan waarom het gebruik van GenAI noodzakelijk is, hoe je nog steeds eigenaar blijft van het werkstuk door zelf iets aan de GenAI-output toe te voegen en hoe je op die manier de leerdoelen bereikt. Wees daarboven steeds transparant over de eigen bijdrage versus de GenAI-output.
- Een duidelijk misbruik van GenAI - waarbij eender welke output grotendeels door GenAI gegenereerd is en je niet **transparant** was over het inzetten van de tool kan volgens het OER als **examenfraude** beschouwd worden.

Dit zegt het Onderwijs en Examenreglement

3.5. Examenfraude
Artikel 57 – definitie

§1 Als examenfraude wordt beschouwd: elk gedrag van een student in het kader van een examen waardoor deze het vormen van een juist oordeel omtrent kennis, inzicht en/of vaardigheden van hemzelf dan wel van andere studenten geheel of gedeeltelijk onmogelijk maakt of poogt te maken, zelfs als de fraude of onregelmatigheid pas achteraf wordt vastgesteld.
§2 Zo wordt het in het bezit zijn van ongeoorloofde middelen in de examenruimte (cursussen, notities, niet geregistreerde rekenmachines, elektronische toestellen zoals GSM's, notebooks, laptops, smartwatch, …) beschouwd als examenfraude. Deze voorbeelden zijn niet exhaustief.
§3 **Onder examenfraude wordt ook verstaan**: het zonder duidelijke bronvermelding geheel of gedeeltelijk overnemen van gegevens bij schriftelijke opdrachten, zoals eindscripties of projecten (het zogenaamde plagiaat) en **het ongeoorloofde gebruik van artificiële intelligentie**.

Hoe kan je GenAI gebruiken?

Een GenAI model kan voor diverse werkstukken gebruikt worden. Bepaald gebruik lijkt heel hard op wat we al doen met andere tools en vraagt niet om extra transparantie. Ander gebruik kan je eigen inbreng mogelijks verdoezelen en is daarom niet toegestaan op enkele uitzonderingen na.

Dit beleidsdocument vormt steeds de basis voor het gebruik van GenAI aan onze instelling. Een docent kan beslissen om af te wijken van dit beleidsdocument en een bepaald gebruik van GenAI niet toe te laten. Controleer altijd of de opdrachtomschrijving (extra) informatie geeft over het gebruik van GenAI.

GenAI: Wat is toegelaten?

- **GenAI gebruiken als taalassistent** voor het herwerken of verbeteren van zelf geschreven teksten, waarbij het model geen nieuwe inhoud toevoegt. Een generatief taalmodel kan gebruikt worden om een tekst die je zelf geschreven hebt te herschrijven en te verbeteren (en te vertalen). Dan is het gebruik gelijk aan de spelling- en grammaticacheckers die we vandaag ook al kennen en hoef je dat gebruik niet expliciet te vermelden.
- **GenAI gebruiken als zoekrobot,** als hulp bij het genereren van nieuwe ideeën**,** om eerste informatie te krijgen over een onderwerp of voor een eerste aanzet om op zoek te gaan naar literatuur. Deze vorm van informatie vergaren is gelijkaardig aan het gebruik van een gewone zoekmachine bij de uitwerking van een opdracht of onderzoek. We dringen aan tot voorzichtigheid bij dergelijk gebruik. Er zijn nadelen verbonden aan het gebruik van generatieve taalmodellen als zoekmachine. In de meeste gevallen is het onmogelijk om de bron van de informatie te achterhalen. Daarnaast blijft er het reële risico op onjuistheden en fictieve referenties. Ga vervolgens zelf op zoek naar wetenschappelijke referenties en voer zelf een analyse van de brondocumenten uit. Interpreteer, analyseer en verwerk de verkregen informatie; kopieer ze niet zomaar. Indien je vervolgens eigenhandig een tekst opstelt, hoef je het gebruik niet te vermelden.
  Let op! Wanneer je bepaalde delen van GenAI output toch letterlijk overneemt (bv. omwille van de aard van de opdracht), dan vermeld je je bronnen en citeer je.
- **Tekst- en beeldgeneratie voor bestaande kennis en nieuwe ideeën:**
  (bv. hulp bij het schrijven van een literatuurstudie of projectaanvraag)
  Een generatief taalmodel kan gebruikt worden om beschrijvingen te genereren van bestaande, algemeen bekende concepten en ook van nieuwe ideeën. We verwachten dat je de gegenereerde teksten op accuraatheid controleert en voorziet van passende, accurate referenties. In ieder geval blijf je verantwoordelijk voor de accuraatheid en samenhang van de tekst, evenals mogelijk verdoken plagiaat.
  Wees transparant over het gebruik van het generatief taalmodel, met inbegrip van referentie.
- **Generatie van synthetische data**, waarbij GenAI modellen bestaande datasets nabootsen. Op deze manier kunnen beperkingen zoals privacy issues en vertrouwelijkheid van reële datasets worden ondervangen. Het is belangrijk om de gegenereerde synthetische data zorgvuldig te evalueren op kwaliteit en mogelijke bias, aangezien de output sterk afhankelijkheid is van de kwaliteit van de data waarop de modellen getraind zijn. We verwachten in ieder geval transparantie: het gebruik van generatieve modellen voor de generatie van synthetische data dient expliciet vermeld te worden (met inbegrip van referentie).
- **Gebruik in de data analyse**: bv. toekomstige Microsoft 365 Co-pilot in Excel (en andere Office producten) om via prompts data te analyseren. We verwachten dat je expliciet refereert naar de gebruikte analytische methoden.

- **Visualisatie van onderzoeksresultaten:** multimodale modellen, die tekst- en beeldmateriaal combineren, bieden onder meer mogelijkheden om onderzoeksresultaten automatisch te visualiseren. We verwachten dat je zulke visualisaties grondig op onjuistheden controleert en het gebruik van het generatief model (met inbegrip van referentie) expliciet vermeldt.
- **GenAI gebruiken als programmeerassistent** voor het genereren van programmeercode. Vermeld steeds je bronnen wanneer je GenAI gebruikt om code te genereren. Verifieer ook of GenAI zich niet baseert op code met een softwarelicentie die niet overeenstemt met jouw gebruik.

GenAI: Wat is zeker niet toegelaten?

- **Elke vorm van letterlijke overname en kopiëren zonder volledige bronvermelding** (citeren, refereren) van eender welk materiaal gegenereerd door GenAI. Academisch werk vereist immers dat de originele bronnen worden nagetrokken. Gebruik van GenAI met bronvermelding kan slechts in uitzonderlijke situaties.
- **Elk gebruik van GenAI waarmee de docent je eigen kunnen niet meer kan natrekken en/of kan evalueren.** Bijvoorbeeld: gebruik van parafraseertools voor teksten die je niet zelf geschreven hebt om plagiaat te verdoezelen, misbruik van vertaalsoftware.
- **Elk gebruik van GenAI tijdens examens op campus.**

Refereren en citeren van GenAI

Hoewel GenAI geen specifiek citatieformaat vereist, is het belangrijk om alle informatie of ideeën die eruit worden gehaald correct toe te schrijven. De richtlijnen m.b.t. het citeren van Generatieve AI voor de verschillende refereerstijlen aan de instelling (APA, Chicago, numeriek) worden opgenomen in de documenten “vorm en bibliografie”, te vinden op Blackboard bij “research methodology”.

Hoe ben je transparant over het gebruik van GenAI?

Refereren (en citeren) is de eerste stap om transparant te zijn. 'Transparant zijn' is essentieel opdat een juist oordeel omtrent je kennis, inzicht en/of vaardigheden steeds mogelijk zou blijven. Hoe je transparant bent is echter moeilijk vast te leggen in algemene richtlijnen. Dit hangt meestal samen met het doel van het werkstuk en kan af te leiden zijn uit de opdrachtomschrijving.

Een goede werkwijze om transparant te kunnen zijn, is zelf bijhouden hoe en waarom je GenAI hebt gebruikt. Dat kan verschillende vormen aannemen, bijvoorbeeld:

- Bewaar de volledige uitwisseling met GenAI door middel van screenshots. Duid hierin eventueel relevante delen aan.
- Geef een toelichting van hoe GenAI werd gebruikt (bijvoorbeeld voor het genereren van ideeën, tekstfragmenten, langere stukken tekst, argumenten, evidentie, illustraties van begrippen, …)
- Noteer waarom GenAI werd gebruikt: om tijd te winnen, om een writer’s block aan te pakken, om het denken te stimuleren, om toenemende stress de baas te kunnen, om een begrip beter te begrijpen, om te vertalen, om te experimenteren met GenAI, …

Hoe je hierover rapporteert kan verschillen. Kijk na of dit wordt verduidelijkt in de opdrachtomschrijving van het werkstuk. Naast bronvermelding en opname in een referentielijst, kan het gewenst zijn om bovenstaande elementen toe te lichten onder het (materialen en) methoden-gedeelte of als een bijlage.

Indien onduidelijk is hoe je over gebruik van GenAI best kan rapporteren, ga in overleg met de verantwoordelijke docent of begeleider.

Detectie van GenAI

Tools zoals SafeAssign kunnen plagiaat opsporen. Daarnaast zijn er ook andere indicatoren (eigen aan hoe GenAI werkt) die kunnen wijzen op gebruik van GenAI.

Weet dat je steeds kan gevraagd worden om toe te lichten hoe (bepaalde delen van) een werkstuk of tekst tot stand zijn gekomen. Op die manier krijgt je de kans om meer duidelijkheid te scheppen en kan worden nagegaan hoe GenAI is gebruikt. Correct gebruik van GenAI en dit kunnen toelichten, is immers een essentieel deel van wetenschappelijk integer te werk gaan.

Tips & tricks voor een verantwoord gebruik van GenAI

- **Overtref GEN AI**
  GenAI-applicaties kunnen soms een hulpmiddel zijn voor werkstukken (visuele, schrijf- en programmeeropdrachten). We verwachten dat je er nog een extra laag aan toevoegt: qua argumentatie of onderbouwing, kritische analyse, volledige en correcte bronvermelding, creatieve inbreng, persoonlijke insteek of reflectie, vernieuwend karakter, connectie met maatschappelijke context, … Zorg ervoor dat je zelf het gesprek of debat kan voeren over je werkstuk.
- **Geef betekenis aan GenAI**
  Wanneer mensen een zin horen, koppelen ze daar een betekenis aan vast, en verbinden ze de taaluiting zo met de werkelijkheid. Een generatief taalmodel heeft enkel toegang tot de vorm, maar kan op geen enkele manier de verbinding met de werkelijkheid leggen. Je hebt dus ook geen enkele garantie dat de teksten van het systeem daadwerkelijk correct zijn. De enige bekommernis van het systeem is om teksten te genereren die zo plausibel mogelijk lijken; waarheid of werkelijkheid zijn daarbij van geen tel.
- **Verifieer GenAI**
  Maak gebruik van GenAI indien toegelaten, maar vertrouw technologie niet blindelings. Hoe meer verantwoordelijkheid je bij het systeem legt, des te meer verificatie, controle en verantwoording nodig is bij het resultaat. Soms lijkt AI-output heel overtuigend, maar het is perfect mogelijk dat het antwoord niet correct of zelfs verzonnen is (‘hallucinating AI’). In de meeste gevallen is het niet mogelijk te achterhalen hoe de algoritmes tot een bepaald resultaat zijn gekomen en is er geen transparantie over gebruikte bronnen ('black box'). De tools gebruiken niet altijd de meest recent beschikbare data waardoor de output achterop kan hinken. Controleer goed wat je

gebruikt en ga zelf op zoek naar bestaand bronmateriaal om naar te refereren. Je blijft namelijk inhoudelijk verantwoordelijk voor het afgeleverde werk.

- **Vermijd plagiaat door GenAI**
  Het is van ChatGPT geweten dat de meeste van de gegenereerde referenties fictief zijn. In de output van GenAI ontbreekt soms transparantie over de gebruikte bronnen, waardoor de kans toeneemt dat je plagiaat pleegt. Ook het risico op inbreuken op het auteursrecht is reëel. Ook GenAI bouwt voort op andermans werk, en daar hoort altijd een correcte bronvermelding bij.
- **Vermenselijk GenAI niet**
  AI-applicaties (chatbots in het bijzonder) zijn niet menselijk, ook al lijkt dat soms zo omdat ze heel interactief reageren. Ze worden niet beïnvloed door persoonlijke ervaringen of hun nabije omgeving. Ze kunnen wel via data ethische principes aangeleerd krijgen. Het is belangrijk je te allen tijde bewust te zijn dat de AI-applicaties slechts technologische hulpmiddelen zijn.
- **Wees bedacht op bias door GenAI**
  AI-applicaties zijn getraind met bepaalde datasets, die soms niet representatief zijn. Er is geen transparantie over de filtering die ze toepassen. Dit doet heel wat ethische vragen rijzen. Bij blindelings gebruik loop je het risico op verder verspreiden van incorrecte informatie, bepaalde stereotypen of vooroordelen.
- **Wees bedacht voor de slechte reproduceerbaarheid van de output van GenAI tools**
  De output die ChatGPT (of gelijkaardige tools) genereren kan bij iedere poging/vraagstelling verschillen. Eenzelfde vraag stellen op een ander moment of een lichtjes andere query kan leiden tot totaal verschillende antwoorden. Ook wanneer men na het krijgen van een antwoord op "Regenerate response" klikt, wordt er een ander antwoord gegenereerd.
- **Voer nooit privacygevoelige informatie in GenAI in**
  Er is vaak weinig transparantie over wat de eigenaars van de AI-toepassingen doen met de ingevoerde data. Veiligheid/vertrouwelijkheid van de ingevoerde data is niet gegarandeerd. Voer geen persoonsgegevens of vertrouwelijke informatie in deze GenAI-applicaties. Ga na of je beschikt over de nodige toestemming of licentie voor het ingeven van auteursrechtelijk beschermd materiaal. Indien je twijfelt over het al dan niet vertrouwelijke karakter van informatie, kan je dit navragen bij de verstrekker van de informatie. Vaak worden de data opgeslagen in niet-Europese cloud toepassingen die niet altijd GDPR conform zijn. Dit is anders wanneer de tools draaien op eigen servers.
  Een invoer van een vinding kan gelijk staan met openbaarmaking, overleg altijd met je promotor of begeleider. Meer informatie over de veiligheid van de technologie wordt soms aangeboden op de website van de GenAI technologie zelf.
- **Vermijd het energieverbruik door GenAI**
  Weet dat de energiekost van de servers voor AI-applicaties heel hoog is, gebruik ze dus alleen wanneer ze meerwaarde kunnen bieden.
- **Rapporteer transparant over GenAI**
  Alleen als je transparant bent over gebruik van AI-output en je iets van jezelf toevoegt, kan je laten zien dat je iets geleerd hebt. Weet dat je kans loopt op een onregelmatigheid die bestraft kan worden als je niet transparant bent. Ga hierover in gesprek met de docent wanneer duidelijke richtlijnen ontbreken.

Bibliografie

KU Leuven. (2023a, 13 juni). *Verantwoord gebruik van Generatieve Artificiële Intelligentie (Academiejaar 2022-2023)*. Student aan KU Leuven. https://www.kuleuven.be/onderwijs/student/onderwijstools/artificiele-intelligentie

KU Leuven. (2023b, 3 juli). *Verantwoord gebruik van Generatieve Artificiële Intelligentie (GenAI) in onderzoek*. Research. https://research.kuleuven.be/en/integrity-ethics/integrity/practices/genAI_NL