



aanpak 
**begeleidings
ethiek**

Verslag workshop
Aanpak Begeleidingsethiek
**Generatieve AI in het
onderwijs**

Hogeschool Windesheim, Journalistiek

ECP

Platform voor de
InformatieSamenleving

Voorwoord

De toenemende digitalisering brengt ingrijpende veranderingen met zich mee, in allerlei sectoren. Nieuwe mogelijkheden voor communicatie, monitoring en analyse roepen vragen op. Wat is de positie van de mens in deze omgeving, hoe staat het met data, met privacy, wat doen algoritmes en wat willen gebruikers eigenlijk? Begeleidingsethiek is ethiek die zich specifiek met dit soort vragen bezighoudt. Hij is gebaseerd op techniekfilosofie, die uitgaat van eeuwenoude verwevenheid tussen mens en technologie.

De Aanpak Begeleidingsethiek is een concrete aanpak waarin betrokkenen met elkaar in dialoog gaan over de effecten van de nieuwe technologie én de waarden die daarbij in het geding komen. Dat zijn vaak waarden gelieerd aan autonomie van de gebruiker, efficiëntie van het proces, transparantie van het algoritme, privacy, et cetera.

Tijdens de sessie komen verschillende stakeholders als gebruikers, ontwikkelaars, beleidsvormers en beslissers met elkaar in gesprek. Na de dialoog hebben de deelnemers ethische handelingsopties gegenereerd, waarvan verschillende vaak direct opgepakt kunnen worden. Er is gezamenlijk gekeken welke waarden we in het digitale domein belangrijk vinden en hoe we die willen verankeren en borgen in digitale processen en handelingen. Dit alles om ook bij verdere digitalisering op het vertrouwen van de samenleving en de participanten kunnen blijven rekenen. Dit is niet in één stap te realiseren het is een continu proces waarin deze workshop een schakel is.



Workshop aanpak begeleidingsethiek

Initiatiefnemers: Erik van Schaik en Wieger van Dalen Hogeschool Windesheim
Moderatoren: Daniël Tijink (ECP) en Wieger van Dalen (Windesheim)

Op 9 november organiseerde de opleiding Journalistiek (JOU) van Windesheim op initiatief van Erik van Schaik en Wieger van Dalen een sessie begeleidingsethiek over de casus Generatieve AI in het onderwijs. Aan de workshop namen 19 deelnemers (zie bijlage voor specificatie) deel onder leiding van Daniël Tijink (ECP) en Wieger van Dalen (Windesheim).

De doelstelling is tweeledig:

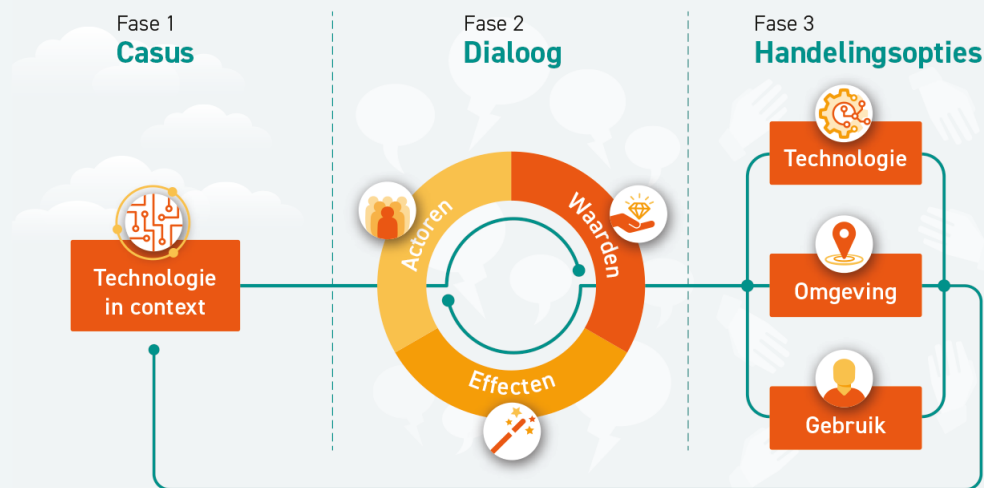
- ▷ Leren tijdens de workshop
- Welke effecten, waarden en actoren worden genoemd
- Welke handelingsopties zijn er om de toepassing te verbeteren
- ▷ Leren over de Aanpak Begeleidingsethiek
- Hoe kan de organisatie deze methodiek gebruiken rondom ethische kwesties en technologie-implementaties

Het verslag bevat de volgende elementen:

- ▷ Toelichting op de Aanpak Begeleidingsethiek
- ▷ Weerslag van de workshop



Aanpak begeleidingsethiek



De aanpak bevat de volgende fasen:

Fase 0 Introductie

Introductie over de doelstelling en een toelichting op het model en het gedachtengoed

Fase 1 Toelichting

Hoe ziet de technologie eruit en in welke context wordt deze toegepast

Fase 2 Dialoog

- ▷ Een korte ronde waarin de deelnemers aan de workshop de betrokken actoren benoemen
- ▷ Brainwrite waar deelnemers mogelijke effecten benoemen en bespreken
- ▷ Benoemen van waarden die een rol spelen bij die effecten

Fase 3 Handelingsopties

In subgroepjes gaan de deelnemers op zoek naar handelingsopties vanuit de technologie, de omgeving en het individu.



Fase 1

Technologie en context

Chat-GPT is een taalmodel ontwikkeld door OpenAI, gebaseerd op de GPT (Generative Pre-trained Transformer) architectuur. Het is ontworpen om mensachtige tekst te begrijpen en te genereren op basis van de ontvangen input. Het model is vooraf getraind op een diverse reeks internetteksten en is in staat om context te begrijpen, samenhangende en contextueel relevante antwoorden te genereren, en vragen in verschillende talen af te ronden. Het heeft geen toegang tot real-time informatie en zijn kennis is beperkt tot wat beschikbaar was tot aan de laatste trainingsafsplitsing in januari 2022.

Gebruikers kunnen met Chat-GPT communiceren door prompts of vragen te verstrekken, en het model genereert reacties op basis van zijn trainingsgegevens. Het wordt gebruikt in verschillende toepassingen, waaronder chatbots, contentgeneratie, taalvertaling en meer.

Hoewel Chat-GPT indrukwekkende mensachtige tekst kan genereren, levert het soms onjuiste of onsamenhangende antwoorden, en het is gevoelig voor de formulering van de input. Gebruikers moeten voorzichtig zijn en kritisch denken bij het interpreteren van de antwoorden.

Tijdens de toelichting vertellen drie studenten hoe zij Chat-GPT inzetten bij hun onderwijs. Een ICT- student vertelde wat meer over de technische achtergrond van Chat-GPT. De twee journalistiek (JOU) studenten vertelden dat ze Chat-GPT benutten bij verschillende onderdelen van het schrijfproces. Als er een werkstuk geschreven moet worden dan vragen ze chat-GPT om inspiratie, voorbeelden of een opzet. Bijvoorbeeld: *'Wat zou jij schrijven over vrijheid van meningsuiting en persvrijheid?'*, daarna werken ze het antwoord van Chat-GPT verder uit. Een samenvatting van een lange tekst is zo gemaakt door Chat-GPT, desgewenst op een specifieke manier, bijvoorbeeld B1 taalniveau, zodat het voor een ieder in de krant goed leesbaar is. Ook is Chat-GPT goed in het maken van samenvattingen van grote lappen tekst. Eén van de studenten vond het boek voor het vak financieel management te duur en besloot de inhoudsopgave te kopiëren en Chat-GPT te vragen hem te helpen bij het leren over die onderwerpen op hbo-niveau. In een 'leergesprek van zo'n 6 uur werkten ze alle onderwerpen door, met veel 'toetsmomenten' in de vorm van leervragen aan de student. Toen de student zei dat hij even pauze nam, reageerde Chat-GPT ondersteunend. Bij terugkomst gaf hij een korte samenvatting. De student sloot het tentamen af met een 7.8. De docent zei dat er in zijn carrière nooit een student was geweest die dit tentamen zonder boek had gehaald. Chat-GPT leert je misschien minder, maar ook beter schrijven. Hij laat zien hoe teksten van jezelf verbeterd kunnen worden. De studenten gaven aan dat ze Chat-GPT zo goed als bij elk vak gebruiken. Ook de gepresenteerde sheets waren door Chat-GPT gecheckt.



Fase 2

Dialogo

In deze tweede fase gaan de deelnemers in gesprek over wie er betrokken zijn bij de inzet en het gebruik van generatieve AI in het onderwijs. Ook buigen zij zich over de positieve en negatieve effecten van deze inzet, en benoemen zij belangrijke waarden waar rekening mee gehouden moet worden bij de inzet van een dergelijke toepassing.

Actoren:

Bij de actoren is de vraag wie er betrokken is of geraakt wordt door de casus. De deelnemers aan tafel vertegenwoordigen al een deel van die actoren. Ze noemen de volgende betrokkenen:

- Journalisten
- Studenten
- Docenten
- Examencommissies
- Curriculumcommissie
- Adviescommissie opleiding JOU (= werkveldadviescommissie, dus mensen uit de praktijk waarmee de opleiding 2x per jaar om tafel zit)
- Het Ministerie van OCW (inhoudelijk, en toetstechnisch, dat zijn verschillende afdelingen op het ministerie)
- Technologie providers van generatieve AI
- Artiesten, kunstenaars
- Uitgevers van onderwijsboeken
- ICT-afdeling Windesheim
- Raad v/d journalistiek
- Nederlands-Vlaamse Accreditatieorganisatie
- Publiek van de journalistiek (de lezer, de kijker, de luisteraar etc.)
- Het ministerie van Binnenlandse zaken (coördineert AI-beleid)

De vraag aan de deelnemers was om in het vervolg van de sessie ook te proberen vanuit het perspectief van deze actoren hun inbreng te geven.



Effecten:

De vraag is hier welke positieve en negatieve effecten de invoering en het gebruik van Generatieve AI (GAI)/Chat-GPT binnen een hogeschool kunnen hebben.

Positieve effecten

- voedt de nieuwsgierigheid
- ondersteunt het leerproces
- kan gebruikt worden als taalfouten checker
- hierdoor worden handgeschreven teksten bijzonder en extra waardevol
- minder controle taken voor docenten
- hogere tekstvaardigheid
- je kunt zo veel meer: meer talen, meer beelden, humor in je teksten
- toegang tot meer perspectieven
- het nodigt uit tot reflectie over wie we als maatschappij zijn, een spiegel
- prettige feedback op je schrijfvaardigheden
- een hulpmiddel voor mensen met dyslexie
- we moeten naar een ander, passender curriculum
- het levert tijdsbesparing zodat docenten meer persoonlijke aandacht voor studenten kunnen hebben
- beter nadenken over de inhoud van vakken
- Chat-GPT-GAI verandert zoveel. Het onderwijs en zeker de hogescholen zijn een prefecte proeftuin
- Andere kijk op kennis

Negatieve effecten

- Kan ongewenst beeldbevestigend werken, is immers gebaseerd op bestaande teksten
- Fraude: door Chat-GPT verslagen te laten maken
- Slechtere educatie
- Nog meer in handen van de machtige big tech
- Aantasting van het informatie-eco-systeem
- Taalverarming, kleine talen verliezen het)
- Culturele context gaat verloren, chat-GPT is 'internationaal', het werkt niet veel lokale kennis/teksten
- De grens tussen mens en GAI vervaagt, wat is nog authentiek?
- Fouten/onwaarheden zijn moeilijk te herkennen tussen de mooie tekst
- Goed voor de mainstream, niet
- Niet iedereen is 'digivaardig' het kunnen inschatten wat de waarde van de uitkomsten is, het goed kunnen werken/prompten van Chat-GPT
- Manipulatie GAI kan je ook in een bepaalde fuik trekken
- Het ondermijnt het leerproces



- AI is ook een reflectie van de maatschappij, hoe die is, niet hoe die zou kunnen zijn
- Digitaal kolonialisme, ofwel heel Engels, westers dominant
- We een heel ander curriculum ontwerpen
- Je moet steeds checken op authenticiteit: heeft de leerling wel echt zelf zijn werk gedaan
- Effecten van AI op privacy, wat je typt is voor hen
- De waarde van je diploma wordt onduidelijk
- Minder kritisch denkvermogen
- Sommige beroepsgroepen verdwijnen (zoals stemacteur, maar later ook journalist???)
- Hogere energiekostengrotere kloof, omdat Chat-GPT niet in gelijke mate voor iedereen toegankelijk is
- Er is nog geen heldere wetgeving of beleid



Waarden:

De relevante waarden zijn door de facilitators geïnventariseerd terwijl de deelnemers bezig waren met het benoemen van effecten. Achter veel effecten gaan waarden schuil. Die lijst is gedeeld met de deelnemers en door hen besproken en aangevuld. Dat leverde de volgende lijst met waarden op:

Waarden

Waarheid
Betrouwbaarheid/validiteit
Nieuwsgierigheid, leergierigheid
Menselijkheid
Leren/ontwikkelen
Wijsheid, rechtvaardigheid
Authenticiteit, originaliteit
Inclusie
Zin in werk/werkplezier
Zelfredzaamheid, zelfbeschikking
individualiteit
Creativiteit
Inspraak, uitdrukkingsvaardigheid
Vernieuwing/innovatie
Transparantie
Gelijkheid, mondiaal
Controleerbaarheid
Efficiëntie
privacy
Duurzaamheid



Fase 3

Handelingsopties

In het proces van het opstellen van de handelingsopties zijn de waarden uit de vorige fase meegenomen. De deelnemers kregen een uitleg over de verschillende categorieën binnen de handelingsopties. Het kan bijvoorbeeld gaan over het ontwerp van de technologie, het aanpassen van de omgeving waarbinnen de toepassing functioneert en het gedrag van mensen. De deelnemers worden in drie werkgroepen verdeeld en gaan ieder aan de slag met een van deze drie gespreksonderwerpen. Dat leidt tot handelingsopties per categorie.

Technologie:

1. Chat-GPT is nu een open platform, waarbij niet duidelijk is met welke data het getraind is, of waar de data blijft. Door naar een 'local hosted' variant te gaan, bijv. binnen Windesheim, kun je meer controle hebben op de data die gebruikt wordt, en mogelijk zelfs bepaalde waarden 'by design' afdwingen.
2. Een variant op 1) is om ook domein-specifieke implementaties te maken.
3. Introduceer een 'socratische modus' of 'doorvraag / coachende modus'. Daarmee leg je niet de druk bij de student neer om via prompt-engineering op een goede manier met Chat-GPT te werken, maar zal Chat-GPT niet slaafs antwoorden proberen te geven op je vragen, maar tot op zekere hoogte door te vragen.
4. Een fact-check optie, bijv. voor referenties: in dat geval wordt een link gelegd met een andere bron, mogelijk door een derde/onafhankelijke partij geverifieerd.
5. We zien dat een deel van de bias ontstaat doordat er veel op/met westerse databronnen is getraind. Gebruik een bredere set aan bronnen => 'inclusie by design'
6. Zorg voor vraagbundeling zoals bij SURF zodat er sprake kan zijn van 'armpje drukken' met de Big Tech, en onze voorwaarden worden overgenomen.
7. Ontwikkel een 'inclusiviteit-readiness' schaal / beoordeling / test.



Omgeving:

8. Heldere afspraken over het gebruik van GAI op Windesheim
9. Verklaring hoe student GAI heeft gebruikt (APA-plus)
10. Gesprek met docenten en studenten per vak over GAI: wat betekent het voor de werkbaarheid, de leercontext?
11. De authenticiteitsverklaring moet weg
12. Toestaan van GAI in het leerproces, aanpassen toetsing en verantwoordelijkheid
13. Kritische houding tav GAI aanleren
14. Een eigen toolmodel voor Windesheim (bv ism SURF).
15. Eventueel allen dit model gebruiken
16. Of een Nederlands LLM-model op open Source
17. Verklaring opstellen over een check op fraude
18. Heldere communicatie over wat we wel-niet doen
19. GA als vak bij elke opleiding
20. Docenten moeten elke toets zo maken, dat Chat-GPT hem niet kan maken

Mens:

21. Gebruik en toepassing van GAI moet een onderdeel zijn van de basis didactische bekwaamheid (bdb); verplichte scholing voor alle docenten
22. Elke docent moet verantwoorden hoe GenAI functioneert in zijn/haar vak
23. Elk tentamen moet door Chat-GPT gemaakt worden, als test.
24. Docenten moeten meer formatief handelen
25. We voeren aanwezigheidsplicht in voor alle vakken
26. Elke student moet kunnen werken met de laatste versie van Chat-GPT , dus mogelijk door de overheid/onderwijsinstelling gefinancierd

	techniek	omgeving	mens
Waarden			
Validiteit, betrouwbaarheid	1, 2, 6	14, 16, 18	22
Ontwikkeling, leergierigheid	3,	10, 12, 13, 19, 20	21, 23, 24, 25
Authenticiteit	4,	8, 9, 11, 17,	
Inclusie	5, 7	15, 1	26



Terugblik en afronding

Aan het einde van de workshop wordt teruggekeken en besproken wat de Aanpak Begeleidingsethiek concreet heeft opgeleverd. De deelnemers geven aan dat het veel nieuwe inzichten oplevert en dat er in korte tijd veel wordt gerealiseerd. Daarnaast wordt de brede groep van deelnemers als meerwaarde ervaren. De combinatie van een duidelijke casus, enthousiaste deelnemers, en een aanpak die zorgt voor duidelijke uitkomsten in een korte tijd was dus een succes.



De deelnemers

1. Emanuel Al-Kahwachi, student EL
2. Armand Bloem, Microsoft
3. Erik Bolhuis, beleidsadviseur CvB Windesheim
4. Yannick Bruintjes, student ICT en specialist GenAI in onderwijs
5. Sakina Elkayouhi, docent JOU
6. Erik Fledderus, lector Digital Business & Society bij domein Techniek Windesheim
7. Roland Gerbers, student JOU
8. Igor ter Halle, docent Communicatiekunde
9. Jasper Kars, Projectleider min BZK
10. Binne Keulen, docent JOU
11. Mariëlle Kuijper, docent JOU
12. Teun Lucassen, docent ICT
13. Bob Reiné, Ministerie OC&W
14. Erik van Schaik, docent JOU
15. Anneke Smits, lector Onderwijsinnovatie en ICT, domein Bewegen en Educatie Windesheim
16. Rob van de Star, docent ICT
17. Mitchell de Vries, student JOU
18. Pascal de Vries, lid examencommissie BMR en onderzoeker constructieve journalism
19. Marloes Wevers, docent JOU







Platform voor de
InformatieSamenleving

Verslag workshop
Aanpak Begeleidingsethiek
Generatieve AI in het
onderwijs

Meer informatie over de Aanpak Begeleidingsethiek,
waaronder dit verslag, vindt u op

www.begeleidingsethiek.nl