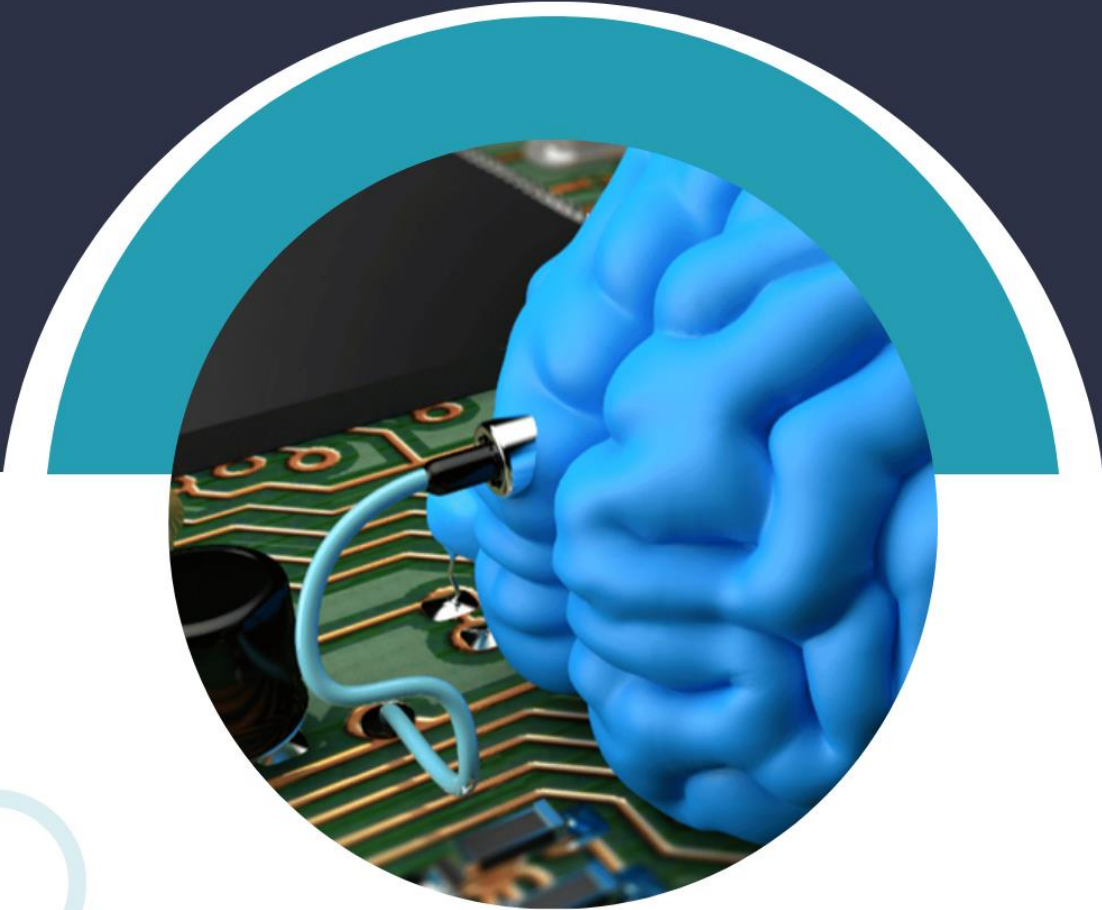




Hogeschool van
Amsterdam



WEEKOPDRACHT DENKEN OVER TECH

Naam:

Theist Tanwir

Studentnummer:

500884055

Minor:

Het internet is stuk, maar we gaan het repareren

Docent:

Hans de Zwart

Datum:

15-9-24

Inhoudsopgave

1. Inleiding	3
1.1 Ethiek en kunstmatige intelligentie in de Google zoekmachine	3
1.1.1 Stap 1: Recognize an ethical issue	3
1.1.2 Stap 2: Get the facts	3
1.1.3 Stap 3: Evaluate alternative actions	4
1.1.4 Stap 4: Make a decision and test it	4
1.1.5 Stap 5: Act and Reflect on the Outcome	4
2. Figure en Ground van ChatGPT	5
2.1 Figure van ChatGPT	5
2.2 Ground van ChatGPT	5
3. Onverwachte gevolgen van AI-antwoorden in Google en hoe dit te verbeteren	6
3.1 Impact van AI op Websiteverkeer en Informatievariëteit	6
3.1.1 Verlies aan verkeer voor websites	6
3.1.2 Beperking van diversiteit aan informatie	6
3.2 Tarot Cards of Tech	6
3.2.1 The Smash Hit	6
3.2.2 The Forgotten	7
3.3 Morele Verbeeldingskracht	7
4. Account(s) op Fediverse	7
Bronnen	8

1. Inleiding

Auteurs van de Inleiding:

Theist Tanwir en Karandeep Singh Gill

1.1 Ethiek en kunstmatige intelligentie in de Google zoekmachine

Wat is Google? Google is een van de grootste informatieplatforms ter wereld en maakt steeds meer gebruik van kunstmatige intelligentie (AI) om zoekresultaten te genereren. Deze AI-systemen geven vaak direct antwoord op vragen zonder menselijk toezicht om te bepalen of de antwoorden correct zijn. Dit kan riskant zijn omdat verkeerde of onvolledige informatie als feit kan worden gepresenteerd en zo verkeerde informatie kan verspreiden en vooroordelen kan versterken.

De vraag is dus of het moreel juist is dat Google AI- gegenereerde antwoorden geeft zonder deze grondig te controleren. In deze opdracht kijken we naar het ethische dilemma door de stappen van het Markkula te volgen. De nadruk zal liggen op de morele aspecten van dit probleem en mogelijke oplossingen.

1.1.1 Stap 1: Recognize an ethical issue

Het ethische probleem in dit geval is of het juist is dat Google antwoorden door AI laat genereren zonder deze te controleren. De kans bestaat dat gebruikers foutieve informatie krijgen, wat kan leiden tot verkeerde beslissingen. Dit kan vooral gevaarlijk zijn wanneer het gaat om gezondheidsinformatie of financiële adviezen. Het probleem raakt niet alleen individuele gebruikers, maar heeft ook bredere gevolgen voor de samenleving. De verspreiding van misinformatie kan negatieve effecten hebben op een groot aantal mensen, zoals het versterken van misverstanden of het vergroten van ongelijkheid in de toegang tot betrouwbare informatie. Dit dilemma gaat verder dan alleen juridische kwesties; het draait om de ethische verantwoordelijkheid die een groot technologiebedrijf als Google heeft.

1.1.2 Stap 2: Get the facts

Om dit dilemma goed te begrijpen, moeten we kijken naar hoe de AI-systemen van Google werken. Deze systemen zijn ontworpen om snel enorme hoeveelheden informatie te verwerken en daaruit antwoorden te genereren. Hoewel dit heel efficiënt is, worden deze antwoorden vaak niet door mensen gecontroleerd voordat ze aan de gebruiker worden getoond. Hierdoor kunnen onnauwkeurige of verouderde antwoorden worden gegeven. Google heeft er belang bij om zoekresultaten zo snel mogelijk te leveren, want dit verhoogt het aantal gebruikers en hun interactie met het platform. Maar gebruikers verwachten wel dat de informatie betrouwbaar en correct is.

Een belangrijke vraag hierbij is of Google een mechanisme heeft om AI-antwoorden te verifiëren en welke gevolgen het heeft als dat mechanisme ontbreekt. Een andere factor om te overwegen is de mate waarin gebruikers begrijpen dat de antwoorden van AI komen en niet altijd gecontroleerd zijn.

1.1.3 Stap 3: Evaluate alternative actions

Om het ethische dilemma op te lossen, kunnen we verschillende opties overwegen. Deze opties kunnen worden beoordeeld aan de hand van verschillende ethische benaderingen:

- **Utilitaristische benadering:** Welke actie levert het meeste goede op voor het grootste aantal mensen? Het invoeren van een systeem om AI-antwoorden te verifiëren zou de algehele betrouwbaarheid van zoekresultaten verbeteren, wat een positief effect heeft op veel mensen.
- **Rechten benadering:** Gebruikers hebben het recht op nauwkeurige en betrouwbare informatie. Een verificatiesysteem zou dit recht respecteren door ervoor te zorgen dat de antwoorden van Google zo correct mogelijk zijn.
- **Rechtvaardigheid benadering:** Deze benadering vraagt of alle gebruikers op een eerlijke manier beschermd worden tegen foutieve informatie. Zonder verificatie lopen sommige gebruikers, bijvoorbeeld mensen die minder ervaring hebben met technologie, meer risico op misleiding.
- **Gemeenschappelijk goedbenadering:** Het welzijn van de maatschappij als geheel is afhankelijk van betrouwbare informatie, vooral bij belangrijke onderwerpen zoals gezondheidszorg. Het waarborgen van de juistheid van informatie is daarom goed voor de samenleving.
- **Deugd ethische benadering:** Google moet zich richten op deugden zoals betrouwbaarheid, transparantie en verantwoordelijkheid. Door deze deugden centraal te stellen, kan Google het vertrouwen van gebruikers versterken en zijn morele verplichtingen nakomen.

1.1.4 Stap 4: Make a decision and test it

De beste optie in dit geval is om een controle- en verificatiesysteem in te voeren voor door AI gegenereerde antwoorden. Dit systeem zou ervoor zorgen dat de antwoorden die Google geeft, gebaseerd zijn op gecontroleerde en betrouwbare bronnen. Deze beslissing zou waarschijnlijk als positief worden gezien door het publiek, omdat het gebruikers in staat stelt om met meer vertrouwen de zoekresultaten te gebruiken. Het zou ook helpen om de morele verantwoordelijkheid van Google te versterken. Als deze beslissing publiekelijk bekendgemaakt wordt, zouden mensen het waarschijnlijk als een positieve stap richting verantwoordelijkheid en ethisch gedrag beschouwen.

1.1.5 Stap 5: Act and Reflect on the Outcome

Bij de uitvoering van deze beslissing moet Google zorgvuldig plannen hoe het verificatiesysteem geïmplementeerd kan worden. Dit betekent investeren in technologie en personeel om de AI-antwoorden te controleren en waar nodig aan te passen. Google moet ook transparant zijn naar gebruikers toe over de stappen die worden genomen om de betrouwbaarheid van de informatie te waarborgen, zodat gebruikers begrijpen welke antwoorden gecontroleerd zijn en welke mogelijk niet.

Na implementatie is het belangrijk om de resultaten van het verificatiesysteem te monitoren. Wordt de betrouwbaarheid van de antwoorden verbeterd? Hoe reageren gebruikers op deze verandering? Google zou ook moeten leren van eventuele fouten die zich voordoen, zodat ze hun systeem voortdurend kunnen verbeteren en kunnen zorgen dat ze aan hun ethische verplichtingen blijven voldoen. {Bron: <https://www.scu.edu/ethics/ethics-resources/a-framework-for-ethical-decision-making/>}

2. Figure en Ground van ChatGPT

ChatGPT is erg populair als het gaat om het stellen van vragen en het krijgen van antwoorden. Bij het gebruik van ChatGPT worden alleen de zichtbare kant van deze technologie ervaren, maar er gebeurt veel op de achtergrond om dit mogelijk te maken.

Zowel “figure” als “ground” spelen hierbij een rol:

2.1 Figure van ChatGPT

De "figure" is de kernfunctionaliteit van ChatGPT, namelijk het genereren van mensachtige tekst gebaseerd op invoer van gebruikers. Dit maakt het mogelijk om teksten te schrijven, vragen te beantwoorden en gesprekken te voeren. Dit is een technologie die veelzijdig en toegankelijk is, en kan worden ingezet voor allerlei praktische toepassingen zoals klantenservice, contentcreatie en educatie.

Met andere woorden, het is de kant van ChatGPT waar interactie mee mogelijk is. Dit is waar de meeste mensen aan denken als ze ChatGPT gebruiken: een slimme assistent die snel en efficiënt antwoordt. {Bron: <https://study.com/learn/lesson/figure-ground-perception-relationship-examples.html#:~:text=Figure%2Dground%20perception%20holds%20that,be%20perceived%20as%20two%20faces.>}

2.2 Ground van ChatGPT

De "ground" gaat verder dan wat nodig is om ChatGPT te laten werken en omvat bredere maatschappelijke effecten. Bijvoorbeeld:

- **Werkgelegenheid:** Door de opkomst van AI zoals ChatGPT kunnen er banen verdwijnen in sectoren waar tekstproductie en klantenservice centraal staan. Dit leidt tot zorgen over automatisering en werkloosheid.
- **Privacy en dataverzameling:** Om goed te functioneren, heeft ChatGPT toegang tot grote hoeveelheden data. Dit roept vragen op over hoe deze data wordt verzameld, opgeslagen en gebruikt, vooral als deze data afkomstig is van gebruikersinteracties.
- **Misinformatie:** AI kan snel en effectief teksten genereren, wat kan leiden tot de verspreiding van misinformatie. Dit kan schadelijk zijn in contexten zoals politiek of gezondheid, waar nauwkeurige informatie cruciaal is.
- **Bias en discriminatie:** De training van ChatGPT is gebaseerd op bestaande teksten op internet, die bias kunnen bevatten. Dit betekent dat ChatGPT onbewust discriminerende of bevooroordeelde antwoorden kan geven, wat kan leiden tot ethische problemen bij gebruik.

{Bron: <https://openai.com/chatgpt/>} {Bron: <https://openai.com/index/gpt-4-research/>}

3. Onverwachte gevolgen van AI-antwoorden in Google en hoe dit te verbeteren

Wanneer Google met AI een direct antwoord geeft op vragen in plaats van naar andere websites te linken, ontstaan er enkele onverwachte gevolgen en externaliteiten.

3.1 Impact van AI op Websiteverkeer en Informatievariëteit

3.1.1 Verlies aan verkeer voor websites

Websites die veel bezoekers kregen via Google-links kunnen nu minder verkeer hebben. Dit is vooral problematisch voor sites die geld verdienen met advertenties, omdat minder bezoekers betekent dat ze minder advertenties kunnen verkopen en dus minder verdienen. Hierdoor kunnen onafhankelijke informatiebronnen verzwakken. Mensen klikken minder vaak door naar deze websites en krijgen in plaats daarvan direct antwoorden van AI-systemen zoals ChatGPT. Dit kan de invloed van deze websites verminderen en de diversiteit van informatiebronnen verkleinen.

3.1.2 Beperking van diversiteit aan informatie

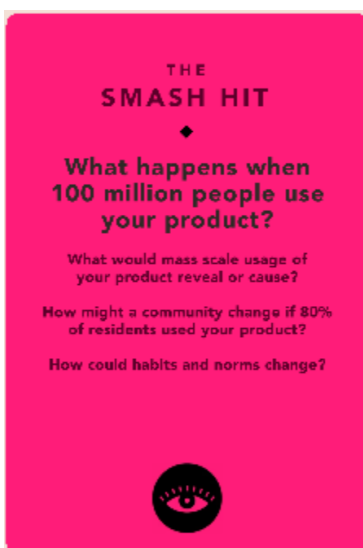
AI-algoritmes geven vaak één standaardantwoord zonder veel ruimte voor verschillende meningen of details. Hierdoor krijgen gebruikers misschien niet alle informatie en perspectieven die ze normaal via verschillende bronnen zouden krijgen. Dit kan de variëteit van informatie beperken en leiden tot een eenzijdig beeld, omdat belangrijke inzichten en details verloren gaan door het ontbreken van diverse bronnen.

3.2 Tarot Cards of Tech

3.2.1 The Smash Hit

Deze kaart verwijst naar de onverwachte problemen die ontstaan wanneer technologie niet werkt zoals bedoeld. Bij AI-systemen zoals die van Google kan dit leiden tot ongewenste effecten zoals een afname van bezoekers voor websites die normaal veel verkeer via Google kregen. Dit kan de diversiteit van informatie verminderen omdat gebruikers minder vaak naar verschillende bronnen doorklikken.

Oplossing: Google zou kunnen zorgen voor een betere balans door AI-antwoorden te combineren met links naar de echte websites. Dit helpt niet alleen om de diversiteit van informatie te behouden, maar zorgt er ook voor dat gebruikers toegang hebben tot meerdere perspectieven en bronnen.



Figuur 1: The Smash Hit

3.2.2 The Forgotten

Deze kaart gaat over de mensen of groepen die door technologie worden gemarginaliseerd of over het hoofd worden gezien. Kleine, onafhankelijke websites en nichesites kunnen vergeten worden als AI alleen populaire en vaak bezochte bronnen gebruikt om antwoorden te genereren. Dit kan leiden tot een versterking van al bestaande machtsstructuren, waarbij grote websites en bedrijven domineren en kleinere spelers irrelevant worden.



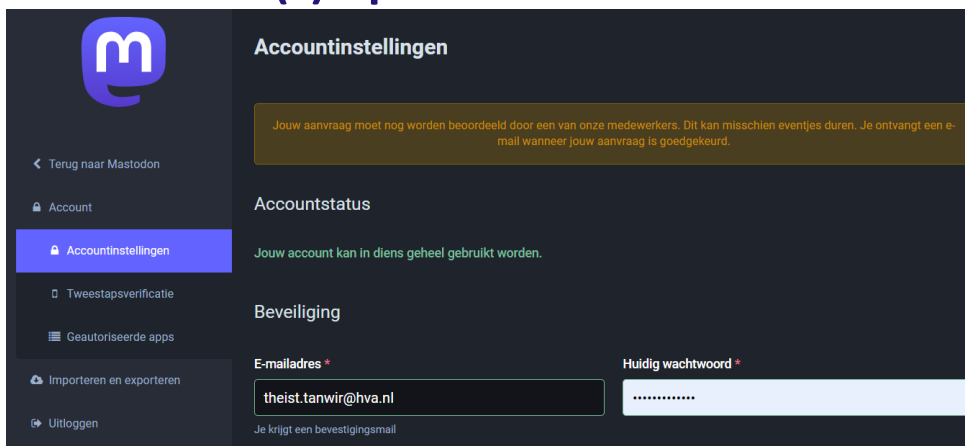
Figuur 2: The Forgotten

Oplossing: Google zou kunnen overwegen om bij elke AI-generatie ook een lijst van alternatieve bronnen en perspectieven te tonen. Dit kan helpen om kleinere websites zichtbaarder te maken en een diverser informatieaanbod te waarborgen. {Bron: <https://tarotcardsoftech.artefactgroup.com/>}

3.3 Morele Verbeeldingskracht

Met morele verbeeldingskracht kunnen we ons voorstellen wat de langetermijneffecten van AI in Google Search zijn op de informatie-economie. Eén oplossing zou kunnen zijn om gebruikers actiever te betrekken bij de keuze voor welk type antwoord ze willen ontvangen: een snel AI-antwoord of een diepgaande lijst van links naar betrouwbare bronnen. Hierdoor wordt de gebruiker empowered en blijft de diversiteit van het internet intact. {Bron: <https://talks.hansdezwart.nl/hva-2023-7/>}

4. Account(s) op Fediverse



Figuur 3: Fediverse

Mijn account moet nog worden beoordeeld.

Bronnen

1. Santa Clara University. (z.d.). *A Framework for Ethical Decision Making*. Markkula Center For Applied Ethics. <https://www.scu.edu/ethics/ethics-resources/a-framework-for-ethical-decision-making/>
2. *Figure-Ground Perception | Definition, Principles & Types - Lesson | Study.com*. (n.d.). study.com. <https://study.com/learn/lesson/figure-ground-perception-relationship-examples.html>
3. *ChatGPT*. (n.d.). OpenAI. <https://openai.com/chatgpt/>
4. *GPT-4*. (n.d.). OpenAI. <https://openai.com/index/gpt-4-research/>
5. *The tarot cards of tech*. (n.d.). <https://tarotcardsoftech.artefactgroup.com/>
6. *Morele verbeeldingskracht*. (n.d.). <https://talks.hansdezwart.nl/hva-2023-7/>