

Etische raamwerk

Ethisch dilemma:

Mag een onofficiële nieuwsplatform bewust nepnieuws aanbieden aan de gebruiker op social media platform zoals YouTube om meer likes te krijgen?

Opties:

- Nee: Een onofficieel nieuwsplatform mag niet zomaar bewust nepnieuws verspreiden, ook niet op socialmediaplatforms zoals YouTube, om meer likes of views te krijgen. Dit valt onder desinformatie.
- Ja (beperkt): Zolang het niet onder specifieke verboden valt, zoals oproepen tot geweld of fraude.
- Ja (absoluut): Als het nepnieuws duidelijk fictief of verzonnen is, zonder de intentie om mensen te misleiden.. Dit hangt af van de context, de intentie en de regels van het platform.

Raamwerken:

The Utilitarian Approach (nee)

-Nepnieuws kan veel mensen misleiden of schade veroorzaken. Het grootste welzijn van de meeste mensen wordt geschaad, dus het mag niet.

The Rights Approach (ja ,absoluut)

-Als nepnieuws duidelijk fictief of verzonnen is en mensen niet misleidt, worden rechten (zoals recht op waarheid en veiligheid) niet geschonden.

The Justice Approach (ja ,beperkt)

-Zolang het nepnieuws niet onder specifieke verboden valt (zoals oproepen tot geweld of fraude), wordt er niemand onrechtvaardig behandeld, omdat het niemand echt schaadt."

Etische raamwerk

The Common Good Approach (nee)

-Nepnieuws tast het vertrouwen van de gemeenschap in informatie en media aan. Het schaadt het algemeen belang → dus het mag niet.

The Virtue Approach (ja ,absoluut)

Als eerlijk persoon zou je nepnieuws dat mensen misleidt vermijden. Fictieve content die duidelijk verzonnen is en bedoeld is om te vermaken, past daarentegen wel bij een eerlijk persoon.

De optie Ja (beperkt) houdt het meeste rekening met de voordelen en minimaliseert de schade.

Ordeel:

Ja (beperkt) → Het gaat om verzonnen verhalen die niet verboden zijn en in de meeste gevallen geen schade veroorzaken. Zolang niemand onrechtvaardig wordt behandeld en er geen illegale activiteiten plaatsvinden, kan beperkt nepnieuws worden toegestaan.

Morele verbeeldingskracht

Ai in google search:

Veiligheid / Vertrouwen

- AI kan automatisch verkeerde of misleidende informatie hoger in de resultaten plaatsen.
- Mensen kunnen beslissingen laten beïnvloeden door wat AI als eerste toont → ze controleren zelf minder kritisch.
- **Filterbubbels ontstaan door gepersonaliseerde resultaten: je ziet steeds dezelfde informatie, waardoor je beperkt wordt in het ontdekken van andere perspectieven.**

Privacy / Transparantie

- **Gebruikers weten niet welke bronnen prioriteit krijgen en waarom een bepaald resultaat bovenaan staat.**
- **Betaalde advertenties of promoties kunnen boven neutrale informatie komen, zonder dat de gebruiker dat duidelijk ziet.**

Macht

- **Commerciële belangen boven publieke waarden: winst kan belangrijker zijn dan objectieve informatie.**

Samenleving

- **Mensen hebben minder kans om alternatieve perspectieven te zien, wat sociale en politieke polarisatie kan versterken.**

Ik heb de link gelegd met de ethische raamwerken. Hiervanuit ben ik systematisch gaan kijken vanuit verschillende benaderingen, en door elk perspectief apart te bekijken ontdekte ik risico's die ik eerst over het hoofd zag. Ik heb de risico's vervolgens geclusterd per thema, zoals in het voorbeeld dat in de link stond.

Algoritmes

Vraag 1: Bedenk een algoritme dat mij vertelt wie van jullie de prijs moet krijgen voor beste student .

- aanwezigheid
- inzet
- werkhouding
- cijfers

Nul meting en eindmeting (verschil = groei)
waardes ook mee nemen, als je groeit en hoog.

Vraag 2: Bedenk een algoritme dat voor mij voorspelt welk eindcijfer elke student gaat behalen voor deze minor.

- Geschiedenis van cijfers
- Wanneer/hoe snel iemand zich heeft ingeschreven
- Geschiedenis aanwezigheid
- Oude docenten geven cijfer aan student
- Antwoorden op intake formulier
- Studie achtergrond/opleiding

Op basis van deze waardes min en plus punten toe kennen.
Iedereen start met een 6 sommige lage scores geven min punten (en anders om).

Fediverse post

Mijn eerste post op het fediverse bestaat uit willekeurige feitjes die elke 24 uur automatisch worden geplaatst.

