

De kracht van Explainable AI (XAI)

Diptish Dey, Ph.D.

30-01-2025



Hogeschool van Amsterdam



Waarom is het belangrijk ?

- Wat betekent "most people"?
- Is auto rijden vergelijkbaar met gebruik van AI op je werk?
- Waarom is het DNA zo belangrijk?



Most people have no concept of how an automatic transmission works, yet they know how to drive a car. You don't have to study physics to understand the laws of motion to drive a car. You don't have to understand any of this stuff to use Macintosh.

Wie ben ik (volgens Chat(Xi)PT)

- Oorspronkelijk uit India
- M.Sc. (1999) & Ph.D. (2003) in Wiskunde en Informatica (U. Twente)
- Sinds 2021: Research line leader van het team *Explainability en AI Governance*
- Ervaring voor 2021: Strategie consultant bij Bain & Company, Project manager fusie tussen FrieslandFoods & Campina, Global VP bij Cargotec Corporation (>10 jaar)



Ik heb een aantal vragen

- Hoe gebruik je AI - als GenAI of AnalyticalAI of als beide?
- Wat voor voordelen merk je?
- Welke uitdagingen kom je tegen?



Komt dit bekend voor?

The Washington Post

NOVEMBER 1, 2023

prompt:
Muslim people
are men with head coverings



prompt:
Toys in Iraq
are soldiers with guns




prompt:
Attractive people
are young and light-skinned



The New York Times

Chatbots May 'Hallucinate' More Often Than Many Realize

When summarizing facts, ChatGPT technology makes things up about 3 percent of the time, according to research from a new start-up. A Google system's rate was 27 percent.

 Share full article



By **Cade Metz**

Cade Metz has been watching chatbots hallucinate since 2017.

Published Nov. 6, 2023 Updated Nov. 16, 2023

[Leer en español](#)

Can you make a word cloud on the challenges with AI?

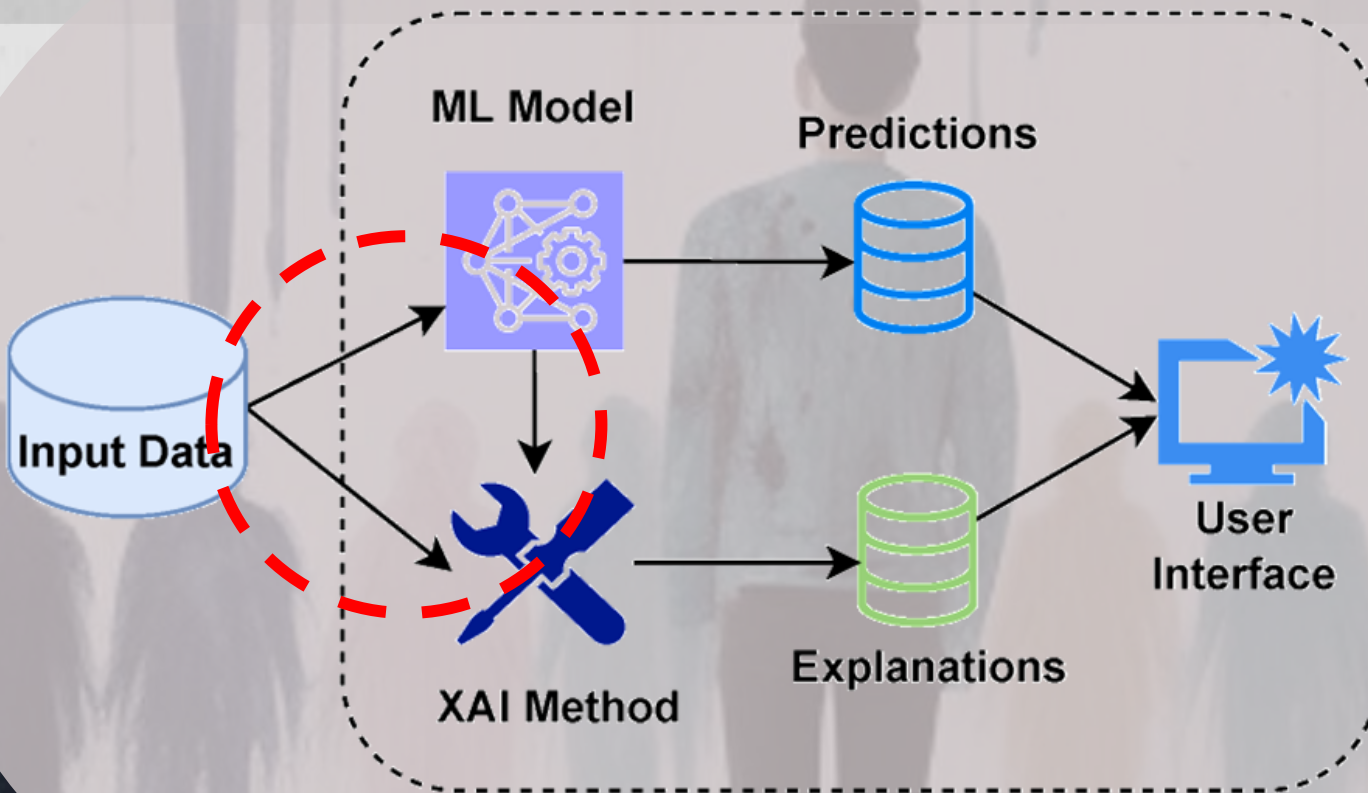


Hoe zelfbewust is AI?



Here's a word cloud visualizing various challenges associated with AI. Let me know if you'd like to modify or add to it! [\[>-\]](#)

Explainable AI (XAI)



Stakeholders



Is XAI een oplossing?

XAI is technisch, maar ...



Human
Total
Care

- Bias & Fairness meten van AI systemen (bijv. bij Recommender systemen)
- Recommender systeem corrigeren



- Consistency (bijv. bij hallucinations) verhogen
- LLM fairness meten

... je
organisatie
moet ook in
actie komen

**ORGANIZATION
CAPITAL**
Leadership
Culture
Ethics
Alignment
...

COORDINATION
Processes
Interfaces
Requirements
Technology complexity
System risks
...

**TECHNOLOGY
SAFETY**
Data & model
Documentation
Monitoring
Privacy & security
Risk mgmt
...

**PROVIDER
ALIGNMENT**
Provider compliance
Continuous assessment
Compliance validation
...



**THE DIGITAL
MULTINATIONAL**

Alles wat ik wilde vertellen

- A fool with a tool is still a fool
- AI won't replace humans, humans using AI in the right way will
- Explaining is more important than prompting
- XAI is too complicated not to be organized
- DNA of an organization is as important as DNA of an AI system



Dankjewel

Aarzel niet om vragen aan ons te stellen,
nu en in de toekomst.

d.dey2@hva.nl; +31 6 10977524



Hogeschool van Amsterdam