

Machine Learning Ethics



W



Terugblik exit tickets



Hoe ging het maken van opdrachten D en E?

- Geef een korte toelichting op de opdracht, waarin alle relevante feiten benoemd worden.
- Benoem een morele vraag passend bij jullie opdracht.
- Benoem de morele uitgangspunten (normen, waarden en deugden) die je relevant vindt. Plaats deze tegenover elkaar en werk voor beide kanten morele uitgangspunt(en) uit. Leg uit hoe en waarom deze uitgangspunten een rol spelen.
- Ga in op de kwestie van verantwoordelijkheid. Wie is in deze kwestie waarvoor verantwoordelijk en waarom?
- Benoem diverse argumenten vóór en tegen het moreel dilemma in kwestie vanuit
 - De plichtethische invalshoek
 - De gevolgenethische invalshoek
 - De deugdethische invalshoek
- Benoem je moreel oordeel. Dit oordeel is een afgewogen antwoord op je morele vraag.
- Reflecteer kritisch op het oordeel.
- Handelingsopties (indien van toepassing): voeg concrete handelingsopties toe m.b.t. de technologie, omgeving en gebruiker.
- Terugblik stappenplan: eigen morele grens

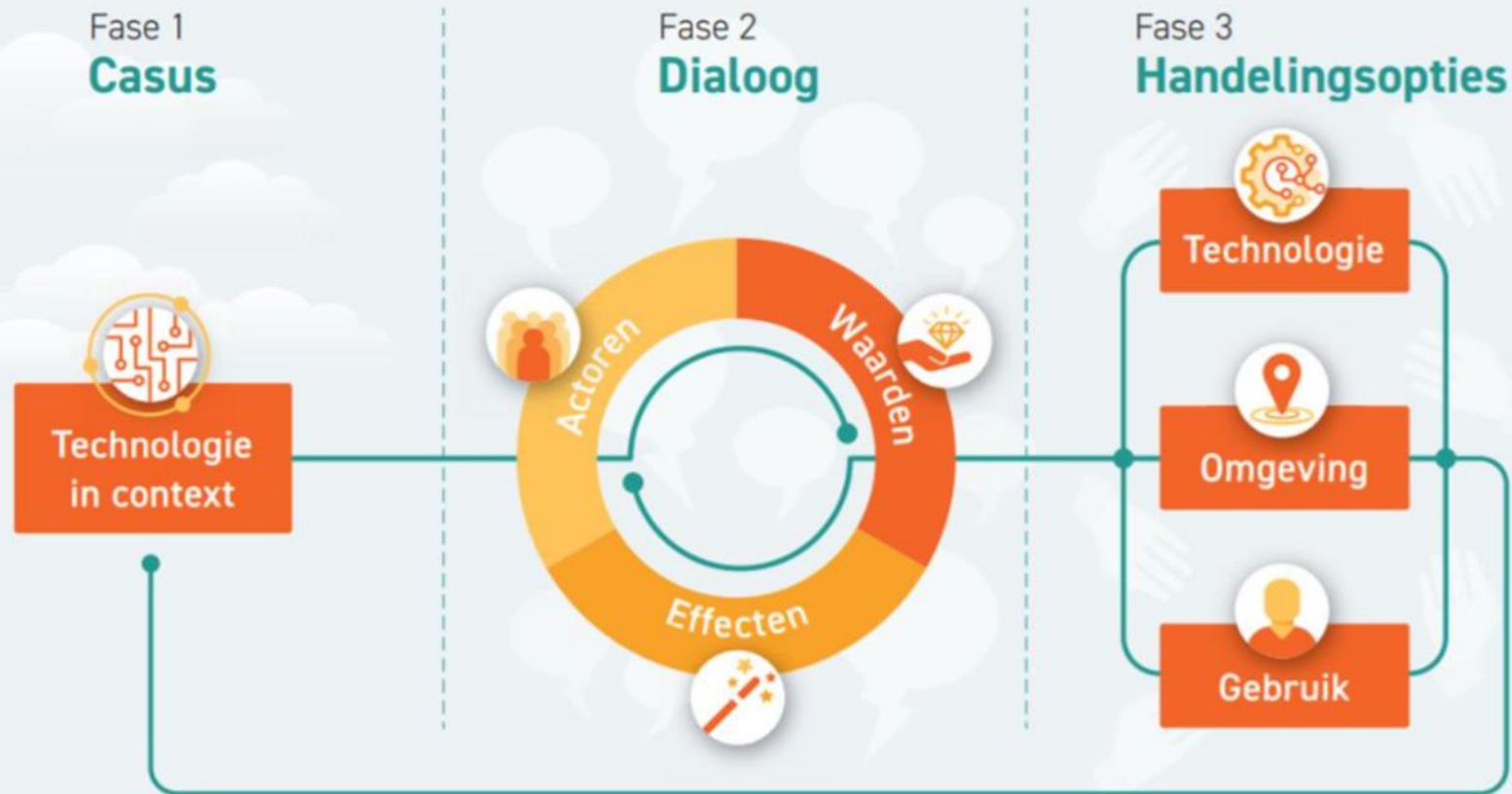


Workshop 4

- Handelingsopties
- Begeleidingsethiek
- Portfolio en project



Vervolgstep: handelingsopties



Een andere benadering

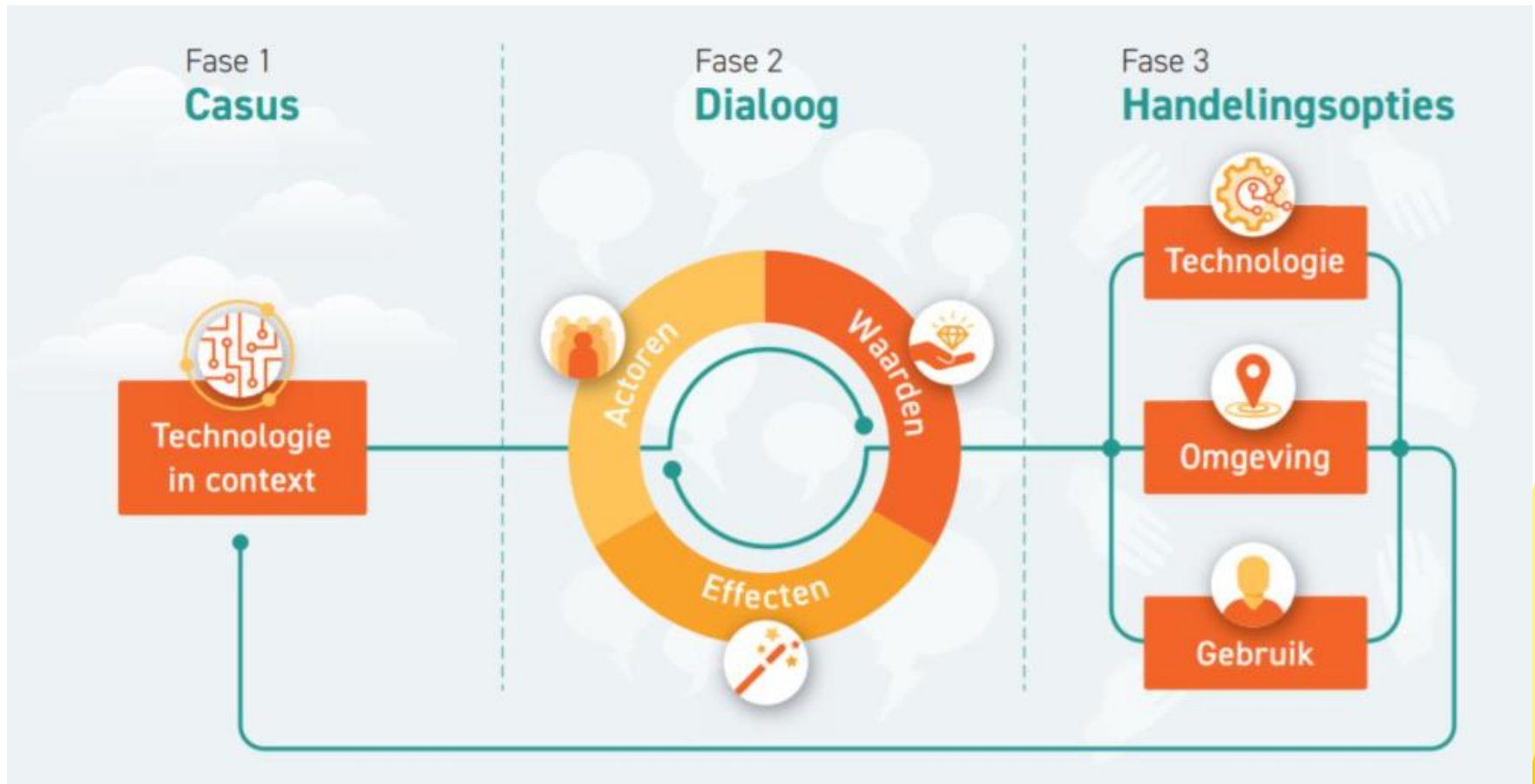
- De begeleidingsethiek

“De centrale vraag in begeleidingsethiek is niet ‘ja of nee?’ maar ‘hoe?’ Ze richt zich niet op afwijzen of aanvaarden, maar op het waardenvolle wijze ontwerpen, implementeren en gebruiken van nieuwe technologie.”

Bron: Peter-Paul Verbeek en Daniël Tijink, *Aanpak begeleidingsethiek*, ECP 2019, 11.



Aanpak begeleidingsethiek



Fase 1: Casus

- Technologie in context
 - Concrete technologie
 - Concrete context
- Nauwkeurige en begrijpelijke beschrijving



Fase 2: Dialoog

- Actoren – betrokkenen
- Effecten – mogelijke gevolgen
- Waarden – die een rol spelen

Dialoog actoren over mogelijke effecten en belangrijke waarden



Fase 3: Handelingsopties

- Ontwerpen van technologie (aansluiten bij waarden)
 - *Ethics by design*
- Omgeving (fysiek inrichten en sociaal, afspraken)
 - *Ethics in context*
- Gebruiker (waardenvol gedrag)
 - *Ethics in use*



Na het moreel oordeel

- Morele vraag is vaak geformuleerd als 'ja / nee'
 - Mag je ... ?
- 'Nee' is helder, dan houdt het (op dit moment) op
- 'Ja' kan echter nog voorwaardelijk zijn
 - Ja, als we eerst ...
 - Ja, maar alleen als ...

(NB het antwoord 'nee, tenzij' is een vorm van 'ja')



Bijvoorbeeld: Mag je AI gebruiken in de rechtspraak?

- Nee, want AI is nog niet geavanceerd genoeg.
- Ja, maar alleen ter ondersteuning van een menselijke rechter
- Ja, maar alleen als de gebruikte data betrouwbaar en transparant is
- Ja, maar alleen als de verdachte bezwaar kan maken
- Etc.



- Machine Learning ontwikkelt zich steeds verder
- Bij iedere fase past steeds opnieuw de 'hoe'-vraag (en 'niet' is ook een antwoord op de vraag)
- Continu nadenken over de voorwaarden die je stelt aan elke nieuwe wijziging / toepassing (waarbij het afwijzen ervan ook een optie blijft), bv. bij risico op function creep



**Wat kun je allemaal met
sentimentanalyses doen?**

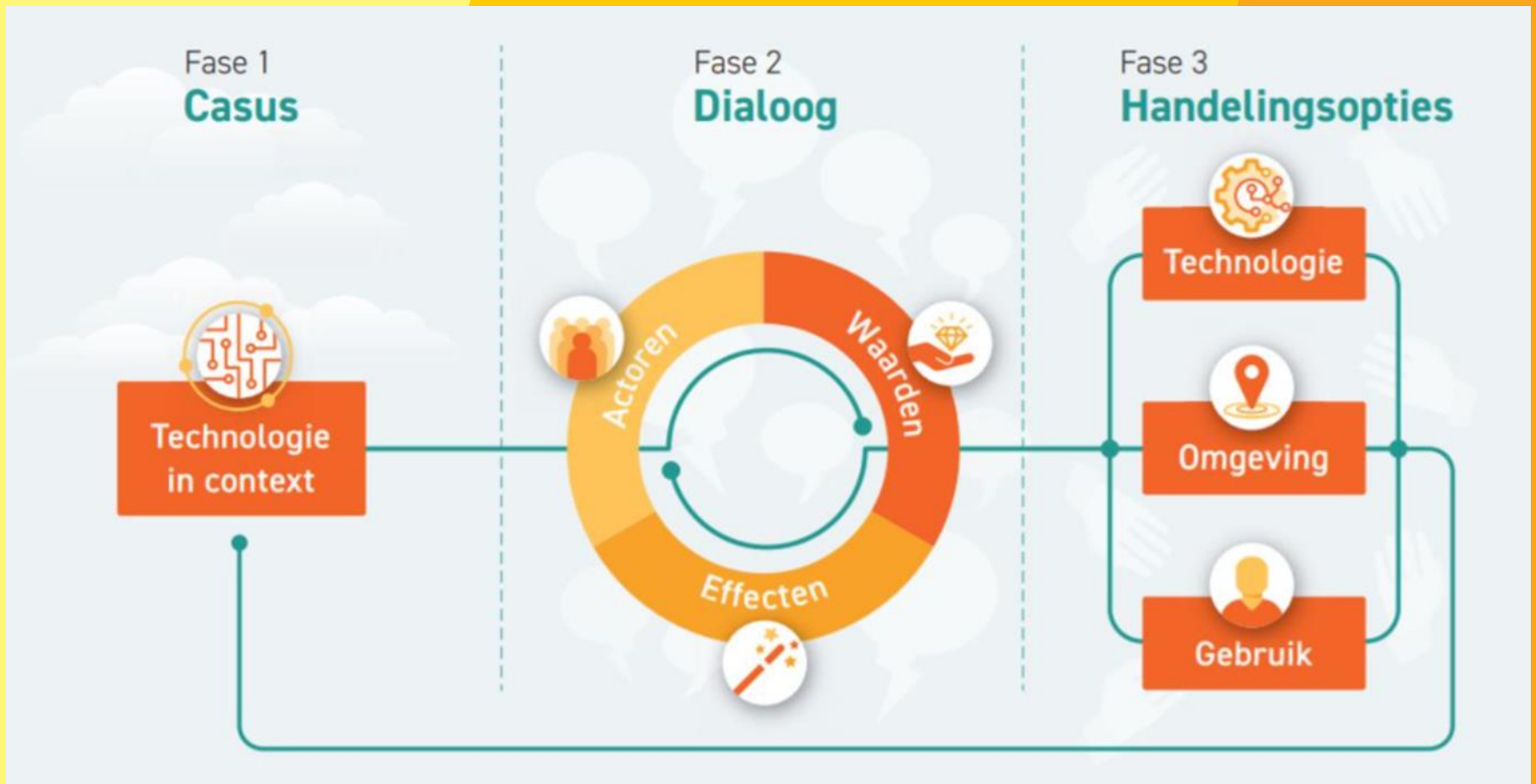
Formuleer een morele vraag

Uitwerking

- Benoem de morele uitgangspunten (normen, waarden en deugden) die jullie relevant vinden. Plaats deze tegenover elkaar en werk voor beide kanten morele uitgangspunt(en) uit. Leg uit hoe en waarom deze uitgangspunten een rol spelen.
- Ga in op de kwestie van verantwoordelijkheid. Wie is in deze kwestie waarvoor verantwoordelijk en waarom?
- Benoem diverse argumenten vóór en tegen het moreel dilemma in kwestie vanuit
 - De plichtethische invalshoek
 - De gevolgenethische invalshoek
 - De deugdethische invalshoek
- Benoem jullie moreel oordeel. Dit oordeel is een afgewogen antwoord op jullie morele vraag.
- Reflecteer kritisch op het oordeel.

	WEL	NIET
Normen, waarden, deugden		
Plichtethisch		
Gevolgenethisch		
Deugdethisch		

- Nuancering: voeg concrete handelingsopties toe m.b.t. de technologie, omgeving en gebruiker.





Technologie



Omgeving



Gebruik

Portfolio

Opdracht na workshop 4 (individueel):

Opdracht F (individueel, na workshop 4 in portfolio):

- Benoem morele vragen bij de derde themaopdracht (chatbot).
- Kies een relevante morele vraag en maak daarvoor een individuele uitwerking van de themaopdracht o.b.v. het stappenplan uit de bijlage. Werk hierbij ook de stap 'handelingsopties' uit.
- Licht je eigen morele grens bij dit thema toe.
- Gebruik bij deze opdracht de *checklist bij het stappenplan ethiek ML*.



Inleveren tussentijds portfolio:

Na de vier workshops wordt het individueel portfolio (waar ethiek een onderdeel van is) op de ELO ingeleverd voor feedback.

- Kies opdracht D/E of F als showcase om de leeruitkomst aan te tonen.
- Voeg in één bijlage alle uitwerkingen van de ethiekopdrachten A t/m F toe.



Portfolio

- Vragen?



Exit tickets

4 workshops ethiek:

- Tops / benefits
- Tips / suggesties

