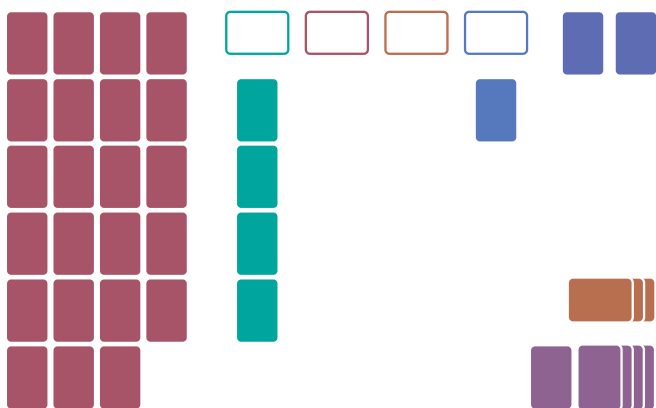


## Doel van het spel

De startup 'Ethics Inc.' heeft jullie benaderd om een betrouwbare AI-toepassing te ontwikkelen die voldoet aan Europese richtlijnen. Samen vormen jullie een denktank om de

startup te adviseren over hoe zij deze toepassing het best kunnen ontwikkelen. Het spel eindigt als er consensus is bereikt na één of meer rondes.

### Speloverzicht:



### Inhoud van de doos:

- 1x Spelregelkaart
- 1x AI-Kaart
- 4x Voorzitterkaart
- 23x Waardenkaarten met belangen
- 16x User storykaarten
- 5x AI-toepassingskaarten met uitdagingen
- 33x Wegingskaarten
- 8x Stakeholderkaarten

Een Uitwisbare stift per deelnemer (niet inbegrepen)

## Vorbereiding

Voordat het ontwerpspel kan beginnen moet worden bepaald wie de voorzitter van de denktank is. De voorzitter loodst de groep door de stappen heen en zal de discussies van de denktank begeleiden. Wanneer de voorzitter is aangewezen kunnen de voorbereidende stappen doorlopen worden:

1. Neem de **Spelregelkaart** en loop het verhaal en de richtlijnen gezamenlijk door.
2. Neem aansluitend de **AI-Kaart** en lees gezamenlijk de definitie van AI door.
3. Pak vervolgens de **AI-toepassingskaarten** met uitdagingen erbij en kies samen met de groep één toepassing uit. Dit bepaalt de uitdaging waarover jullie je advies als denktank gaan uitbrengen. Gevorderde spelers kunnen ook zelf een toepassing en uitdaging inbrengen.
4. Pak dan de **Stakeholderkaarten** erbij. Selecteer gezamenlijk de vier méést relevante stakeholders bij deze uitdaging. Gebruik de 'jokerkaart' om zelf waar nodig een stakeholder in te brengen.

5. Neem daarna de **Waardenkaarten met belangen**. Verspreid deze over een helft van de tafel. Leg alle kaarten dusdanig neer dat de overkoepelende waarden zichtbaar zijn (zoals menselijke autonomie en transparantie) en de gespecificeerde belangen gesloten op tafel liggen. Deze waarden en belangen komen overeen met de vereisten van de 'Ethics guidelines for trustworthy AI' van de Europese Commissie uit 2019.
6. Pak vervolgens de **User storykaarten**. Verdeel de **Invulkaarten** over de spelers. Leg de overige vier storykaarten op tafel: **Als... wil ik... om zo... bij...**
7. Deel dan aan iedere speler een set **Wegingskaarten** uit (S, M, L en XL). Leg de koffiekaart op tafel.

### Spelverloop

Nu kan het ontwerpspel echt beginnen! Als denktank moeten jullie samen discussiëren over de ideale AI-toepassing voor 'Ethics Inc.' en hierover tot consensus komen. De voorzitter leidt de discussie zo neutraal mogelijk.

### Totstandkoming spel

Het ethisch ontwerpspel is ontstaan uit een samenwerking tussen:

#### Stichting Toekomstbeeld der Techniek (STT)

STT is in 1968 opgericht door het Koninklijk Instituut van Ingenieurs (KIVI). STT voert al meer dan 50 jaar brede toekomstverkenningen uit op het snijvlak van technologie en samenleving die domeinoverstijgend en interdisciplinair zijn.

#### Stichting Koninklijk Nederlands Normalisatie Instituut (NEN)

NEN verbindt partijen zodat zij komen tot afspraken die worden vastgelegd in (inter)nationale normen, standaarden en richtlijnen. Daarnaast ondersteunt NEN partijen bij de toepassing en het gebruik van normen in de praktijk, door middel van trainingen en praktijkgidsen.

#### Lectoraat Artificial Intelligence van de Hogeschool Utrecht (HU)

Het Lectoraat Artificial Intelligence onderzoekt toepassingen van AI en data-gedreven innovatie. Het is van maatschappelijk belang de ontwikkeling van AI in goede banen te leiden en de grenzen van AI te verkennen en waar nodig vast te leggen.

## 1. Match de stakeholders en de belangen

*Welke belangen hebben de stakeholders bij deze uitdaging?*

- Maak voor elke stakeholder verschillende combinaties door een waardenkaart om te draaien en een user story te formuleren: **Als...(stakeholder) wil ik...(belang) om zo...(invuloefening) bij...(uitdaging)**. Schrijf bij de invuloefening de motivatie voor de stakeholder op met behulp van de invulkaarten en leg de kaarten op de juiste plek. Zorg ervoor dat je voor iedere waarde tenminste één combinatie maakt.
- Het is mogelijk dat een belang beter bij een andere stakeholder past. Het kan ook voorkomen dat er voor sommige combinaties geen user stories te formuleren zijn. Leg deze waardenkaarten weg.
- Doe dit totdat je voor elke stakeholder drie user stories hebt geformuleerd (of er 30 minuten voorbij zijn. Stel de klok bijvoorbeeld in op je mobiele telefoon).
- Bepaal of jullie als groep alles gezamenlijk doen of de stakeholders onderling verdelen.

## 2. Weeg de belangen

*Hoe zwaar wegen de belangen?*

- Bepaal voor elke user story hoe zwaar het belang en de motivatie wegen door gebruik te maken van de **Wegingskaarten (S,M,L,XL)**. Toe aan een break? Zet dan de koffiekaart in en het hele team neemt een pauze.
- Elke speler legt een **Wegingskaart** gesloten op de user story. Als iedereen een kaart heeft gelegd worden de kaarten gelijktijdig omgedraaid. Herhaal dit voor alle user stories. De voorzitter houdt de scores bij.
- Discussie: hierbij worden de spelers met de hoogste en laagste weging uitgenodigd om hun keuze te verklaren aan de rest van de groep. Is er al consensus mogelijk? Zo niet, dan volgt een tweede ronde waarin spelers opnieuw kaarten op kunnen leggen. Mogelijk moeten user stories worden aangepast. De voorzitter begeleidt deze stappen: doel is dat er consensus ontstaat en er voor elk belang een weging is waar de groep zich als geheel in kan vinden. Er zijn geen individuele winnaars.

## 3. Vergelijk de waarden onderling

*Hoe zwaar wegen de belangen onderling?*

- Bekijk de uiteindelijke scores waar consensus over is:
  - S en M: leg deze belangen even apart. Deze komen in een latere stap weer aan bod.
  - L en XL: deze belangen staan in deze stap centraal.
- Maak combinaties tussen de verschillende belangen: weeg voor elk belang hoe zwaar deze weegt ten opzichte van het belang van een andere waarde. Welk belang weegt dan zwaarder? Maak opnieuw gebruik van de **Wegingskaarten**. Herhaal dit voor meerdere combinaties.
- Discussie: Tussen welke waarden ontstaan er conflicterende belangen? Tussen welke waarden zijn overeenkomstige belangen? Welke waarden kunnen verenigd worden? Ook hier moet gestreefd worden naar consensus.

## 4. Construeer de ideale ontwerpprincipes

*Hoe ziet de ideale AI-toepassing eruit?*

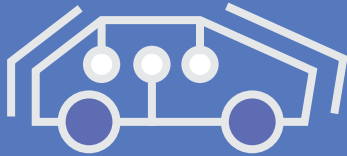
- Leg de **Belangenkaarten** die overblijven bij elkaar
- Hoe ziet de ideale AI-toepassing er nu uit, gegeven de uitdaging? Wat is er minimaal nodig voor een betrouwbare toepassing? (Minimum Viable Product).
- Pak de belangen met de S- en M-weging er weer bij: welke belangen kunnen geïntegreerd worden in de ideale toepassing?
- Kom tot een gezamenlijk advies voor de startup. Desgewenst kan de ideale AI-toepassing gevisualiseerd worden op een flipover.

## 5. Einde van het spel

*Welke inzichten kunnen we hieruit halen?*

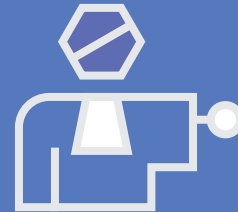
- Ga naar [www.ethicsinc.nl](http://www.ethicsinc.nl) en vul het formulier in. Hoe heb je het doorlopen van het proces ervaren? En welke inzichten levert dit op? De inzichten worden gebruikt voor verder onderzoek.

## Zelfrijdende auto



Een autonoom voertuig brengt ons van A naar B, waarbij de bestuurder niet meer nodig is om alles in de gaten te houden. De bestuurder moet echter wel op ieder moment kunnen ingrijpen, wanneer het voertuig daarom vraagt.

## Robotrechter



De robotrechter helpt een handje in de rechtspraak door automatisch de zaak van een verdachte te toetsen aan het wetboek en te vergelijken met miljoenen soortgelijke zaken.

## DNA-analyse



Het menselijk genoom bestaat uit miljarden basenparen, die met behulp van AI-technieken kunnen worden geanalyseerd. Zo kan voor iedereen een persoonlijke DNA-kaart worden gemaakt.

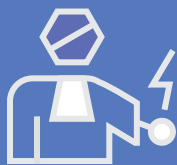
## Slimme deurbel



Steeds meer deurbellen zijn uitgerust met een camera. Zo kan met behulp van gezichtsherkenning worden bepaald wie er voor de deur staat en kan de deur automatisch openen.

## Onterecht schuldig

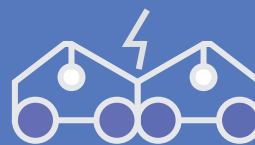
Robotrechter



De robotrechter beoordeelt een verdachte als schuldig en de rechter neemt dit oordeel een-op-een over. Later blijkt echter dat de veroordeling onterecht was. Hoe gaan we ethisch verantwoord met deze situatie om? En wat kunnen we hieruit leren bij de ontwikkeling van betere robotrechters in de toekomst?

## Een botsing

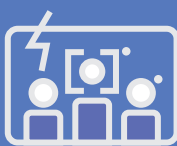
Zelfrijdende  
auto



Door de zelfrijdende auto zijn er bijna geen files meer en zijn er steeds minder ongelukken. Toch kunnen niet alle ongelukken voorkomen worden. Twee autonome voertuigen botsen op elkaar. Hoe gaan we ethisch verantwoord met deze situatie om? En wat kunnen we hieruit leren bij de ontwikkeling van betere autonome voertuigen in de toekomst?

## Iedereen in beeld

Slimme  
deurbel



Zelfs als je op vakantie bent kun je zien wie er voor de deur staat. De camera van de deurbel staat altijd aan en de fabrikant ontvangt alle beelden. De gezichten van mensen die toevallig op straat voorbij lopen worden echter ook herkend. Hoe gaan we ethisch verantwoord met deze situatie om? En wat kunnen we hieruit leren bij de ontwikkeling van betere gezichtsherkenningsoftware in de toekomst?

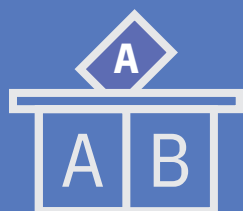
## Nooit meer ziek

DNA-analyse



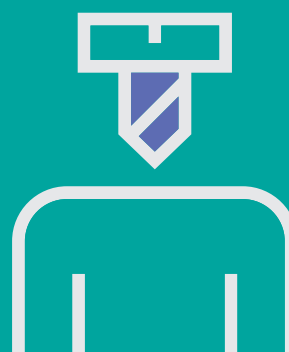
Mensen kunnen laten voorspellen of ze een verhoogd risico hebben op ernstige ziektes, of waar ze een grote kans op hebben om aan te komen overlijden. In sommige gevallen heeft het systeem het echter mis. Hoe gaan we ethisch verantwoord met deze situatie om? En wat kunnen we hieruit leren bij de ontwikkeling van betere DNA-analyses in de toekomst?

## Intelligente stemhulp

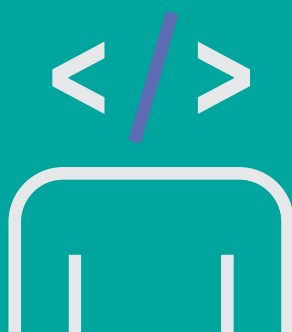


Op basis van een persoonlijk data-profiel kan een intelligente stemhulp met hoge mate van zekerheid bepalen welke politieke partij het beste bij iemand past. Dit vormt de basis voor het stemadvies.

## Fabrikant



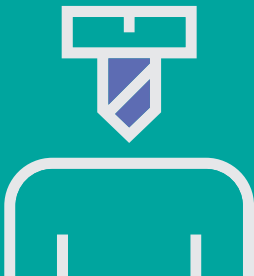
## Programmeur



## Aanbieder



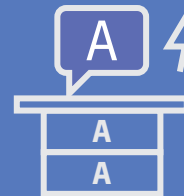
## Fabrikant



De fabrikant maakt nieuwe producten of diensten en houdt daarbij goed rekening met ethische en juridische kaders. Dat is een lastige klus, omdat niet altijd duidelijk is hoe de producten of diensten zullen worden toegepast door de aanbieder, of in welke context de gebruiker ermee zal werken.

## Opnieuw de favoriet

Intelligente  
stembulp



Een intelligente stembulp bespaart burgers veel tijd tijdens verkiezingen. Bij elke verkiezing moet de stembulp opnieuw worden vastgesteld aan de hand van de verkiezingsprogramma's. Sommige politieke partijen weten echter steeds beter wat ze in hun programma's moeten zetten om als favoriet naar boven te komen. Hoe gaan we ethisch verantwoord met deze situatie om? En wat kunnen we hieruit leren bij de ontwikkeling van betere stembulpen in de toekomst?

## Aanbieder



De aanbieder brengt het AI-product van de fabrikant naar de gebruiker en past deze aan voor de context van de gebruiker. De aanbieder weet vaak beter dan de fabrikant hoe het product zal worden gebruikt, maar wordt tegelijkertijd beperkt door de mogelijkheden die de fabrikant het product meegeeft.

## Programmeur

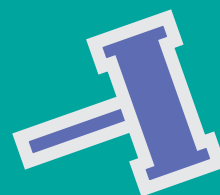


De programmeur is bekend met de mogelijkheden en beperkingen van de AI-software en weet op welke wijze data in het product wordt gebruikt. Dit stelt hem of haar in staat om problemen te signaleren en oplossingen te bedenken die anderen ontgaan.

Gebruiker



Wetgever



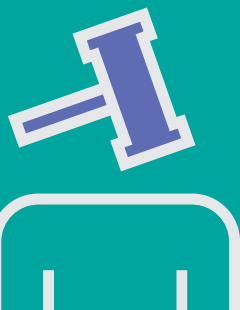
Toezichthouder



Maatschappij



## Wetgever



De wetgever kan de toepassing van AI stimuleren en kan door middel van wetten burgers en bedrijven beschermen tegen onwenselijke gevolgen. Daarnaast maakt de overheid als wetgever zelf in toenemende mate gebruik van AI, waarbij ze zich natuurlijk ook aan deze wetten dient te houden.

## Gebruiker



De gebruiker is een persoon die interacteert met het product waar de AI in is verwerkt. Vaak gaat dit in ruil voor betaling aan de producent of aanbieder, of is er een andere partij die voor het gebruik betaalt. In sommige gevallen wordt data van de gebruiker verzameld.

## Maatschappij



De maatschappij kan direct of indirect gevolgen ondervinden van vrijwel iedere AI-toepassing. Vaak is de indirecte impact niet meteen helder. Dit dwingt alle stakeholders diep na te denken over de maatschappelijke gevolgen voor burgers, samenleving, economie en klimaat.

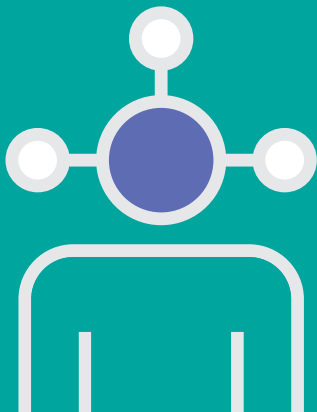
## Toezichthouder



Wanneer er negatieve gevolgen optreden door het gebruik van AI-toepassingen kan de toezichthouder bepalen welke stakeholders aansprakelijk gesteld moeten worden. De toezichthouder kan er daarnaast op toezien (ondersteund door toezichthoudende organen) dat wetten door alle betrokken stakeholders worden nageleefd.



## Jokerkaart



## Menselijke autonomie



AI-systemen moeten menselijke autonomie en beslissingen ondersteunen. AI-systemen mogen mensen niet onbewust misleiden, manipuleren of dwingen. Bij het ontwerp van AI-toepassingen moeten menselijke keuzes en mogelijkheden centraal staan.

## Menselijke autonomie



AI-systemen moeten menselijke autonomie en beslissingen ondersteunen. AI-systemen mogen mensen niet onbewust misleiden, manipuleren of dwingen. Bij het ontwerp van AI-toepassingen moeten menselijke keuzes en mogelijkheden centraal staan.

## Menselijke autonomie



AI-systemen moeten menselijke autonomie en beslissingen ondersteunen. AI-systemen mogen mensen niet onbewust misleiden, manipuleren of dwingen. Bij het ontwerp van AI-toepassingen moeten menselijke keuzes en mogelijkheden centraal staan.

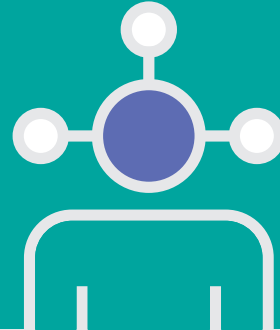
## Grondrechten

Menselijke  
autonomie



AI-systemen kunnen menselijke grondrechten zowel bevorderen als belemmeren. Zo kan het bijvoorbeeld ons recht op privacy eerbiedigen (door het beschermen van onze persoonsgegevens), maar ook schenden (door het misbruiken van onze persoonsgegevens). Bij de ontwikkeling van AI-systemen moeten mogelijke risico's beperkt of gerechtvaardigd kunnen worden om zo de rechten en vrijheden van mensen in een democratische samenleving te respecteren.

## Jokerkaart



Beschrijf zelf een stakeholder die je mist in het spel.

## Menselijk toezicht

Menselijke  
autonomie



Een AI-systeem mag de menselijke autonomie niet ondermijnen. Daarom is menselijk toezicht van groot belang. Dit kan worden verwezenlijkt door menselijke interventie te waarborgen in zowel de besluitcyclus van het systeem, als in het ontwerpproces van het systeem. Het is aan de mens om te bepalen wanneer en hoe het systeem in iedere specifieke situatie wordt gebruikt. Dit kan dus ook betekenen dat een keuze van een AI-systeem in een bepaalde situatie niet wordt opgevolgd.

## Menselijke controle

Menselijke  
autonomie



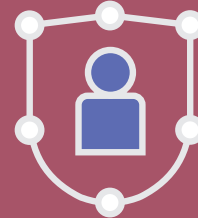
Gebruikers van AI-systemen moeten in staat zijn om het systeem in voldoende mate zelf te kunnen controleren of indien nodig beslissingen van AI-systemen aan te kunnen vechten. Mensen mogen niet onderworpen worden aan systemen die uitsluitend geautomatiseerde besluiten nemen, waarbij gebruikers aanzienlijk getroffen kunnen worden of waarbij het besluit van het systeem rechtsgevolgen kan hebben voor de gebruiker.

## Technische robustheid en veiligheid



AI-systemen moeten in staat zijn om onacceptabele schade te voorkomen en onbedoelde en onverwachte schade zoveel mogelijk te beperken. Hierbij moet de fysieke en mentale gezondheid van de mens te allen tijde gewaarborgd blijven.

## Technische robustheid en veiligheid



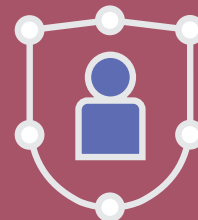
AI-systemen moeten in staat zijn om onacceptabele schade te voorkomen en onbedoelde en onverwachte schade zoveel mogelijk te beperken. Hierbij moet de fysieke en mentale gezondheid van de mens te allen tijde gewaarborgd blijven.

## Technische robustheid en veiligheid



AI-systemen moeten in staat zijn om onacceptabele schade te voorkomen en onbedoelde en onverwachte schade zoveel mogelijk te beperken. Hierbij moet de fysieke en mentale gezondheid van de mens te allen tijde gewaarborgd blijven.

## Technische robustheid en veiligheid



AI-systemen moeten in staat zijn om onacceptabele schade te voorkomen en onbedoelde en onverwachte schade zoveel mogelijk te beperken. Hierbij moet de fysieke en mentale gezondheid van de mens te allen tijde gewaarborgd blijven.

## Uitwijkplan en algemene veiligheid

Technische  
robuustheid  
en veiligheid



AI-systemen moeten in staat zijn om bij problemen een uitwijkplan in werking te stellen. Dit kan betekenen dat het systeem om menselijke tussenkomst vraagt voordat het verder gaat. Er moeten processen worden ingesteld om potentiële risico's in verschillende toepassingsdomeinen te verduidelijken en te beoordelen. Bij hoge risico's moeten er proactief veiligheidsmaatregelen worden ontwikkeld en getest. Onbedoelde gevolgen en fouten moeten worden geminimaliseerd.

## Weerbaarheid en beveiliging

Technische  
robuustheid  
en veiligheid



AI-systemen moeten weerbaar zijn tegen vijandige aanvallen, zodat ze niet door kwaadwillende individuen of organisaties misbruikt kunnen worden. Er moet daarom rekening worden gehouden met mogelijke onbedoelde toepassingen van AI en potentieel misbruik van een AI-systeem door kwaadwillende actoren. Er moeten stappen worden ondernomen om kwetsbaarheden in het systeem op te sporen en mogelijk misbruik van het systeem te voorkomen en te beperken.

## Betrouwbaarheid en reproduceerbaarheid

Technische  
robuustheid  
en veiligheid



AI-systemen moeten betrouwbaar zijn en optimaal kunnen werken met verschillende soorten input en in verschillende situaties. Op deze wijze kan het systeem worden gecontroleerd en kan onbedoelde schade voorkomen worden. AI-systemen moeten daarbij reproduceerbaar zijn en moeten hetzelfde gedrag vertonen wanneer een experiment onder gelijke omstandigheden wordt herhaald. Dat maakt het proces van het testen en reproduceren van gedrag eenvoudiger.

## Nauwkeurigheid

Technische  
robuustheid  
en veiligheid



AI-systemen moeten in staat zijn om op basis van gegevens en modellen correcte voorspellingen en aanbevelingen te doen en beslissingen te nemen. Door een gedegen ontwikkelings- en evaluatieproces kunnen onbedoelde risico's vanwege onjuiste voorspellingen worden beperkt en gecorrigeerd. Het AI-systeem moet hierbij aangeven hoe groot de kans op dergelijke fouten is.

## Privacy en datagovernance



AI-systemen moeten persoonsgegevens afschermen en privacyschade voorkomen. Hiervoor is het noodzakelijk om de kwaliteit en integriteit van de gebruikte gegevens te waarborgen. Om het recht op privacy te beschermen kunnen onder andere toegangsprotocollen worden ingesteld.

## Privacy en datagovernance



AI-systemen moeten persoonsgegevens afschermen en privacyschade voorkomen. Hiervoor is het noodzakelijk om de kwaliteit en integriteit van de gebruikte gegevens te waarborgen. Om het recht op privacy te beschermen kunnen onder andere toegangsprotocollen worden ingesteld.

## Privacy en datagovernance



AI-systemen moeten persoonsgegevens afschermen en privacyschade voorkomen. Hiervoor is het noodzakelijk om de kwaliteit en integriteit van de gebruikte gegevens te waarborgen. Om het recht op privacy te beschermen kunnen onder andere toegangsprotocollen worden ingesteld.

## Transparantie



Alle elementen die relevant zijn voor een AI-systeem moeten transparant zijn. Zowel de gegevens en het systeem als de bedrijfsmodellen. Dit betekent dat de processen, capaciteiten, doelen en beslissingen van AI-systemen inzichtelijk moeten zijn voor belanghebbenden.

## Kwaliteit en integriteit van gegevens

Privacy  
en data-  
governance



De kwaliteit van de gebruikte gegevenssets is cruciaal voor de prestaties van AI-systemen. Bij de verzameling van gegevens moeten vertekeningen, onnauwkeurigheden, fouten en vergissingen voorkomen worden vóórdat het AI-systeem met dergelijke gegevenssets getraind wordt. De gebruikte gegevenssets moeten daarom bij elke stap in het proces worden getest en gedocumenteerd. Er moet hierbij worden getracht zo min mogelijk gevoelige gegevens te gebruiken.

## Privacy en gegevensbescherming

Privacy  
en data-  
governance



AI-systemen moeten de privacy en gegevensbescherming van gebruikers gedurende de volledige levenscyclus van het systeem garanderen. Dit omvat niet alleen de informatie die oorspronkelijk door de gebruiker is aangeleverd, maar ook de informatie die in de loop van de interactie met het systeem over de gebruiker is gegenereerd. Het AI-systeem moet ervoor zorgen dat de verzamelde gegevens niet worden ingezet om gebruikers onwettig of onrechtvaardig te discrimineren.

## Traceerbaarheid

Transparantie



De beslissingen van AI-systemen moeten herleidbaar zijn. Dit betekent dat de gegevenssets en de processen waaruit de beslissing van het AI-systeem voortkomt zo goed mogelijk gedocumenteerd moeten worden. Dit geldt zowel voor de verzameling en indeling van de gegevens, als voor de gebruikte algoritmen. Dit maakt de controleerbaarheid van AI-systemen mogelijk, waardoor toekomstige fouten voorkomen kunnen worden.

## Toegang tot gegevens

Privacy  
en data-  
governance



Om persoonsgegevens te kunnen beschermen moet de toegang tot gegevens beheerd kunnen worden. Hiervoor moeten gegevensprotocollen worden ingesteld. In deze protocollen moet worden beschreven wie onder welke omstandigheden gegevens kan inzien. Alleen gekwalificeerd personeel met de juiste bevoegdheid en noodzaak om gegevens van personen in te zien, mag toegang tot dergelijke gegevens verkrijgen.

## Transparantie



Alle elementen die relevant zijn voor een AI-systeem moeten transparant zijn. Zowel de gegevens en het systeem als de bedrijfsmodellen. Dit betekent dat de processen, capaciteiten, doelen en beslissingen van AI-systemen inzichtelijk moeten zijn voor belanghebbenden.

## Transparantie



Alle elementen die relevant zijn voor een AI-systeem moeten transparant zijn. Zowel de gegevens en het systeem als de bedrijfsmodellen. Dit betekent dat de processen, capaciteiten, doelen en beslissingen van AI-systemen inzichtelijk moeten zijn voor belanghebbenden.

## Diversiteit en rechtvaardigheid



Bij de inzet van AI-systemen moeten individuen en groeperingen vrij zijn van onrechtvaardige discriminatie en stigmatisering. Gedurende de gehele levenscyclus van het AI-systeem moeten inclusie en diversiteit gewaarborgd worden. Alle betrokkenen moeten hierbij gelijk behandeld worden.

## Diversiteit en rechtvaardigheid



Bij de inzet van AI-systemen moeten individuen en groeperingen vrij zijn van onrechtvaardige discriminatie en stigmatisering. Gedurende de gehele levenscyclus van het AI-systeem moeten inclusie en diversiteit gewaarborgd worden. Alle betrokkenen moeten hierbij gelijk behandeld worden.

## Communicatie

Transparantie



Gebruikers van AI-systemen hebben het recht om te weten dat ze met een AI-systeem te maken hebben. Dit betekent dat AI-systemen zich niet als mensen mogen voordoen en herkenbaar moeten zijn. De capaciteiten en beperkingen van een AI-systeem moeten daarbij tijdig aan gebruikers gecommuniceerd worden. Denk hierbij aan de nauwkeurigheid van het systeem. Gebruikers moeten op elk moment op de hoogte gesteld kunnen worden van de redenen achter de resultaten van het AI-systeem.

## Verklaarbaarheid

Transparantie



De beslissingen van een AI-systeem moeten door mensen begrepen kunnen worden. Hiervoor moet een geschikte verklaring van het besluitvormingsproces van het AI-systeem tijdig beschikbaar zijn. Dit geldt zowel voor de technische processen van een AI-systeem, als de daaraan gerelateerde menselijke beslissingen. Deze verklaring moet afgestemd zijn op de mate van deskundigheid van de betrokken belanghebbende (bijv. leek, regelgever of onderzoeker).

## Toegankelijkheid en universeel ontwerp

Diversiteit en rechtvaardigheid



Mensen moeten ongeacht hun leeftijd, geslacht, vermogens of eigenschappen gebruik kunnen maken van AI-toepassingen. Het is in het bijzonder van belang dat deze technologie toegankelijk is voor mensen met een beperking. Het ontwerp van AI-systemen moet zijn gericht op het breedst mogelijke scala aan gebruikers, volgens relevante toegankelijkheidsnormen. Daardoor wordt voor alle mensen gelijke toegang tot AI-systemen mogelijk.

## Voorkomen van onrechtvaardige vertekening

Diversiteit en rechtvaardigheid



Bij de ontwikkeling en toepassing van AI-systemen moet vertekening worden tegengegaan. Dit kan onder andere door onvolledige gegevenssets niet te gebruiken. Op deze wijze kunnen onbedoelde vooroordelen en discriminatie tegen bepaalde individuen of groepen worden voorkomen. Daarnaast kan de diversiteit van meningen worden gewaarborgd door personeel met diverse achtergronden en uit verschillende culturen en disciplines in te zetten bij de ontwikkeling van AI-systemen.



## Diversiteit en rechtvaardigheid



Bij de inzet van AI-systemen moeten individuen en groeperingen vrij zijn van onrechtvaardige discriminatie en stigmatisering. Gedurende de gehele levenscyclus van het AI-systeem moeten inclusie en diversiteit gewaarborgd worden. Alle betrokkenen moeten hierbij gelijk behandeld worden.

## Maatschappelijk en milieuwelzijn



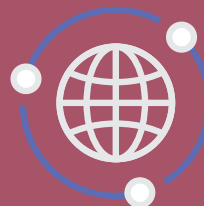
AI-systemen moeten worden gebruikt in het voordeel van alle mensen, inclusief toekomstige generaties. Daarnaast moet AI de samenleving in bredere zin bedienen, dus ook de natuur en het milieu. Duurzaamheid en ecologische verantwoordelijkheid van AI-systemen moeten worden aangemoedigd.

## Maatschappelijk en milieuwelzijn



AI-systemen moeten worden gebruikt in het voordeel van alle mensen, inclusief toekomstige generaties. Daarnaast moet AI de samenleving in bredere zin bedienen, dus ook de natuur en het milieu. Duurzaamheid en ecologische verantwoordelijkheid van AI-systemen moeten worden aangemoedigd.

## Maatschappelijk en milieuwelzijn



AI-systemen moeten worden gebruikt in het voordeel van alle mensen, inclusief toekomstige generaties. Daarnaast moet AI de samenleving in bredere zin bedienen, dus ook de natuur en het milieu. Duurzaamheid en ecologische verantwoordelijkheid van AI-systemen moeten worden aangemoedigd.

## Duurzame en milieuvriendelijke AI

Maatschappelijk en milieuvriendelijk



Het ontwikkelings-, installatie- en gebruiksproces van AI-systemen moeten zo milieuvriendelijk mogelijk gebeuren. Maatregelen om de milieuvriendelijkheid van de volledige ontwikkelings- en toeleveringsketen van het AI-systeem te waarborgen moeten worden gestimuleerd. Door kritisch onderzoek te doen naar het gebruik van hulpbronnen en de energieconsumptie tijdens de training kan bijvoorbeeld gekozen worden voor minder schadelijke opties.

## Participatie van belanghebbenden

Diversiteit en rechtvaardigheid



Bij de ontwikkeling van AI-systemen moeten belanghebbenden gedurende de gehele levenscyclus van het systeem geraadpleegd worden. Zo kunnen bijvoorbeeld gebruikers ook na de ingebruikname van het systeem om feedback gevraagd worden. Daarbij moeten mechanismen ingesteld worden om de participatie en betrokkenheid van ontwikkelaars te bevorderen bij het volledige ontwikkelingsproces van AI-systemen. Dit draagt bij aan een betrouwbare ontwikkeling van AI.

## Samenleving en democratie

Maatschappelijk en milieuvriendelijk



Bij de inzet van AI-systemen moet rekening worden gehouden met het effect op instellingen, de democratie en de samenleving als geheel. Het gebruik van AI-systemen moet zorgvuldig worden afgewogen, met name in situaties die verband houden met het democratische proces. Denk hierbij aan politieke besluitvorming en verkiezingen. Hierbij moeten de bredere maatschappelijke gevolgen van het gebruik van het AI-systeem worden onderzocht.

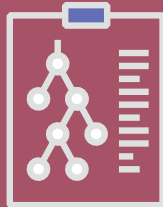
## Sociale gevolgen

Maatschappelijk en milieuvriendelijk



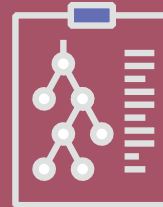
AI-systemen kunnen bijdragen aan de verslechtering van menselijke vaardigheden, zoals arbeidsvaardigheden en sociale vaardigheden. Wanneer mensen steeds meer interactie hebben met sociale AI-systemen kan dit gevolgen hebben voor onze sociale relaties en mentale gezondheid. De sociale gevolgen van AI-systemen moeten daarom zorgvuldig worden gemonitord en afgewogen. Denk hierbij ook aan de economische gevolgen, zoals het verlies van banen.

## Verantwoording



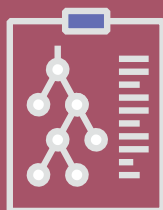
Beslissingen die door AI-systemen genomen worden moeten, afhankelijk van de context en de ernst van de gevolgen, door gebruikers aangevochten kunnen worden. Hiervoor is het noodzakelijk dat de entiteiten die verantwoordelijk zijn voor de beslissing identificeerbaar zijn.

## Verantwoording



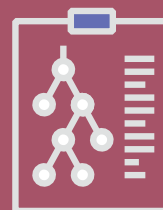
Beslissingen die door AI-systemen genomen worden moeten, afhankelijk van de context en de ernst van de gevolgen, door gebruikers aangevochten kunnen worden. Hiervoor is het noodzakelijk dat de entiteiten die verantwoordelijk zijn voor de beslissing identificeerbaar zijn.

## Verantwoording



Beslissingen die door AI-systemen genomen worden moeten, afhankelijk van de context en de ernst van de gevolgen, door gebruikers aangevochten kunnen worden. Hiervoor is het noodzakelijk dat de entiteiten die verantwoordelijk zijn voor de beslissing identificeerbaar zijn.

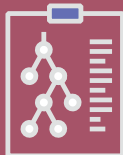
## Verantwoording



Beslissingen die door AI-systemen genomen worden moeten, afhankelijk van de context en de ernst van de gevolgen, door gebruikers aangevochten kunnen worden. Hiervoor is het noodzakelijk dat de entiteiten die verantwoordelijk zijn voor de beslissing identificeerbaar zijn.

## Verslaglegging van negatieve gevolgen

Verantwoording



De vaststelling, beoordeling en verslaglegging van de potentiële negatieve effecten van AI-systemen moeten worden gewaarborgd. Ngo's, vakverenigingen, klokkenluiders en andere entiteiten moeten worden beschermd wanneer zij gerechtvaardigde zorgen uiten over een AI-systeem. Het gebruik van effectbeoordelingen tijdens de ontwikkeling, de installatie en het gebruik van AI-systemen, kan helpen bij het minimaliseren van negatieve gevolgen.

## Controleerbaarheid

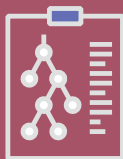
Verantwoording



De algoritmen, gegevens en ontwerpprocessen van AI-systemen moeten gecontroleerd kunnen worden. Dat betekent niet dat informatie over bedrijfsmodellen en intellectueel eigendom altijd openbaar beschikbaar moet zijn. Interne en externe controleurs moeten een evaluatieverslag kunnen opstellen die indien nodig openbaar gemaakt kan worden. Bij toepassingen die van invloed zijn op grondrechten moeten AI-systemen onafhankelijk kunnen worden gecontroleerd.

## Beroep

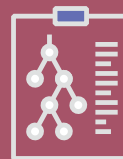
Verantwoording



Wanneer er bij het gebruik van een AI-systeem iets verkeerd gaat en er negatieve gevolgen zijn, dan moeten de betrokkenen in staat zijn om in beroep te gaan tegen de beslissing van een bestuursorgaan. Er moeten hiervoor voldoende mogelijkheden zijn om een bezwaarschrift in te kunnen dienen bij de rechtbank om de beslissing aan te vechten. Er moet hierbij bijzondere aandacht worden besteed aan kwetsbare personen of groepen.

## Afwegingen

Verantwoording



Bij de ontwikkeling van AI-systemen kunnen spanningen ontstaan tussen verschillende belangen en waarden, die tot onvermijdelijke afwegingen kunnen leiden. Alle beslissingen over te maken afwegingen moeten worden onderbouwd en goed worden gedocumenteerd. In situaties waarin geen ethisch acceptabele compromissen kunnen worden gevonden, mag de verdere ontwikkeling en het gebruik van het AI-systeem niet in die vorm worden voortgezet.

Onbelangrijk

S

Onbelangrijk

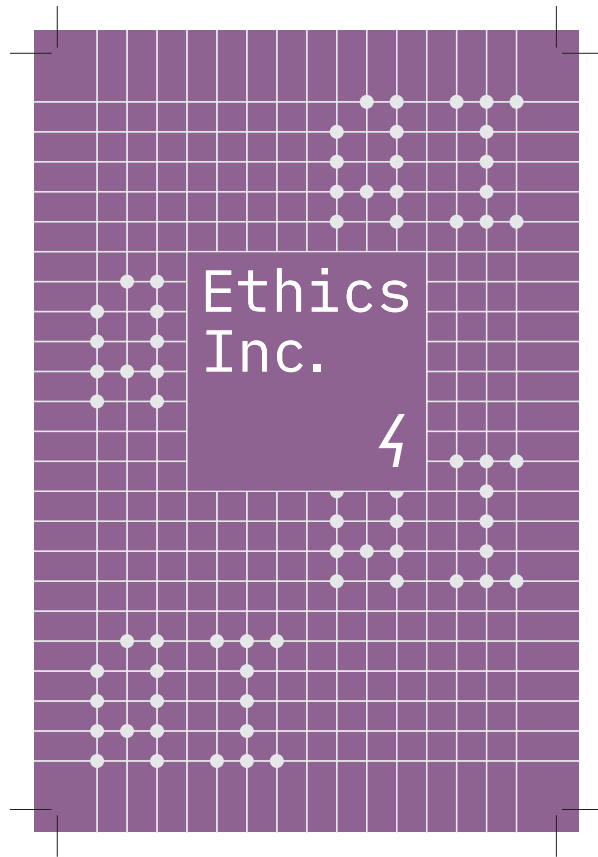
S

Onbelangrijk

S

Onbelangrijk

S



Onbelangrijk

S

Onbelangrijk

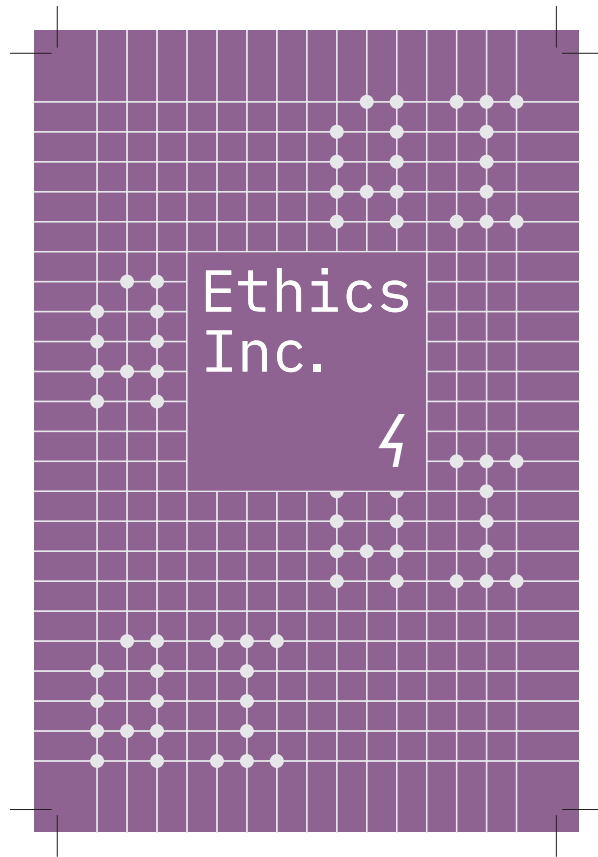
S

Onbelangrijk

S

Onbelangrijk

S





Niet erg  
belangrijk



Niet erg  
belangrijk

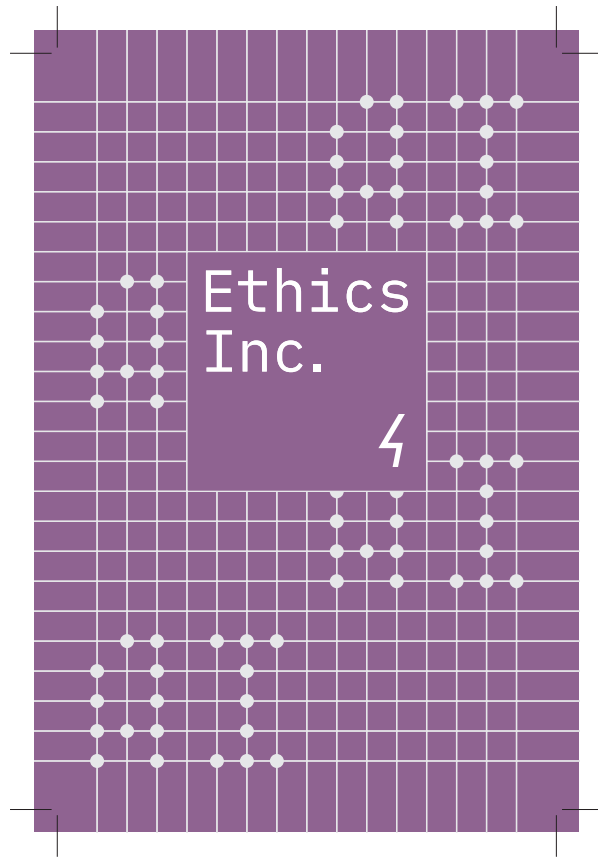


Niet erg  
belangrijk



Niet erg  
belangrijk





Niet erg  
belangrijk



Niet erg  
belangrijk



Niet erg  
belangrijk



Niet erg  
belangrijk





Belangrijk



Belangrijk

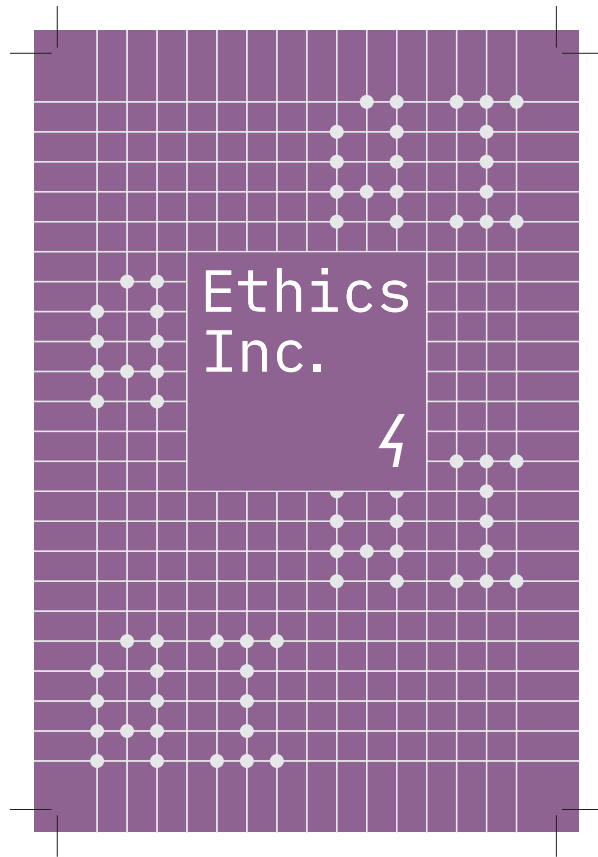


Belangrijk



Belangrijk





Belangrijk



Belangrijk



Belangrijk



Belangrijk







Heel erg  
belangrijk

XL

Heel erg  
belangrijk

XL

Heel erg  
belangrijk

XL

Heel erg  
belangrijk

XL



Heel erg  
belangrijk

XL

Heel erg  
belangrijk

XL

Heel erg  
belangrijk

XL

Heel erg  
belangrijk

XL



Tijd voor  
pauze



ALS...

om ZO...

WIL ik...



Print dit document dubbelzijdig, spiegelen aan de lange zijde.

...rij

Invulkaart

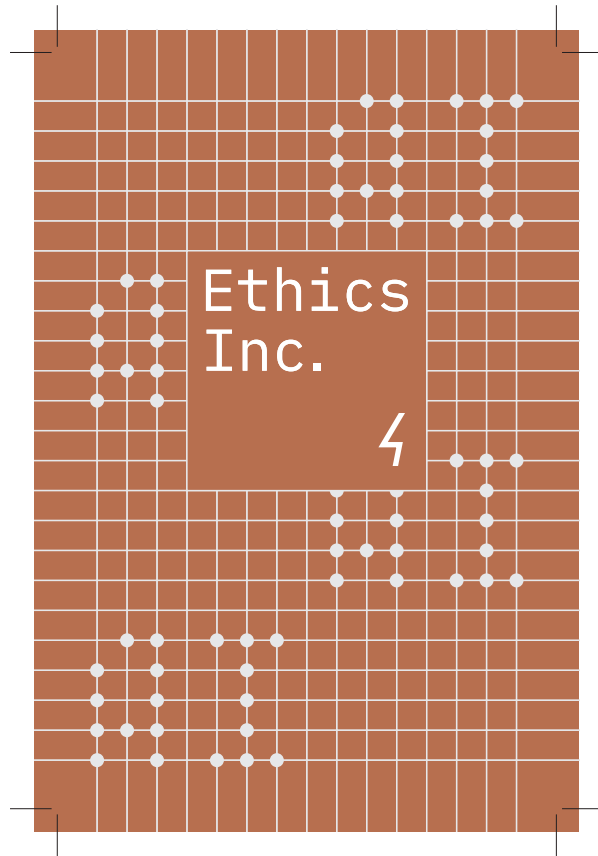
-----  
-----  
-----  
-----  
-----  
-----

Invulkaart

-----  
-----  
-----  
-----  
-----  
-----

Invulkaart

-----  
-----  
-----  
-----  
-----  
-----





Invulkaart

-----  
-----  
-----  
-----  
-----  
-----

Invulkaart

-----  
-----  
-----  
-----  
-----  
-----

Invulkaart

-----  
-----  
-----  
-----  
-----  
-----

Invulkaart

-----  
-----  
-----  
-----  
-----  
-----



Invulkaart

-----  
-----  
-----  
-----  
-----  
-----

Invulkaart

-----  
-----  
-----  
-----  
-----  
-----

Invulkaart

-----  
-----  
-----  
-----  
-----  
-----

Invulkaart

-----  
-----  
-----  
-----  
-----  
-----



## Invulkaart

---

---

---

---

---

---

---

---

1. Match de stakeholders en de belangen



2. Weeg de belangen



3. Vergelijk de waarden onderling



## Tip voor de voorzitter

Stap 1



Zo zou een user story eruit kunnen zien bij de zelfrijdende auto:

Als fabrikant wil ik de traceerbaarheid vergroten om zo de systeemfout te kunnen opsporen en verhelpen bij een botsing



## Tip voor de voorzitter

Stap 3



Mocht de discussie niet van de grond komen dan kun je de discussie aanjagen door dilemma's in te brengen, zoals het trolleyprobleem bij de zelfrijdende auto, een onterechte levenslange gevangenisstraf bij de robotrechter, het 'shoppen' van DNA voor je baby bij de DNA-analyse, het preventief oppakken van criminelen bij de slimme deurbel of misbruik en manipulatie bij de intelligente stembulp.

## Tip voor de voorzitter

Stap 2



Mocht de discussie niet lopen dan kun je de discussie aanjagen door een 'actualiteit' in te brengen: denk aan een fictief nieuwsbericht waarin een rechter een bepaalde uitspraak heeft gedaan of dat er onethische praktijken aan het licht zijn gekomen bij een organisatie.

## 4. Construeer de ideale ontwerp-principes



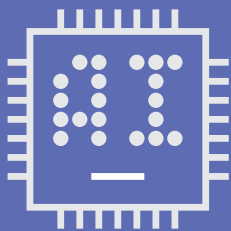
## Het verhaal

Ethics  
Inc.  
4

Spelregels

De startup 'Ethics Inc.' heeft jullie benaderd om een betrouwbare AI-toepassing te ontwikkelen die voldoet aan Europese richtlijnen. Samen vormen jullie een denktank om de startup te adviseren over hoe zij deze toepassing het best kunnen ontwikkelen.

## Artificiële Intelligentie (AI)

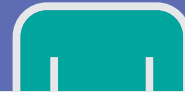


Artificiële Intelligentie (AI) verwijst naar het gedrag van machines dat door observatie en analyse van data leidt tot uitkomsten en acties die passend zijn voor het bereiken van specifieke doelen.

## De richtlijnen

- Je strijdt niet tegen elkaar, maar werkt met elkaar samen
- Iedereen heeft een gelijke stem in de denktank
- Werk toe naar consensus, pas dan kun je verder in het spel
- Iedere mening telt. Ook als je minder ervaring hebt met AI
- Probeer stereotypen over stakeholders te vermijden
- Ook de kleinere belangen tellen mee
- Je kunt een slechte daad niet wegstrepen door een goede daad
- Streef naar het best mogelijke resultaat
- Volg de instructies van de voorzitter op

## Tip voor de voorzitter



Stap 4

Zijn er bestaande standaarden waar de AI-toepassing aan moet voldoen? Over ethische afwegingen voor de ideale AI-toepassing zijn vereisten, richtlijnen en best practices opgenomen in internationale standaarden voor AI. Je vindt ze via [www.nen.nl/ai](http://www.nen.nl/ai). Andersom kunnen nieuwe inzichten weer in de standaarden worden verwerkt. Dat gaat via de norm-commissie Artificial intelligence & big data bij NEN.

Ethics  
Inc.

4