

aanpak begeleidingsethiek

Verslag workshop

Aanpak Begeleidingsethiek

Ongevallenhandboek

Nederlandse Arbeidsinspectie

ECP
Platform voor de InformatieSamenleving

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Nederlandse Arbeidsinspectie
Ministerie van Sociale Zaken en Werkgelegenheid

Voorwoord

De toenemende digitalisering brengt ingrijpende veranderingen met zich mee, in allerlei sectoren. Nieuwe mogelijkheden voor communicatie, monitoring en analyse roepen vragen op. Wat is de positie van de mens in deze omgeving, hoe staat het met data, met privacy, wat doen algoritmes en wat willen gebruikers eigenlijk? Begeleidingsethiek is ethiek die zich specifiek met dit soort vragen bezighoudt. Hij is gebaseerd op techniekfilosofie, die uitgaat van eeuwenoude verwevenheid tussen mens en technologie.

De Aanpak Begeleidingsethiek is een concrete aanpak waarin betrokkenen met elkaar in dialoog gaan over de effecten van de nieuwe technologie én de waarden die daarbij in het geding komen. Dat zijn vaak waarden gelieerd aan autonomie van de gebruiker, efficiëntie van het proces, transparantie van het algoritme, privacy, et cetera.

Tijdens de sessie komen verschillende stakeholders als gebruikers, ontwikkelaars, beleidsvormers en beslissers met elkaar in gesprek. Na de dialoog hebben de deelnemers ethische handelingsopties gegenereerd, waarvan verschillende vaak direct opgepakt kunnen worden. Er is gezamenlijk gekeken welke waarden we in het digitale domein belangrijk vinden en hoe we die willen verankeren en borgen in digitale processen en handelingen. Dit alles om ook bij verdere digitalisering op het vertrouwen van de samenleving en de participanten kunnen blijven rekenen. Dit is niet in één stap te realiseren het is een continu proces waarin deze workshop een schakel is.

Workshop aanpak begeleidingsethiek

Initiatiefnemers: Joris Sauter en Marjolein Ham (Nederlandse Arbeidsinspectie)
Moderatoren: Daniël Tijink (ECP) en Edwin Borst (ECP)

Op maandag 24 februari 2025 organiseerde de Nederlandse Arbeidsinspectie (hierna: NLA) op initiatief van Joris Sauter (NLA) een sessie begeleidingsethiek over de casus Ongevallenhandboek. Aan de workshop namen 13 deelnemers (zie bijlage voor specificatie) deel onder leiding van Daniël Tijink (ECP) en Edwin Borst (ECP).

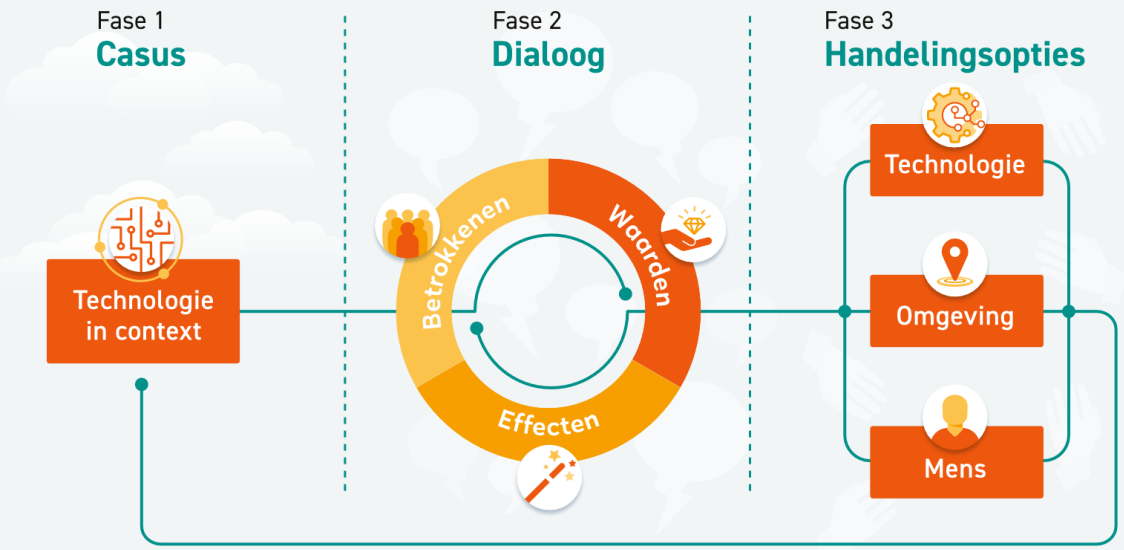
De doelstelling is tweeledig:

- Leren tijdens de workshop
  - Welke effecten, waarden en actoren worden genoemd
  - Welke handelingsopties zijn er om de toepassing te verbeteren
- Leren over de Aanpak Begeleidingsethiek
  - Hoe kan de organisatie deze methodiek gebruiken rondom ethische kwesties en technologie-implementaties

Het verslag bevat de volgende elementen:

- Toelichting op de Aanpak Begeleidingsethiek
- Weerslag van de workshop

Aanpak begeleidingsethiek



De aanpak bevat de volgende fasen:

Fase 0 Introductie

Introductie over de doelstelling en een toelichting op het model en het gedachtengoed

Fase 1 Toelichting

Hoe ziet de technologie eruit en in welke context wordt deze toegepast

Fase 2 Dialoog

- Een korte ronde waarin de deelnemers aan de workshop de betrokken actoren benoemen
- Brainwrite waar deelnemers mogelijke effecten benoemen en bespreken
- Benoemen van waarden die een rol spelen bij die effecten

Fase 3 Handelingsopties

In subgroepjes gaan de deelnemers op zoek naar handelingsopties vanuit de technologie, de omgeving en het individu.

Fase 1

Technologie en context

Uitleg over de chatbot Ongevallenhandboek

De casus betreft een generatieve AI-toepassing om het Ongevallenhandboek toegankelijker te maken voor professionals van de Arbeidsinspectie (NLA). Het Ongevallenhandboek beschrijft de handelswijze van de inspecteur bij het inspecteren van een arbeidsongeval. Het handboek bestaat uit praktische adviezen en richtlijnen alsmede wettelijke kaders. Het handboek heeft op dit moment de vorm van een (doorzoekbaar) document op het intranet van de NLA.

De (demo) generatieve AI toepassing die is gebouwd is te gebruiken voor meerdere use cases, in de sessie hebben we ons gericht op het Ongevallenhandboek. Er is gebruik gemaakt een bestaand Large Language Model en daar is een toevoeging voor de NLA op gebouwd op basis van het handboek. Tijdens de sessie werd dit gepresenteerd. Aan de hand van een prompt komt het model tot antwoorden in de Nederlandse taal gebaseerd op de kennis van het handboek.

Tijdens de toelichting werden ook veel andere suggesties gedaan voor mogelijke toepassingen van generatieve AI binnen de NLA, bijvoorbeeld het beoordelen van inspectieformulieren of het schrijven van adviezen. Er is bewust gekozen om te beginnen met het handboek als een relatief eenvoudige en veilige toepassing. Dit vormde ook de scope van deze sessie.

Fase 2

Dialoog

In deze tweede fase gaan de deelnemers in gesprek over wie er betrokken zijn bij de inzet en het gebruik van het Ongevallenhandboek. Ook buigen zij zich over de positieve en negatieve effecten van deze inzet, en benoemen zij belangrijke waarden waar rekening mee gehouden moet worden bij de inzet van een dergelijke toepassing.

Betrokkenen:

Bij de betrokkenen is de vraag wie er betrokken is of geraakt wordt door de case. De deelnemers aan tafel vertegenwoordigen al een deel van die betrokkenen. Ze noemen de volgende betrokkenen:

- Eindgebruiker (inspecteur)
- Projectleider
- Deployer (Data Science Lab)
- Kenniscentrum (bronhouder)
- Juristen
- Informatiehuishouding
- Geïnspecteerden
- Beveiligingsautoriteit (BVA)
- Directeur Informatievoorziening
- Directeur-generaal
- Slachtoffer ongeval
- Privacy- en informatiebeveiligingsfunctionarissen
- Teammanager
- Ondernemingsraad
- Opleidingen
- Supportafdeling
- Architectuurraad
- Communicatie
- Ministerie van Sociale Zaken en Werkgelegenheid

De vraag aan de deelnemers was om in het vervolg van de sessie ook te proberen vanuit deze perspectieven hun inbreng te geven.

Effecten:

De vraag is hier welke positieve en negatieve effecten de invoering en het gebruik van het Ongevallenhandboek hebben.

Positieve effecten

- Maakt het werk makkelijker (concreet vraag leidt tot een concreet antwoord)
- Leidt tot een beter geïnformeerde inspecteur
- Betere focus inspecteur, GAI pakt de andere zaken op
- Uniformiteit in toepassing handboek
- Bevorderen handelingssnelheid inspecteur
- Makkelijker – je krijgt sneller, preciezere info uit het handboek
- Je sluit goed aan bij maatschappelijke verandering, de NLA is up-to-date
- Professionaliteit (je hebt je zaakjes beter op orde)
- Pilot als geschikte vorm van stap-voor-stap innovatie
- GAI laat zien hoe het tot een antwoord komt

Negatieve effecten

- GAI maakt allerlei onderscheiden, gebeurt dat wel goed
- Missen van informatie als het handboek niet goed werkt
- Is Ctrl + F niet goed genoeg?
- Als mensen blind de NLA volgen, wat doet dat met de betrouwbaarheid van het oordeel van de NLA.
- Als GAI opleiding vervangt, neemt de kennis van de NLA af.
- Risico op hack
- Afhankelijkheid bij inspecteur van tool
- Gemakzucht bij inspecteur
- Mogelijk in strijd met privacy en wettelijke afspraken
- Minder uniformiteit
- Gebrek aan eigenaarschap, uitkomst is “niet van de inspecteur”
- GAI laat niet zien hoe het tot een antwoord komt

Waarden:

De relevante waarden zijn door de facilitators geïnventariseerd terwijl de deelnemers bezig waren met het benoemen van effecten. Achter veel effecten gaan waarden schuil. Die lijst is gedeeld met de deelnemers en door hen besproken en aangevuld. Dat leverde de volgende lijst met waarden op, waarbij de vetgedrukte waarden zijn geprioriteerd door de deelnemers:

Waarden	
Arbeidsvreugde	Veiligheid
Kwaliteit van handhaving	Continuïteit van het proces
Efficiëntie	Legaliteit
Rechtvaardigheid	Reputatie
Continuïteit organisatie	Innovativiteit
Vertrouwen	**Betrouwbaarheid**
Onafhankelijkheid	Transparantie
Vakmanschap	Zorgvuldigheid
Onpartijdigheid	

Fase 3

Handelingsopties

In het proces van het opstellen van de handelingsopties zijn de waarden uit de vorige fase meegenomen. De deelnemers kregen een uitleg over de verschillende categorieën binnen de handelingsopties. Het kan bijvoorbeeld gaan over het ontwerp van de technologie, het aanpassen van de omgeving waarbinnen de toepassing functioneert en het gedrag van mensen. De deelnemers worden in drie werkgroepen verdeeld en gaan ieder aan de slag met een van deze drie gespreksonderwerpen. Dat leidt tot handelingsopties per categorie.

Technologie:

- Geef de tool een naam (Davy, Daivy)
- Integreer voice
- Validatie -> gebruiker geeft feedback (eventueel direct via interface chatbot)
- A/B-test onder inspecteurs: met en zonder tool
- Toegankelijk maken in aparte app
- Continuous development
- Doorlopend opleiden van inspecteurs in gebruik van tool
- Uitgangspunt: tool begeleidt inspecteur, is assistent. Vraagt bijvoorbeeld door, en geeft suggesties: heb je hier al aan gedacht?
- Redeneerstappen intrainen
- Snel en gemakkelijk in gebruik (bijvoorbeeld te gebruiken tijdens reistijd in auto)
- Koppeling met zaaksysteem, bijvoorbeeld met soortgelijke zaken (later)
- Audit trail
- Omgaan met hallucinatie, onder andere:
    - Laat model op de achtergrond 5 keer de vraag herschrijven en beantwoorden, destilleer een overeenkomend antwoord, toon dat aan de gebruiker
    - Neem mee in opleiding
    - Laat de inspecteur in de lead
    - Check altijd url's die in een antwoord staan
- Aangeven van bron

Omgeving:

Betrouwbaarheid

- Er moet een testsysteem komen dat de betrouwbaarheid test. Met experts, met steekproeven
- Het is belangrijk een goede inwerkperiode te organiseren
- Check ook het prompten door verschillende inspecteurs. Wat is de impact van een persoonlijke touch
- Maak beleid voor data, waar komt het terecht, wie mag erbij etc.

Efficiëntie

- Kwaliteitssysteem dat checkt po echte efficiëntie, dus niet alleen van de directe effecten, maar ook de neveneffecten
- Zorg voor een goede trouble-shooter, dat de feedback die er zeker komen gaat bij praktijkgebruik, goed wordt opgepakt en leidt tot verbeteringen
- Wat doe je met de (efficiency)winst: besteden aan andere dingen binnen NLA: bijvoorbeeld oppakken actieve dossiers: grotere inspectiedekking, met minder mensen toekunnen

Arbeidsvreugde

- Feedbacksysteem inrichten wat de invoering ervan doet met de inspecteurs
- Zorgen voor een goede opleiding voor het gebruik en de randvoorwaarden van GAI
- Zorg dat de GAI echt helpt in het proces van inspectie: dus direct toepasbaar, informatie helpt om sneller naar de volgende stap te komen
- Afspraak om het allemaal te gaan gebruiken

Kwaliteit van handhaving

- Zorgen voor goede opleiding
- Toetsen kwaliteit van handhaving
- Een beeld krijgen van het verschil met en zonder GAI – Mens

Mens:

Arbeidsvreugde

- Testen:
    - Wat maakt wie blij? Wat is de behoefte van de gebruiker?
    - Evalueren, doet het product wat je verwacht?
    - Plan-do-check-act-cyclus: 0-meting, helpt om de toegevoegde waarde te bepalen voor gebruikers.

Kwaliteit van handhaving

- Training is key (herhalen) -> breder opleiden dan alleen inspecteurs, ook bijvoorbeeld leidinggevenden.
    - Wat voer je in? En wat krijg je voor antwoorden? Hoe kun je daarvan leren?
- Onderdeel maken van bestaande opleidingstrajecten
- Mensenkennis vs digitale informatie, blijf vertrouwen op menselijk oordeel en dus:
    - Kwaliteitscheck
    - Borging controle door projectleider

Efficiëntie

- Training + intervisie, ervaringen delen zie ook hierboven.
- Meerwaarde creëren en zichtbaar maken
- Juiste taken laten uitvoeren door AI
    - Wat kan het wel en wat kan het niet
- Beoordeling van kwaliteit moet makkelijk zijn/ondersteund worden
- Begrijpelijke taal (B1)
- Ambassadeurs die helpen om het project een succes te maken

Betrouwbaarheid

- Leren om de juiste vraag te stellen via training
- Menselijke maat: altijd mens die oordeelt of AI antwoord juist is
- Zorgen voor een duidelijke disclaimer bij de resultaten (vertrouw op jezelf)
- Laten weten waar het vandaan komt: zorgen voor betrouwbare links naar achterliggende stukken documentatie

Terugblik en afronding

Aan het einde van de workshop wordt teruggekeken en besproken wat de Aanpak Begeleidingsethiek concreet heeft opgeleverd. De deelnemers geven aan dat het veel nieuwe inzichten oplevert en dat er in korte tijd veel wordt gerealiseerd. Daarnaast wordt de brede groep van deelnemers als meerwaarde ervaren. De combinatie van een duidelijke casus, enthousiaste deelnemers, en een aanpak die zorgt voor duidelijke uitkomsten in een korte tijd was dus een succes.

De deelnemers

1. **Jered Vermeent** NLA
2. **Dith Zandbergen** NLA
3. **Linde Grootswagers** NLA
4. **Joris Sauter** NLA
5. **Marcel Philipse** NLA
6. **Noud Ketelaars** NLA
7. **Davy van Aalst** NLA
8. **Sebastiaan Jacobs** NLA
9. **Anonieme deelnemer**
10. **Foske Kleima** NLA
11. **Ron Dotsch** NLA
12. **Lourens Snoek** NLA
13. **Wesley van der Zaag** NLA

Moderatoren:
Daniël Tijink (ECP), Edwin Borst (ECP)

aanpak begeleidings ethiek

ECP Platform voor de InformatieSamenleving

Verslag workshop
Aanpak Begeleidingsethiek
Ongevallenhandboek

Meer informatie over de Aanpak Begeleidingsethiek, waaronder dit verslag, vindt u op **www.begeleidingsethiek.nl**

Foto omslag: Shutterstock.com