



**aanpak**   
**begeleidings  
ethiek**

**Verslag workshop**  
Aanpak Begeleidingsethiek  
**RVO - Casus**  
**GenAI voor content**

Rijksdienst voor Ondernemend Nederland

**ECP**

Platform voor de  
InformatieSamenleving



Rijksdienst voor Ondernemend  
Nederland



Rijksdienst voor Ondernemend  
Nederland



Platform voor de  
InformatieSamenleving



Ministerie van Binnenlandse Zaken en  
Koninkrijksrelaties

# 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 geformuleerd, 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 te 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:** Thijs van der Wal (RVO), Joleen van der Zwan (RVO)

**Moderatoren:** Edwin Borst – ECP (moderator), Iris Visser – Digitaal in Balans (co-moderator)

Op 14 januari 2025 organiseerde RVO op initiatief van Thijs van der Wal (RVO) en Joleen van der Zwan (RVO) een sessie begeleidingsethiek over de casus 'generatieve AI voor contentcreatie'. Aan de workshop namen 11 deelnemers (zie bijlage voor specificatie) deel onder leiding van Edwin Borst (ECP) en Iris Visser (Digitaal in Balans).

## **De doelstelling is tweeledig:**

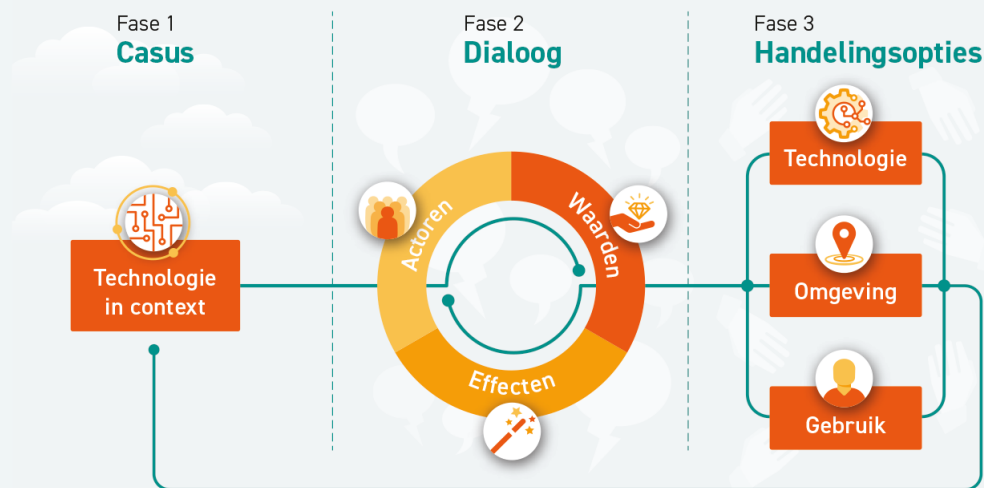
- ▷ Leren tijdens de workshop
  - Welke effecten, waarden en actoren worden genoemd
  - Welke handelingsopties zijn er om de toepassing te verbeteren
- ▷ Leren over de Aanpak Begeleidingsethiek
  - Hoe kan de organisatie deze methodiek gebruiken rondom ethische kwesties en technologie-implementaties

## **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

RVO heeft 14 teams met contentspecialisten. Zij werken voor klanten aan content (tekst, beeld, social posts) op de website (bijvoorbeeld pagina's waarop regelingen worden uitgelegd, blogs en praktijkverhalen), social media (LinkedIn, Instagram, YouTube) en redigeren brieven. Daarnaast zijn er corporate contentspecialisten voor arbeidsmarkt- en interne communicatie.

De contentspecialisten ervaren een hoge werkdruk. Er is continu behoefte aan hoogwaardige content die voldoet aan alle eisen. Nieuwe content moet vaak snel af. Zij werken volgens vaste afspraken over formats, taalniveau en taalgebruik.

De verwachting is dat generatieve AI de contentspecialisten kan helpen kwalitatief betere content te maken, in kortere tijd. Om dit te onderzoeken is een eerste test gedaan met een prototype op een veilige omgeving, geleverd door DICTU. De usecase: een tool die tekst kan schrijven en verbeteren op basis van ruwe input en volgens de RVO-richtlijnen en schrijfwijzer.

De resultaten van deze eerste test zijn positief: de toepassing is goed in staat eerste versies van de teksten op te leveren op basis van ruwe input. De contentspecialisten maakten de eerste versies gereed voor publicatie.

De volgende stap is een pilot met een grotere groep contentspecialisten en meer typen teksten.



## Fase 2

# Dialogoog

In deze tweede fase gaan de deelnemers in gesprek over wie er betrokken zijn bij de inzet en het gebruik van generatieve AI voor contentcreatie. 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:

- Contentspecialisten van RVO
- Ondernemers
- Andere burgers
- Hoofdredactie
- Ketenpartners – adviseurs inhoudelijke kennis
- DICTU
- Medewerkers automatisering/techniek - zoals technisch beheer en ondersteuning
- Vertalers
- Microsoft
- Open AI
- Data scientists
- Tegenlezers
- Content coördinator
- IMP Regisseur
- Communicatie adviseur
- Opdrachtgevers van RVO
- Toezichthouders (AP)
- Concurrenten open AI
- Directie RVO
- Juridische Zaken
- CIP
- Processpecialisten
- Team klantkanalen en middelen
- BZK
- Politiek/Tweede Kamer
- Publiek/Media
- Data-‘labellers’

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 generatieve AI bij contentcreatie hebben.

| Positief effect                                                                    | Waarde                                                |
|------------------------------------------------------------------------------------|-------------------------------------------------------|
| Tool geeft houvast                                                                 | Vakmatige kwaliteit                                   |
| Tijdswinst                                                                         | Efficiëntie                                           |
| Ideaal hulpmiddel                                                                  | Effectiviteit                                         |
| Standaardisatie van prompts over AI middelen heen                                  | Stabiliteit, effectiviteit, efficiëntie, uniformiteit |
| Meer inspiratie                                                                    | Werkplezier, Creativiteit                             |
| Meer tijd voor kwaliteit                                                           | Efficiëntie, werkplezier                              |
| Opschaling, meer impact                                                            | Effectiviteit                                         |
| Tijd om bestaande content goed te onderhouden                                      | Efficiëntie, kwaliteit                                |
| Kostenbesparing                                                                    | Efficiëntie                                           |
| Afname gebruik andere AI                                                           | Veiligheid                                            |
| Meer werkplezier – aantrekkelijke werkgever, imago                                 | Continuïteit, ondernemend                             |
| Direct feedback op teksten                                                         | Efficiëntie, kwaliteit,                               |
| Meer uniforme output, in stijl en inhoudelijk                                      | Vakmatige kwaliteit, betrouwbaarheid                  |
| Voelt veiliger van chatGPT                                                         | Veiligheid                                            |
| Afgeschermdde omgeving                                                             | Veiligheid                                            |
| Minder richtlijnen nodig dan bij chatGPT omdat je nu in een veilige omgeving werkt | Efficiëntie/werkplezier                               |
| Nieuwe technologie inzetten                                                        | Innovatie, ondernemend                                |
| AI-leveranciers verdienen hier geld aan                                            | Winstgevendheid/ continuïteit                         |



| Negatief effect                                                           | Waarde                                             |
|---------------------------------------------------------------------------|----------------------------------------------------|
| Beeldvorming van gebruik van AI in teksten                                | Eerlijkheid, betrouwbaarheid                       |
| Content specialist kan overbodig worden                                   | Bestaanszekerheid, waardering                      |
| Het is niet ideaal. Menselijke moderatie is essentieel                    | Vakmatige kwaliteit                                |
| Afhankelijkheid van AI leveranciers                                       | Onafhankelijkheid                                  |
| Frustratie door complexiteit<br>Frustratie door tool die niet goed werkt  | Efficiëntie, onafhankelijkheid, vrijheid           |
| Nieuw proces voor functioneel beheer                                      | Vernieuwing, werkplezier, verantwoordelijkheid     |
| Minder werkplezier wanneer je de eerste draft niet zelf kunt maken        | Werkplezier                                        |
| Minder bevlogen medewerkers                                               | Werkplezier                                        |
| Belasting van milieu door AI                                              | Duurzaamheid                                       |
| Hogere eisen aan/werkdruk voor medewerkers                                | Efficiëntie/ Autonomie, werkplezier, menselijkheid |
| Mogelijk moet je het erbij vermelden dat teksten door AI worden gegeneerd | Transparantie, betrouwbaarheid, uitlegbaarheid     |
| Mogelijke schending van auteursrechten                                    | Eerlijkheid, rechtvaardigheid, rechtmatigheid      |

#### 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 rechts bovenstaande waarden op. De deelnemers kozen 2 á 3 waarden die voor hen, in deze casus, het meest waardevol zijn:

| Waarden          |                       |                                   |
|------------------|-----------------------|-----------------------------------|
| Kwaliteit (4x)   | Transparantie (1x)    | Onafhankelijkheid (1x)            |
| Efficiëntie (3x) | Duurzaamheid (1x)     | Bestaanszekerheid/waardering (1x) |
| Werkplezier (3x) | Veiligheid (1x)       |                                   |
| Betrouwbaar (2x) | Rechtvaardigheid (1x) |                                   |
| Innovatie (2x)   | Effectiviteit (1x)    |                                   |

Geprioriteerde kernwaarden zijn: **Kwaliteit, efficiëntie en werkplezier.**



## 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:

- Werkinstructie voor medewerkers over gebruik tool
- Makkelijkere interface maken met knop voor informatie
- Meerdere documenten en formats erin verwerken
- Meerdere modellen te selecteren (en dat vraagt extra uitleg)
- Meerdere output mogelijkheden en bestandsformaten mogelijk maken
- Meerdere modellen testen
- Meerdere uitkomsten van verschillende modellen integreren (maar dit geeft ook consistentie problemen)
- Technologie blijven updaten, monitoren, ontwikkelingen volgen (team paraat hebben die daar voor verantwoordelijkheid draagt)
- Snelheid bevorderen (dat is nu een trade-off met kwaliteit)
- Spelen met prompts ingeven (en die prompts kunnen opslaan en delen delen met collega's). Dit vraagt om promptmanagement, bijvoorbeeld met een super-user
- Tekst meteen kunnen zien en kunnen kopiëren
- Integratie bouwen met Drupal en/of Office
- Logging (hier botst mogelijk transparantie met werkplezier)



## Omgeving:

- Visie op gebruik generatieve AI voor contentcreatie + bijhorende werkinstructie
  - Verwachtingen uitdragen: is het gebruik van de tool verplicht of niet?
  - Uitdragen hoe de technologie voldoet aan wet- en regelgeving en veilig is in gebruik
- Intervisie en evaluatiemomenten organiseren om feedback op te halen bij medewerkers
- De effectiviteit en efficiëntie monitoren om de tijdswinst inzichtelijk te maken
- Een kwaliteitsaudit uitvoeren
- Waardering klanten meten
- De tool doorontwikkelen t.b.v. betere kwaliteit, effectiviteit en efficiëntie (een ontwikkelaar toevoegen aan het team voor structurele doorontwikkeling)

## Mens:

### Werkplezier:

- Via HR/management de optie bieden om doorlopend aan te geven hoe het ervoor staat qua werkplezier - monitoring
- Goede uitleg en mensen laten meedenken en betrekken
- Zorgen dat systeem zo intuïtief werkt dat je geen handleiding nodig hebt (handelingsoptie voor technologie)
- Afspraken over vrij te komen tijd
- Goede instructie + training + begeleiding bij implementatie
- Zorgen dat werk dat blijft bij contentspecialisten de creativiteit aanspreekt

### Kwaliteit:

- Opleiding RVO-richtlijnen en standaarden + relatie tot AI tool
- Principe: tekst wordt altijd door contentspecialist nabewerkt en gaat vervolgens door in het huidige proces (vaak check bij communicatie en inhoud, altijd eindredactie)

### Efficiëntie:

- Er als contentspecialist een gewoonte van maken om de tool op het juiste moment in het schrijfproces in te zetten

### Betrouwbaarheid:



- Goed begrip van generatieve AI en verwachtingen t.a.v. feitelijke juistheid

# 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.

Een deelnemer is onder de indruk van hoeveel betrokken partijen er zijn en had niet voorzien dat er zoveel handelingsperspectieven uit deze sessie zouden komen. Het is een mooie oogst waarmee we zeker verder kunnen.

Wat nog nodig is, is het aanvullen van de handelingsperspectieven op basis van de overige waarden. De top 3 waarden zijn natuurlijk belangrijk, maar we willen graag compleet zijn. Het totaal aan handelingsperspectieven zal worden vertaald naar een praktisch plan dat onderdeel wordt van het implementatieplan.

In de sessie worden geen besluiten genomen over het al dan wel of niet operationaliseren van generatieve AI voor contentspecialisten. De opgedane kennis is weergegeven in dit verslag, welke vervolgens in formele besluitvormingsprocessen en organen zal worden ingebracht. Deelnemers aan deze sessie worden van dit proces op de hoogte gehouden.



# De deelnemers

- Digitaal Innovatie Lab (RVO)
- iNova (RVO)
- CxO werkgroep Generatieve AI (RVO)
- Data en Ethiek Raadkamer (RVO)
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK)

1. **Thijs van der Wal** RVO - Enterprise Architect - AI & Innovatie
2. **Joleen van der Zwan** RVO – Sogeti – Projectleider Innovatie – Digitaal Innovatie Lab
3. **Henk van der Zeijden** RVO
4. **Fiënna Veenstra** RVO
5. **Joost Mol** RVO
6. **Gijs Leppers** RVO
7. **Anita Vocht** RVO
8. **Esther Besselink** RVO
9. **Humphrey van Berkel** RVO
10. **Simone Stubenitsky** RVO
11. **Mildo van Staden** Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

## **Moderatoren:**

Edwin Borst (ECP), Iris Visser (Digitaal in Balans)





Platform voor de  
InformatieSamenleving

**Verslag workshop**  
**Aanpak Begeleidingsethiek**  
**RVO - Casus**  
**GenAI voor content**

Meer informatie over de Aanpak Begeleidingsethiek,  
waaronder dit verslag, vindt u op  
[www.begeleidingsethiek.nl](http://www.begeleidingsethiek.nl)