





Rijksdienst voor Ondernemend
Nederland



Dienst ICT Uitvoering
Ministerie van Economische Zaken

ECP

Platform voor de
InformatieSamenleving

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: Thijs van der Wal (RVO), Moshin El Hachioui (RVO)

Moderatoren: Edwin Borst (ECP), Marleen van Leengoed (zelfstandig adviseur)

Op woensdag 28 augustus organiseerde RVO op initiatief van Thijs van der Wal (RVO), een sessie begeleidingsethiek over de casus Microsoft Copilot in Office 365 (Copilot). Aan de workshop namen 14 deelnemers (zie bijlage voor specificatie) deel onder leiding van Edwin Borst (ECP) en Marleen van Leengoed (zelfstandig adviseur).

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

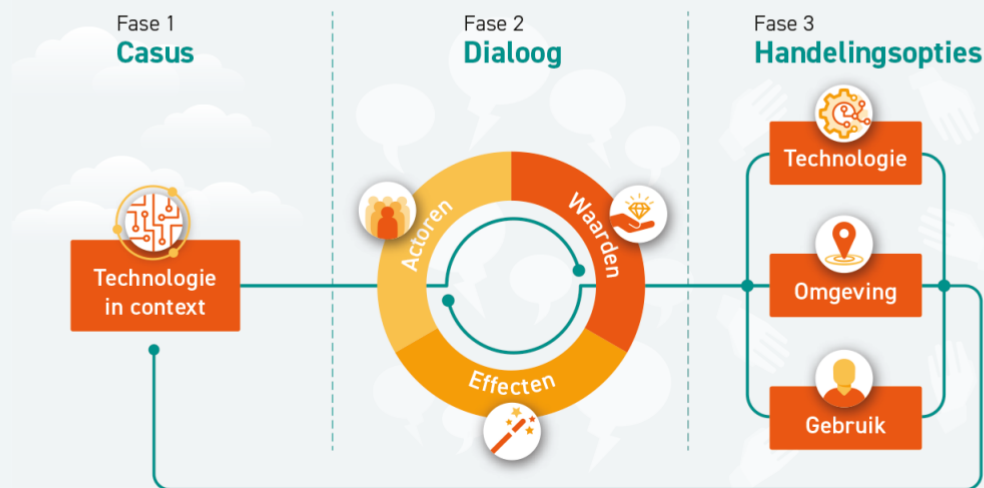
Het verslag bevat de volgende elementen:

- ▷ 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

Copilot in Office 365 is een generatieve AI-assistent van Microsoft, ontwikkeld om de productiviteit te verhogen in toepassingen zoals Word, Excel, PowerPoint, Outlook en Teams. Het is geïntegreerd in de Microsoft 365 suite en helpt gebruikers bij hun dagelijkse taken door middel van kunstmatige intelligentie (AI) en machine learning. Naast een eenvoudige ondersteuning, zoals spellingscontrole of tekstsuggesties, gaat het een stap verder door proactief oplossingen te bieden.

In Word kan Copilot bijvoorbeeld teksten herschrijven, samenvattingen maken of suggesties geven om de toon van een document te verbeteren. In Excel helpt het bij het analyseren van gegevens en het destilleren van inzichten uit de data. In PowerPoint helpt het met het automatisch ontwerpen van dia's en het maken van presentaties op basis van ingevoerde gegevens of teksten. In Outlook kan het e-mails samenvatten, antwoorden voorstellen. In Teams kan het gespreksverslagen maken en automatisch notulen en actiepunten terugkoppelen.

Het doel van Copilot is om tijd te besparen en de nauwkeurigheid van het werk te vergroten door routinetaken te automatiseren.

Bij de RVO is tussen 31 mei en 21 augustus een experiment uitgevoerd met Copilot in Office 365 (Copilot), waarbij het doel was om te leren hoe Copilot het dagelijks werk bij de RVO kan ondersteunen en verbeteren. In het experiment is de software 3 maanden gebruikt door 50 participanten. Daarbij is twee keer een enquête afgenomen waarbij een inventarisatie is gemaakt van de effecten op onder meer productiviteit en bij welke taken Copilot heeft geholpen. Ook is onderzocht of men alternatieve gratis generatieve AI-toepassingen (chatdiensten zoals ChatGPT) minder is gaan gebruiken. Daarnaast is ook het effect op de waardering van RVO als werkgever onderzocht.

Een aantal deelnemers van het experiment was bij deze sessie begeleidingsethiek aanwezig.



Fase 2

Dialog

In deze tweede fase gaan de deelnemers in gesprek over wie er betrokken zijn bij de inzet en het gebruik van Copilot. 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.

Actoren

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

- Medewerkers van RVO
 - Gebruikers van Copilot
 - Niet-gebruikers van Copilot, incl. mensen die het niet willen gebruiken
- Medewerkers overheid automatisering/techniek – technisch ondersteuners van Copilot
 - Beheerders
 - Helpdesk
- Klanten RVO
- Opdrachtgevers RVO
- Leverancier van Copilot: DICTU, Microsoft en SoftwareOne
- Data beveiliging
- IV Architecten
- Werkplekeigenaar
- Toezichthouders
- Politiek
- Publiek/Media
- Juridische partijen
- CIA - Amerikaanse overheid
- NGO's
 - Privacy
 - Ethiek
 - Milieu
- Data-'labellers'
- Data-eigenaren
 - Intern
 - Extern auteurs, content creators
- OR
- Planeet/milieu

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 Copilot hebben.

Positieve effecten

- Mate van compliance is beter inzichtelijk, betere beveiliging
- Sneller en leuker werken
- Hogere kwaliteit
- Minder mensen nodig
- Toegang tot meer informatie, vollediger
- Niet-menselijke verrassing
- Beter op de hoogte
- Interessante werkgever voor nieuwe generatie
- Stimulans voor informatiehuishouding
- Verbeterd leerproces/ doorgeven informatie
- Vergroot vermogen door andere voorbereiding

Negatieve effecten

- Niet meer naar collega voor overleg?
- Wordt moeilijker om compliant te zijn
- Minder mensen nodig
- Transparantie en uitlegbaarheid wordt minder
- Geen inzicht in totstandkoming van advies
- Geen balans tussen dom en slim werk
- AI kan hallucineren, bronherleiding is lastig, er is controle nodig
- Neiging om voor waar aan te nemen
- Minder persoonlijk
- Ecologische footprint- Natuur, energie, water, kans op 'Whitewashing ethics'
- Er ontstaat een kloof tussen vaardig en niet-vaardig
- Investering in training nodig
- Wat doe je met efficiency/ winst?
- Verslavend
- Informatiehuishouding moet op orde, veel informatie is niet meer relevant
- Wat AI produceert wordt ook weer gebruikt om het te trainen, zichzelf versterkend
- Afkeuring door medewerkers
- Angst voor 'big-tech' en privacy
- Nuances raken verloren
- Transcriptie laagste vorm van vastleggen
- Uniformiteit (minder creativiteit)
- Bias (polarisatie, versterkend effect)



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:

Waarden	
Vakmatige kwaliteit	Uniformiteit
Betrouwbaarheid	Waarborgen van de continuïteit van de organisatie
Rechtvaardigheid	Privacy/ veiligheid
Verantwoordelijkheid	Onafhankelijkheid
Compliant zijn	Autonomie/ vrijheid
Effectiviteit	Creativiteit
Werkplezier	Menselijkheid
Milieuwelzijn/ duurzaamheid	Uitlegbaarheid/ Transparantie



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 één van deze drie gespreksonderwerpen. Dat leidt tot handelingsopties per categorie.

Technologie

- In generieke zin, Copilot op verschillende manieren kunnen aanpassen als organisatie
- Meer guard-rails zodat er geen hallucinatie optreedt, en geen toxiciteit. Bijv. Schadelijke informatie = geen output
- Idealiter model niet getraind met ongewenste biased data
 - Ethische ladder: niet alle modellen zijn gelijk, keuze voor alternatief model
- Geen datalekken, contractueel vastleggen consequenties
- Fine-tuning van model op behoefte Rijksoverheid:
 - Voor het goed kunnen transcriberen (openbaar bestuur)
 - Nederlandse taal beter ondersteund
 - Ook dialecten ondersteunen (spraakinput)
 - Opt-out per individu per usecase (bijv. in geval van transcriptie)
 - Anoniem worden, maar toch transparant zijn als overheid
 - RVO versie van Copilot of ander LLM – op basis van RVO informatie getraind
- Bronvermeldingen toevoegen
- Nadenken (vooraf!) over exit-strategie
- Controle over locatie server- dataopslag
- Beslissing maken over dataopslag: alles op one-drive?
- Meer transparantie vragen/eisen van Microsoft
- Waarschuwing indien gebruikt voor dingen die ook op eenvoudigere minder milieuvriendelijke manier kunnen: ‘Let op, dit kost een boom’
- Meer invloed kunnen uitoefenen op balans gebruiksvriendelijkheid vs RVO richtlijnen
- Waarschuwing wanneer je het onjuist gebruikt?
- Feedbackloop naar ontvangende partij, zichtbaar wanneer deels met AI gegenereerd. En dat ze verzoek in kunnen dienen voor menselijke check



Omgeving

- Educatie over gebruik Copilot en AI in algemeen geven:
 - Richtlijnen hoe je het moet gebruiken
 - Hoe kan je goede vragen leren stellen (prompting) om te weten of de uitkomsten kloppen?
 - Welke skills heb je nodig om te beoordelen of uitkomst betrouwbaar is en van voldoende kwaliteit?
 - Milieubewust gebruik stimuleren. Was het milieuschade waard? Bijvoorbeeld: wanneer gebruik je je zoekfunctie in de mail en wanneer stel je een vraag aan Copilot?
- Moet heel duidelijk zijn wanneer je in Copilot omgeving werkt
- Coaches in teams die goed weten wat wel/niet kan
- FAQ op muismatten met do's en don'ts
- Goed met elkaar in gesprek blijven hoe je het gebruikt
- 4 ogen principe hanteren
- Nudges geven aan de mens om kritisch te blijven op output: "Check inhoud"
- Uitkomsten AI gecheckt door mens of uitkomsten mens gecheckt door AI?
- Disclaimer bij gegenereerde content: "Created by AI"
- Informatiehuishouding moet op orde zijn: zoals beter versiebeheer en er zou een bepaalde drempelwaarde voor kwaliteit moeten zijn
- Duidelijkheid over wat er wordt verwacht van medewerkers:
 - Nieuwe medewerkers
 - Oude medewerkers moeten keuze houden om het niet te gebruiken
 - Waarborgen dat je wordt beoordeeld op eigen kwaliteiten en niet eigen kwaliteiten +AI
 - Productiviteit denken vs. Toegevoegde waarde: is de effectiviteitswinst persoonlijke winst of is deze voor het bedrijf?
 - Goede voorbeeld geven als manager
- Uitleg naar organisatie waarom de keuze voor Copilot, en uitleg hoe veiligheid gewaarborgd is



Mens

- Algemeen bewustwording van medewerkers vergroten, zodat deze Copilot geïnformeerd kunnen gebruiken
- Stap 1: Gebruikersgroepen formeren en daarmee gedifferentieerde kaders, instructies en richtlijnen voor deze gebruiksgroepen definiëren met gebruik van de waarden
- Stap 2: Klantgroepen formeren (idem als 1)
- Stap 3: Groepen voorzien van deze instructies, kaders en richtlijnen
 - ‘Medewerkers die medewerkers begeleiden’ en maatwerk kunnen leveren. E.g. Opleiders, mentoren, superusers
- Stap 4: Groepen verplichten en verleiden tot (goed) gebruik
 - Ambassadeurs, super users, bekeerlingen, visie makers, DT/MT
- Stap 5: Feedbackloop over toepassing en effecten
 - Medewerkers die feedback geven: met collega’s overleggen bij vragen/twijfel, naar management, naar ontwerpers, naar beleids- en visiemakers, met DERK en EC.



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.

In de sessie worden geen besluiten genomen over het al dan wel of niet operationaliseren van Copilot. De opgedane kennis is weergegeven in dit verslag, die 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)
- Ethische Commissie (RVO)
- Data en Ethiek Raadkamer (RVO)
- Kamer Versterken ethisch denken en handelen (RVO)
- Dienst ICT Uitvoering (DICTU)
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK)

1. **Thijs van der Wal** RVO - Strategisch CIO-Adviseur - Digitale Transformatie & Innovatie

2. **Mohsin el Hachioui** RVO - Projectleider Innovatie- Digitaal Innovatie Lab

3. **Marcello Balistieri** RVO

4. **Wim Brama** RVO

5. **Shielta Dielbandhoesing** BZK

6. **Ronald Heukers** DICTU

7. **Jorinde Kruisdijk** RVO

8. **Martijn van Lingen** RVO

9. **Marion Mulder** RVO

10. **Olaf Pijl** RVO

11. **Simone Stubenitsky** RVO

12. **Anita Vocht** RVO

13. **Anna Vogel** RVO

14. **John Wilcke** DICTU

Moderatoren

Edwin Borst (ECP), Marleen van Leengoed (zelfstandig adviseur)



aanpak 
begeleidings
ethiek

ECP
Platform voor de
InformatieSamenleving

Verslag workshop
Aanpak Begeleidingsethiek
RVO - Casus

Microsoft Copilot in Office 365

Meer informatie over de Aanpak Begeleidingsethiek,
waaronder dit verslag, vindt u op

www.begeleidingsethiek.nl