Aanpak begeleidings ethiek

Door Peter-Paul Verbeek
en Daniël Tijink

Een dialoog over technologie met handelingsperspectief

ECP
Platform voor de
InformatieSamenleving

THINK BEFORE YOU PRINT

Colofon

© 2019 ECP | Platform voor de InformatieSamenleving

Auteurs
Daniël Tijink en
Peter-Paul Verbeek

Ontwerp
Brand new something

Fotografie
Unsplash, Shutterstock

Druk
Veldhuizen grafisch effect

Met dank aan
Turnaround Communicatie

Inhoud

Introductie 5
Aanpak begeleidingsethiek

Hoofdstuk 1 9
Begeleidingsethiek

Hoofdstuk 2 13
Vier cases

Case 1 AI in de psychiatrische zorg 14
Case 2 Facebook: algoritmische tijdslijn/newsfeed 15
Case 3 Eetrobot 16
Case 4 Ondermijning 18

Hoofdstuk 3 21
Uitgangspunten voor concretisering begeleidingsethiek

Hoofdstuk 4 31
Toelichting aanpak begeleidingsethiek

Fase 1 Casus: technologie in context 33
Fase 2 Dialoog: actoren, effecten, waarden 33
Fase 3 Handelingsopties 38

Hoofdstuk 5 47
Gebruik van de aanpak begeleidingsethiek

Hoofdstuk 6 55
Vervolg

Bijlagen 59

Introductie

Aanpak begeleidingsethiek

We leven in een samenleving waar digitalisering zorgt voor grote veranderingen. De smartphone in ieders hand, algoritmes die slimmere beslissingen suggereren of soms nemen, robots die mensen ondersteunen of een deel van hun werk vervangen. Veel van die ontwikkelingen roepen ethische vragen op. Zijn het goede ontwikkelingen? Passen ze bij onze waarden? Kunnen we ze een verantwoorde richting geven?

Deze vragen hebben een hoge urgentie. Met de ontwikkeling van digitale technologie zijn niet alleen grote maatschappelijke maar ook economische belangen gemoeid. Bedrijven willen mee met de ontwikkelingen, Europa mag niet verder achterblijven bij China en Amerika. Mensen vragen om nieuwe oplossingen, maar er leven ook zorgen over de impact van de snelle digitalisering. Hoe kunnen we ethiek en innovatie vruchtbaar met elkaar verbinden?

Rondom deze vraag is een samenwerkingsverband ontstaan tussen ECP|Platform voor de InformatieSamenleving en techniekfilosoof Peter-Paul Verbeek. Professor Verbeek is een toonaangevend denker in de (inter)nationale discussie over ethiek en technologie. Hij heeft de kennis vanuit de techniekfilosofie in de ethische discussie gebracht. De basis van zijn gedachtegoed, ondersteund door veel onderzoek, is dat het beter is mens en technologie te beschouwen als intrinsiek verbonden in plaats van ze tegenover elkaar te stellen. Dit heeft vergaande consequenties voor de manier waarop ethisch denken wordt vormgegeven. Dit alternatief voor de klassieke ethiek noemt hij begeleidingsethiek. De technologie wordt in de maatschappij begeleid en omgekeerd wordt de maatschappij begeleid in de omgang met technologie. In hoofdstuk 1 wordt dit gedachtegoed uitgewerkt.

ECP nam het initiatief tot samenwerking met professor Verbeek met als doel begeleidingsethiek toepasbaar te maken in de pratijk. Daartoe werd een werkgroep samengesteld met partijen vanuit het (ICT)-bedrijfsleven (KPN, IBM, Facebook, Microsoft), vanuit de overheid (de Ministeries van BZK, EZK, J&V en de VNG) en vanuit verschillende sectoren: Alliander (energie), de Nationale Politie (veiligheid), Siza/Academy Het Dorp en Rijnstate (zorg). Ook was er een deelnemer van de Dienst Koninklijk Huis. Vanuit de ethiek namen onderzoekers deel van de RU Leiden, de TU Delft, Tilburg University en de Erasmus Universiteit.

In deze publicatie wordt de ontwikkelde aanpak toegelicht en uitgewerkt. Na een inleidend hoofdstuk over de basisgedachten van de benadering (Hoofdstuk 1) worden vier cases gepresenteerd, die door de rest van het document als voorbeeld dienen (Hoofdstuk 2). Vervolgens werken we uit wat de uitgangspunten zijn voor de concretisering van begeleidingsethiek (Hoofdstuk 3) om daarna de 'toelichting aanpak begeleidingsethiek' te presenteren (Hoofdstuk 4). Aansluitend onderzoeken we wanneer en hoe de aanpak gebruikt kan worden (Hoofdstuk 5). Het laatste hoofdstuk bevat een korte terugblik en een blik op de toekomst (Hoofdstuk 6).

Veel technologische ontwikkelingen roepen ethische vragen op

Hoofdstuk 1
Begeleidingsethiek

Ethische discussies rondom technologie hebben vaak het karakter van een dilemma: is het wel of niet aanvaardbaar om deze technologie toe te passen? In de publieke discussie leidt ethische zorg over technologie dan ook vooral tot radicale kritiek: 'de macht van Big Tech', 'de spionerende overheid', 'het einde van arbeid door AI', et cetera. Hierdoor ontbreekt er een kader om ethische vragen en zorgen constructief in te zetten en om technologische ontwikkelingen een wenselijke richting te geven in plaats van ze alleen volledig te omarmen of te veroordelen.

De invulling van techniek-ethiek als het aanvaarden of afwijzen van technologie, plaatst techniek en samenleving tegenover elkaar. Technologie vormt in die benadering een potentiële bedreiging voor de samenleving en het is de verantwoordelijkheid van de ethiek om te bepalen welke techniek toegelaten mag worden en welke niet. Maar dit beeld is onjuist. Techniek en samenleving zijn juist fundamenteel met elkaar verweven. Technologie wordt door mensen ontwikkeld met het oog op een bepaalde rol van die technologie in de samenleving. En de samenleving heeft altijd al vorm gekregen door interactie met technologie, soms ook op manieren die niet expliciet bedoeld waren door ontwerpers. De boekdrukkunst leidde bijvoorbeeld niet alleen tot de mogelijkheid om gemakkelijker teksten te vermenigvuldigen maar droeg ook bij aan de reformatie, de opkomst van de moderne wetenschap en de universiteit, het belang van kennis et cetera. We zijn net zo verbonden met technologie als met de taal of de zwaartekracht: technologie maakt ons mede tot de mensen die we zijn. Techniek en samenleving geven vorm aan elkaar: dat is de les die we uit de afgelopen 50 jaar wetenschaps- en techniekonderzoek kunnen leren.

Deze verwevenheid van techniek en samenleving brengt een andere rol voor de ethiek met zich mee. Het standaardmodel van 'ethische beoordeling', waarin op basis van een normatieve theorie wordt bepaald of een technologie aanvaardbaar is of niet, doet onvoldoende recht aan deze verwevenheid. De vraag is in de meeste gevallen niet of een technologie toegelaten moet worden of niet, maar hoe we er op een verantwoorde manier mee om kunnen gaan. Een gevolg van de verwevenheid van mens en techniek is bovendien dat ook de ethische kaders waarmee we technologie beoordelen zich ontwikkelen in interactie met die technologie. Wat we verstaan onder 'privacy', bijvoorbeeld, ontwikkelt zich hand in hand met de technologieën die de grens tussen privé en openbaar doen verschuiven.

In plaats van ethiek te zien als 'beoordelen' zou ze ook gezien kunnen worden als het normatief 'begeleiden' van technologie in de samenleving. Tegelijkertijd kan ethiek de samenleving ook begeleiden in de omgang met de technologie. Zo'n benadering plaatst de ethiek niet buiten de techniek, maar er middenin. De ethiek is dan niet primair gericht op de vraag of een technologie aanvaardbaar is of niet, maar op de vraag of en onder welke condities een technologie een verantwoorde plaats in de samenleving kan krijgen. 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.

In plaats van ethiek te zien als 'beoordelen' zou ze ook gezien kunnen worden als het normatief 'begeleiden' van technologie in de samenleving

Centraal in die begeleidingsethiek staat het inventariseren van de mogelijke maatschappelijke implicaties van een technologie, en de centrale waarden die daarbij op het spel staan. Dat gebeurt in een deliberatief proces. Bij begeleidingsethiek is het belangrijk altijd uit te gaan van concrete technologieën en hun specifieke uitwerking en effecten. Het gaat niet om een generieke analyse van digitalisering of artificiële intelligentie (AI) maar om de concrete toepassing daarvan in een maatschappelijk domein. Het gaat in deze ethische benadering immers altijd om de concrete wisselwerking tussen mens, technologie en maatschappij.

Bij het inventariseren van effecten is het belangrijk om niet alleen te kijken naar de gevolgen voor de individuele gebruiker, maar ook naar de maatschappelijke gevolgen voor onder meer het onderwijs, de zorg, de rechtspraak, wetgeving, opsporing en handhaving. Tenslotte hebben technologieën ook invloed op interpretatiekaders (bijvoorbeeld de betekenis van centrale waarden als privacy en autonomie en veranderende maatschappelijke structuren).

Hoofdstuk 2
Vier cases

Het perspectief van verbondenheid tussen techniekontwikkeling en mens en maatschappij is de basis voor de aanpak begeleidingsethiek die in dit document wordt uitgewerkt. Het is goed die aanpak te illustreren aan de hand van cases. Dat past ook bij het kernidee van begeleidingsethiek, dat zich niet richt op abstracte technologieën, maar op concrete toepassingen die functioneren binnen een bestaande context.

Case 1 **AI in de psychiatrische zorg** [1]

UMC Utrecht onderzoekt een AI-toepassing die het risico op agressie door patiënten op de afdeling psychiatrie kan inschatten. De toepassing wordt nog niet gebruikt, maar is wel succesvol getest. De afdeling is klaar om het in de praktijk te brengen. Voordeel van goede inschatting is dat de professional samen met de patiënt (en eventueel familie) kan kiezen uit een aantal maatregelen om in te zetten, bijvoorbeeld medicatie, meer sporten of familie die kan blijven slapen.

Agressie, zowel in verbale als in fysieke vorm komt veel voor; in deze instelling zijn er honderden meldingen per jaar. Het heeft impact op de patiënt zelf maar ook op de naasten en zorgprofessionals. Daarom hebben professionals, patiënten en naasten ervoor gekozen om te zien of analyse van data hier kan helpen.

De AI-toepassing krijgt als input geanonimiseerde klinische teksten (intakeverslag, incidentrapporten en andere verslagen). Deze worden in wiskundige representatie omgezet, waarna machine learning kan worden toegepast en dit geeft als uitkomst de kans dat er (bijvoorbeeld) binnen dertig dagen agressie zal optreden.

Het inschatten van risico op agressie gebeurt nu door artsen/verpleegkundigen. Het blijkt dat de AI-toepassing iets beter voorspelt dan de vragenlijsten die hiervoor nu speciaal worden ingevuld en beduidend beter dan de subjectieve beoordeling door zorgprofessionals. De tool is na goede resultaten in het UMC Utrecht gevalideerd in een grote GGZ-instelling met vergelijkbare goede resultaten.

1 De eerste case is besproken tijdens een workshop georganiseerd door ECP en NICTIZ, de drie anderen zijn in de werkgroep ethiek en digitalisering ingebracht. We danken respectievelijk Karin Haagoort (UMCU), Edo Haveman (Facebook), Brigitte Boon (Siza, Academy Het Dorp) en Marc Noordhoek (BZK) voor de inbreng van hun case en de toelichting erop. De verantwoordelijkheid voor de tekst over de cases ligt bij de schrijvers van dit document.

Case 2 **Facebook: algoritmische tijdslijn/newsfeed**

Iedereen kent Facebook. Ooit opgericht als smoelenboek is het nu het grootste sociale netwerk ter wereld. De technologische toepassing die in deze de case besproken wordt is newsfeed. Activiteiten die gebruikers op hun newsfeed laten zien (door vrienden te maken en pagina's te volgen) gaan in een algoritme dat ze in een volgorde zet. Engagement is daarin een belangrijke factor. Door het algoritme komen zaken bovenaan te staan, die mensen echt niet willen missen, zoals trouwfoto's en babyfoto's.

Deze technologie maakt het volgende mogelijk: user generated content, vrijheid van informatie (iedereen kan bijna alles posten), democratisering en aanspreekbaarheid van politici, snelle vragen en antwoorden, mobilisatie (women's march) en aan de kaak stellen van nieuwe zaken (#metoo).

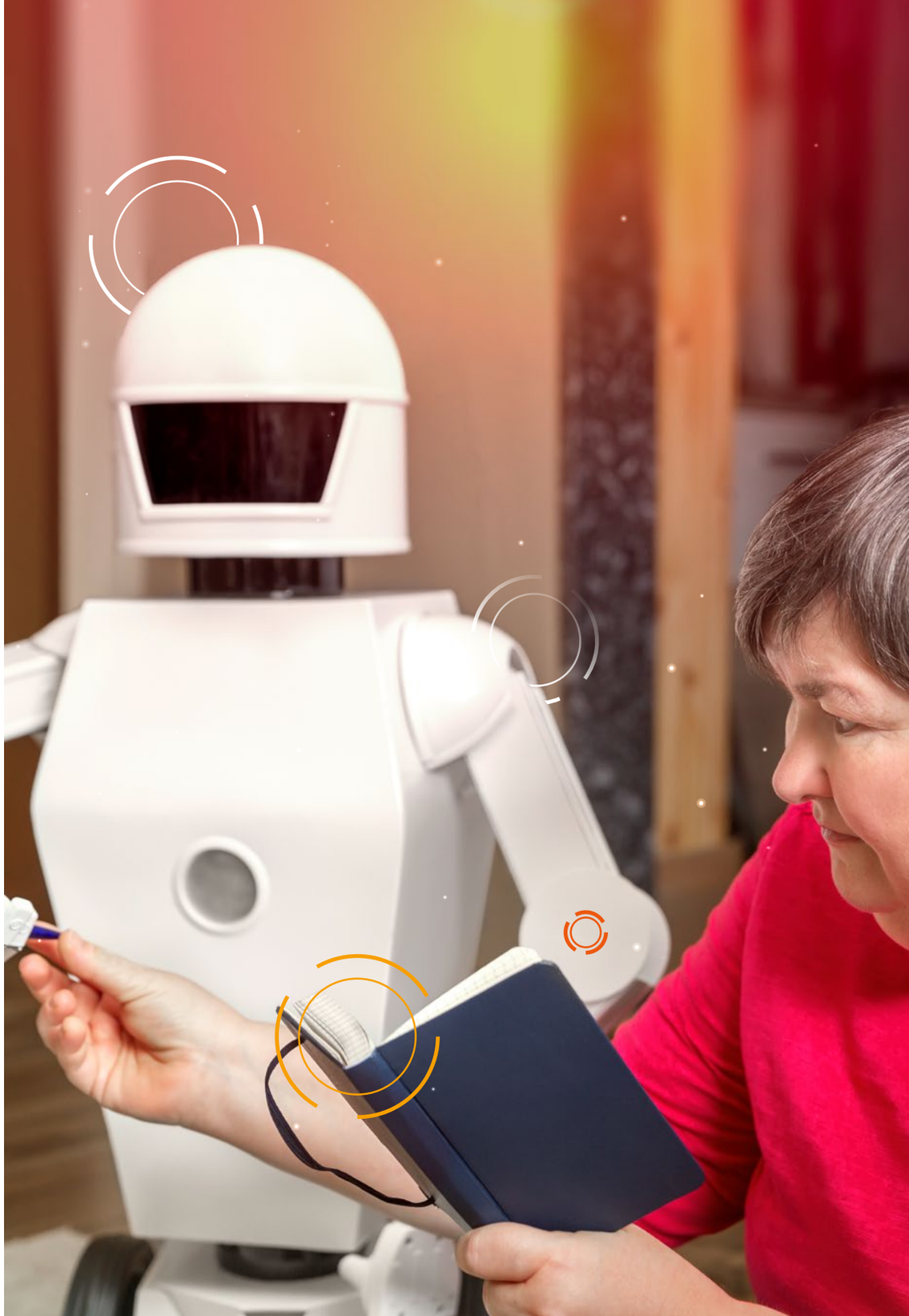
Door de ontwikkeling van het internet en platformen verschuift een deel van de poortwachtersrol van bijvoorbeeld kranten of tv-nieuwsprogramma's naar algoritmes en de gebruiker. Zij bepalen deels de relevantie en betrouwbaarheid van informatie. Door mensen te geven wat ze wensen, neemt de kans toe dat intellectuele content lager komt of sensatieberichten hoger. Als het gaat om misinformatie en polariserende content is dit onwenselijk.

Case 3 **Eetrobot**

Academy Het Dorp ondersteunt instellingen in de Langdurige Zorg met (onderzoek naar) de inzet van nieuwe technologieën. Ook onderzoek naar de ethische waarden die een rol spelen bij de inzet van technologie maakt daar deel van uit. Eén van de zorginstellingen, en oprichter van Academy Het Dorp, is Siza. Siza is een gehandicaptenzorginstelling die zorg biedt aan mensen met verschillende typen beperkingen, waaronder mensen met lichamelijke beperkingen, niet-aangeboren hersenletsel, verstandelijke beperkingen en ernstige meervoudige beperkingen. Siza wil haar cliënten zo veel mogelijk zelfstandigheid bieden, en zet daarbij nieuwe technologieën in.

Cliënten met lichamelijke beperkingen en niet-aangeboren hersenletsel wonen in Het Dorp, één van de locaties van Siza, in een eigen huis dat zoveel mogelijk op hun specifieke behoeften is aangepast. Sommige cliënten kunnen niet zelf eten en hebben daar hulp bij nodig. Men maakt naast de ondersteuning door zorgmedewerkers gebruik van verschillende soorten technologie. Te denken valt aan een eetrobot of een robotarm (die gemonteerd is op een rolstoel). Met de robot(arm) kan het eten of drinken naar de mond van de cliënt worden gebracht. De zorgmedewerker hoeft nu alleen het eten klaar te zetten en na afloop op te ruimen, en de cliënt kan zijn maaltijd gebruiken zoals hij het wil: op eigen tempo, met privacy – of, als hij daarvoor kiest met een zorgmedewerker die hem alsnog helpt met eten.

Algoritmes en robots worden steeds meer in 'normale' werkprocessen gebruikt

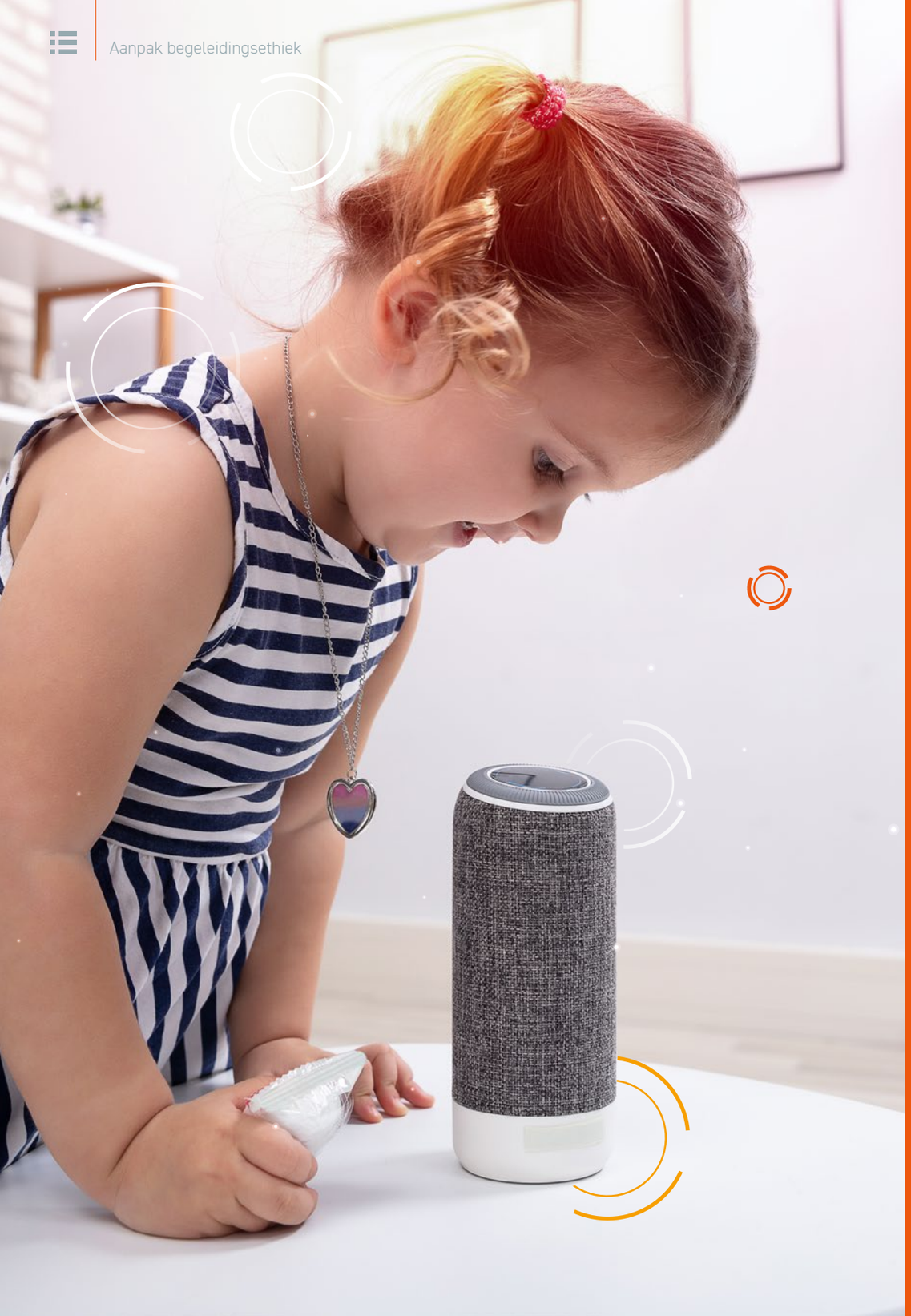


Case 4 **Ondermijning**

Ondermijnende criminaliteit is een verzamelterm voor delicten die overheids- en maatschappelijke structuren evenals het vertrouwen daarin schaden. Een belangrijk element van dergelijke criminaliteit is het misbruik van legale/bovengrondse personen, organisaties en instituties.

In het project 'ondermijning' wil men met behulp van (big) data uit verschillende bronnen meer zicht krijgen op waar ondermijning zich concentreert. Een groeiend aantal grotere gemeenten werkt hieraan mee samen met het OM en de Ministeries van Binnenlandse Zaken, Justitie & Veiligheid en Financiën. Bronnen die gebruikt worden zijn bijvoorbeeld het Kadaster en het Centraal Bureau voor de Statistiek. De voornemens en randvoorwaarden zijn vastgelegd in de Staatscourant onder de naam: CITY DEAL zicht op ondermijning. Daarmee is een heldere afbakening en een goede juridische grondslag gewaarborgd. Zo zijn de analyses wetenschappelijk controleerbaar, voldoet het proces aan de AVG en worden de algoritmes op validiteit getest. Er is geen sprake van een black box.

Er is ingezoomd op drie thema's: drugs, malafide stichtingen en vastgoedfraude. Een voorbeeld van dat laatste zijn de zogenaamde windvangers. Mensen zonder inkomen en zonder vermogen die toch vastgoed bezitten. De analyse brengt tot op wijkniveau in kaart hoe groot de concentratie is.



Hoofdstuk 3

Uitgangspunten voor concretisering begeleidingsethiek

Het doel van deze publicatie is om vanuit het gedachtegoed van begeleidingsethiek te komen tot een methode om dat gedachtegoed in de praktijk te kunnen toepassen. Daarvoor zetten we eerst een tussenstap. We benoemen de uitgangspunten die horen bij deze manier van kijken naar ontwikkeling van technologie en van maatschappij. Bij de bespreking van die uitgangspunten gebruiken we ter illustratie elementen uit de eerdere cases.

De hoe-vraag centraal

Mens en technologie zijn verbonden en beïnvloeden elkaar continu, als twee partners in een dans. De kern van de begeleidingsethiek is daarom niet de menselijke beoordeling van de technologische ontwikkeling, in tegenstelling tot vele andere ethische benaderingen. De kern is het samenspel, de 'hoe'-vraag, in plaats van de 'ja of nee'-vraag: hoe kunnen mens en technologie zich op een waarden-volle wijze ontwikkelen?

Als we naar de case over de eetrobot kijken, gaat het dus niet in eerste instantie om de vraag of een eetrobot wel ingezet mag worden, maar hoe we dat het beste kunnen doen. Kunnen we een dialoog voeren over hoe we met de eetrobot omgaan in plaats van discussie wat goed en fout is, kunnen we handelingsopties vinden voor goed gebruik? In plaats van de vraag 'mogen we de zorg voor kwetsbare mensen wel aan machines delegeren?' komt dan de vraag: 'is er een manier om de waarden die centraal staan in de zorg als uitgangspunt te nemen bij het ontwikkelen en gebruiken van zorgrobots?'.

De focus op de hoe-vraag wil niet zeggen dat er nooit nee gezegd kan worden en dat elke technologie in principe overal kan worden gebruikt of ingevoerd, maar de focus ligt in eerste instantie niet op het 'ja of nee'- oordeel. Die vraag is snel beperkend, waardoor handelingsopties onbesproken blijven. Vaak worden ook waarden of belangen van bepaalde groepen onvoldoende meegenomen. Door de aandacht te richten op de hoe-vraag wordt het mogelijk om ethiek daadwerkelijk te verbinden met techniek. Deze vraag maakt ruimte om naar condities te zoeken waaronder een technologie op een verantwoorde manier kan functioneren. En die condities bevinden zich in het ontwerp van de technologie zelf, in de maatschappelijke inbedding ervan en in de manier waarop mensen er gebruik van maken. Maar uiteindelijk blijft ook 'niet' een antwoord op de 'hoe'-vraag. Als blijkt dat de technologie niet verenigbaar is met onze waarden, geeft het model als uitkomst dat de technologie beter niet gebruikt kan worden.

Kleine, doorgaande stappen

Nauw verbonden met het stellen van de hoe-vraag is dat verbetering komt in (kleine) stappen. Als je zoekt naar een 'ja of nee', 'mag dit of mag dit niet', zoek je naar een ultiem antwoord. Vanuit het paradigma van begeleidingsethiek zie je dat de met elkaar verbonden technologie en mens zich ontwikkelen in kleine, steeds doorgaande stappen. De technologie past zich steeds aan op basis van wat mensen aangeven, maar ook de gebruiker van technologie past zich aan de nieuwe mogelijkheden aan.

Bij de eetrobot zijn er in de ontwikkeling allerlei kleine stappen te zien. De robot wordt steeds preciezer, mede op basis van de input van de gebruikers. Maar ook de rol van de eetrobot in het zorgproces en wat er van de zorgprofessionals wordt verwacht, verandert.

Sommige ethische stromingen hebben het idee dat de ontwikkeling van mens en technologie kan worden stopgezet en dat de ethiek dat hoort te bepalen. Vanuit de begeleidingsethiek denken we dat die situatie zelden voorkomt. Wat wel mogelijk is, is een continue overweging wat er kan worden verbeterd en daarin concrete stappen zetten.

Technologie en samenleving bevinden zich in een proces van continu wederzijdse aanpassing

Technologie in context

Omdat mens en technologie zo nauw met elkaar verbonden zijn, heeft het weinig zin te spreken over de technologie zonder daarbij de mens en diens context te betrekken. Het maakt namelijk uit waar en door wie de dialoog wordt gevoerd.

Vanuit de begeleidingsethiek willen we het gesprek aangaan over concrete technologieën die functioneren binnen een concrete context, met echte mensen. We praten daarom liever niet over dé technologie of dé mens. Discussies over AI of blockchain kunnen interessant zijn, maar worden pas relevant als ze de praktijk raken.

Dit stelt eisen aan het abstractieniveau dat we kiezen om over een technologie te praten. De casus over de eetrobot is een mooi voorbeeld. We spreken niet over robotica (te algemeen) of over een merk robot van een bepaald serienummer (te specifiek), maar over 'eetrobots'. En wel binnen een bepaalde context, in dit geval een instelling voor gehandicaptenzorg. Geleerde lessen zullen (groten)deels ook gelden voor het gebruik van andere eetrobots in andere instellingen. Maar mogelijk ook breder voor andere e-health in deze context of voor bredere robotica of juist beperkter; toch heel specifiek voor deze instelling en deze eetrobot.

Zonder context kunnen we eigenlijk niet oordelen. De patronen die bij ondermijning worden opgemerkt, krijgen pas betekenis als agenten en beleidsmedewerkers er iets mee kunnen, als analisten er een interpretatie aan geven en als gesprekken met bewoners en andere betrokkenen ze duiden. De ethische spanning zit net zo goed in het netwerk om de technologie heen als in de data-analyse zelf.

Menselijke waarden

Het doel van begeleidingsethiek is dat de ontwikkeling van de technologie en de omgang ermee steeds beter voldoet aan de menselijke waarden, zoals rechtvaardigheid, autonomie, snelheid, duurzaamheid, veiligheid, effectiviteit, etc.

Emoties, zowel positieve als negatieve, kunnen een belangrijke rol spelen in het zoeken naar de waarden die centraal staan rondom een bepaalde technologie. Angst, enthousiasme, verbazing, bezorgdheid: het zijn allemaal indicaties dat de technologie iets waardevols op het spel zet of mogelijk maakt. En daarmee vormen emoties een indicator voor de normatieve kaders die een plek zouden moeten krijgen in het ontwerpen, implementeren en gebruiken van deze technologie.

Welke waarden relevant zijn, is afhankelijk van de specifieke technologie en context. In de Chinese cultuur heersen andere waarden dan in die van de Verenigde Staten, en in de zorg andere dan in de bouw.

De komst van technologie kan waarden ook nog eens veranderen. Bij de introductie van de mobiele telefonie hadden waarden als 'bereikbaarheid' en 'aandacht' een andere invulling dan bij het huidige gebruik ervan.

Er is bijna altijd spanning tussen de verschillende waarden die spelen bij een bepaalde technologie binnen één context. De newsfeed van Facebook heeft de spanning van de vrijheid om alles te publiceren en kwalijke gevolgen van versterking van uitingen die minder relevant of zelfs onwaar zijn. Bij het project ondermijning zit er een spanning tussen de wens zo precies mogelijke opsporingsinformatie te verkrijgen en privacybescherming van individuen.

Het streven naar een *waarden-volle* ontwikkeling van technologie en mens is kortom complex. Het is belangrijk deze complexiteit te onderkennen en er een omgang mee te vinden. Redeneren vanuit één bepaalde waarde, zonder het besef van die complexiteit, brengt niet zoveel.

Wat zal de toekomst brengen: positief én negatief

Van nieuwe technologie weten we niet wat die ons gaat brengen, of waar wij de technologie gaan brengen. Als we terugkijken naar toekomstvoorspellingen blijken die er vaak naast te zitten. De opkomst van internet is in de jaren vijftig nauwelijks voorspeld, vliegende auto's wel. Toch zijn beelden van de toekomst nodig, ze zijn een belangrijk onderdeel van de wisselwerking van technologieontwikkeling en maatschappelijke verandering.

Veel ethische discussies richten zich op de mogelijke nadelen van een technologie. Vanuit de begeleidingsethiek gaan we ervan uit dat technologie zowel positieve als negatieve gevolgen heeft. Het is belangrijk beide voldoende ruimte te geven in een dialoog. Als we alleen praten over hoe een eetrobot patiënten nooit het menselijke contact kan bieden als een zorgprofessional dat kan, missen we de kans die de robot patiënten biedt om juist iets van hun menselijke waardigheid terug te krijgen door weer alleen te kunnen eten.

Mogelijk is hier een parallel met bewegingen in andere wetenschapsgebieden: positief design (op de nieuwe mogelijkheden ontwerpen), positieve psychologie (focus op iemands floreren en niet op iemands problemen), positieve gezondheid (niet alleen focus op wat er mis is en niet kan, maar ook op wat er wel kan).

Handelingsopties

De begeleidingsethiek-aanpak zoekt naar concrete handelingsopties om te komen tot een waardenvoller samenspel tussen mens, maatschappij en technologie. We onderscheiden drie soorten handelingsopties om die waardenvolle technologie-maatschappelijke ontwikkeling te bereiken.

Ontwerpen van technologie: *ethics by design*

Elke technologie heeft waarden als het ware ingebouwd, technologie nodigt uit tot bepaald gedrag. Een technologie kan dus zo ontworpen worden dat zij beter aansluit bij bepaalde waarden. Zo kan de waarde van privacy worden gewaarborgd door de gebruiker invloed te geven op de cookies die opgeslagen worden, of door fotocamera's op drones voor crowd control een lage resolutie te geven die het wel mogelijk maakt om aantallen mensen te tellen maar niet om individuele mensen te herkennen.

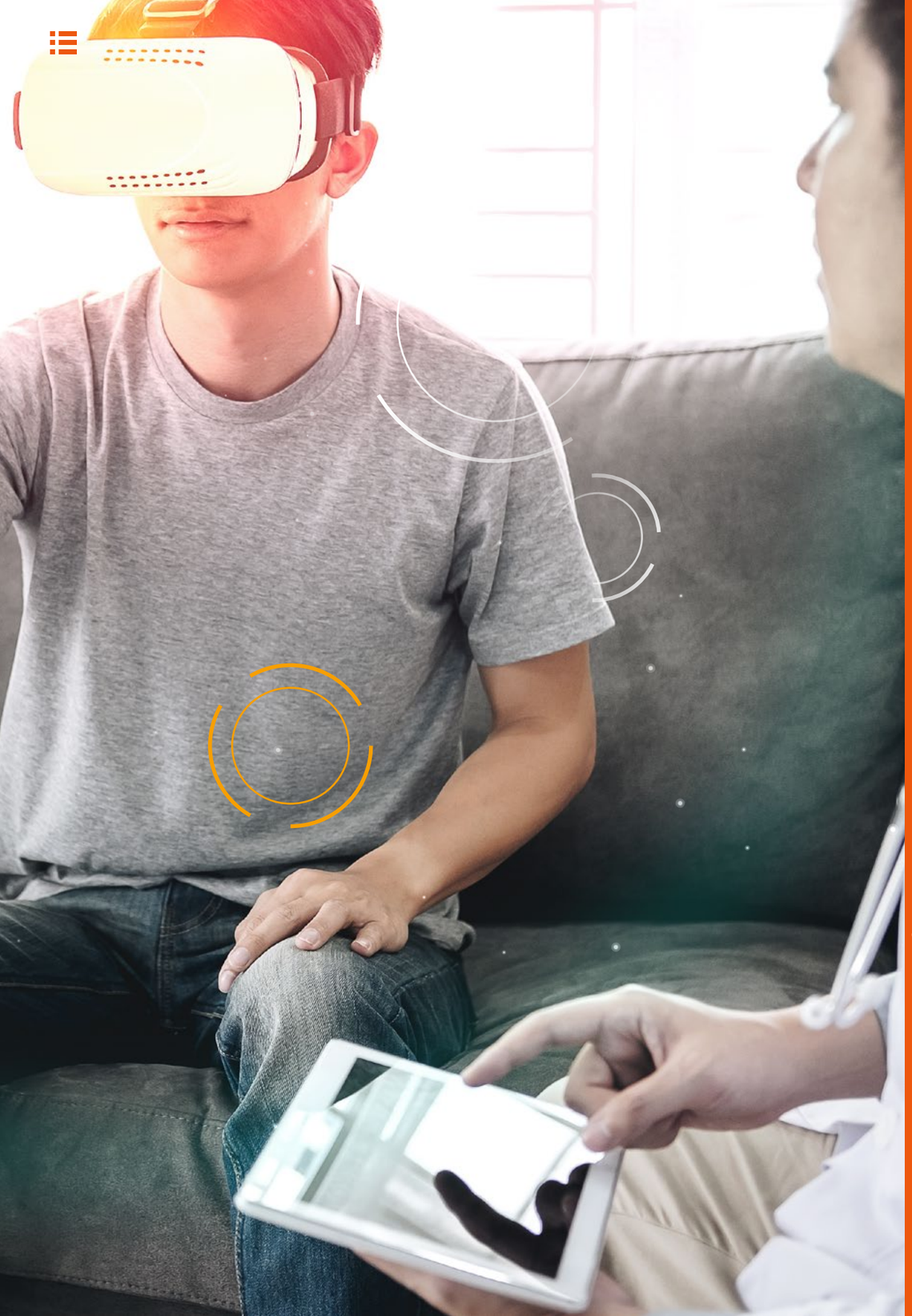
Omgeving: *inrichten (fysiek) en afspraken maken (sociaal); ethics in context*

Elke technologie wordt gebruikt in een context: fysiek, sociaal, beleidsmatig/juridisch. Bij nieuwe technologieën wordt die omgeving ook aangepast. Bij het in gebruik nemen van de auto kwamen er (fysiek) stoepen en stoplichten. Ook sociaal kwamen er nieuwe afspraken: voetgangers op de stoep, auto's op de weg, zebrapad en vluchtheuvel veilige plekken voor voetgangers. Die afspraken zijn ook gejuridificeerd via verkeerswetten.

Gebruiker: *bewust worden en gedrag aanpassen: ethics by user*

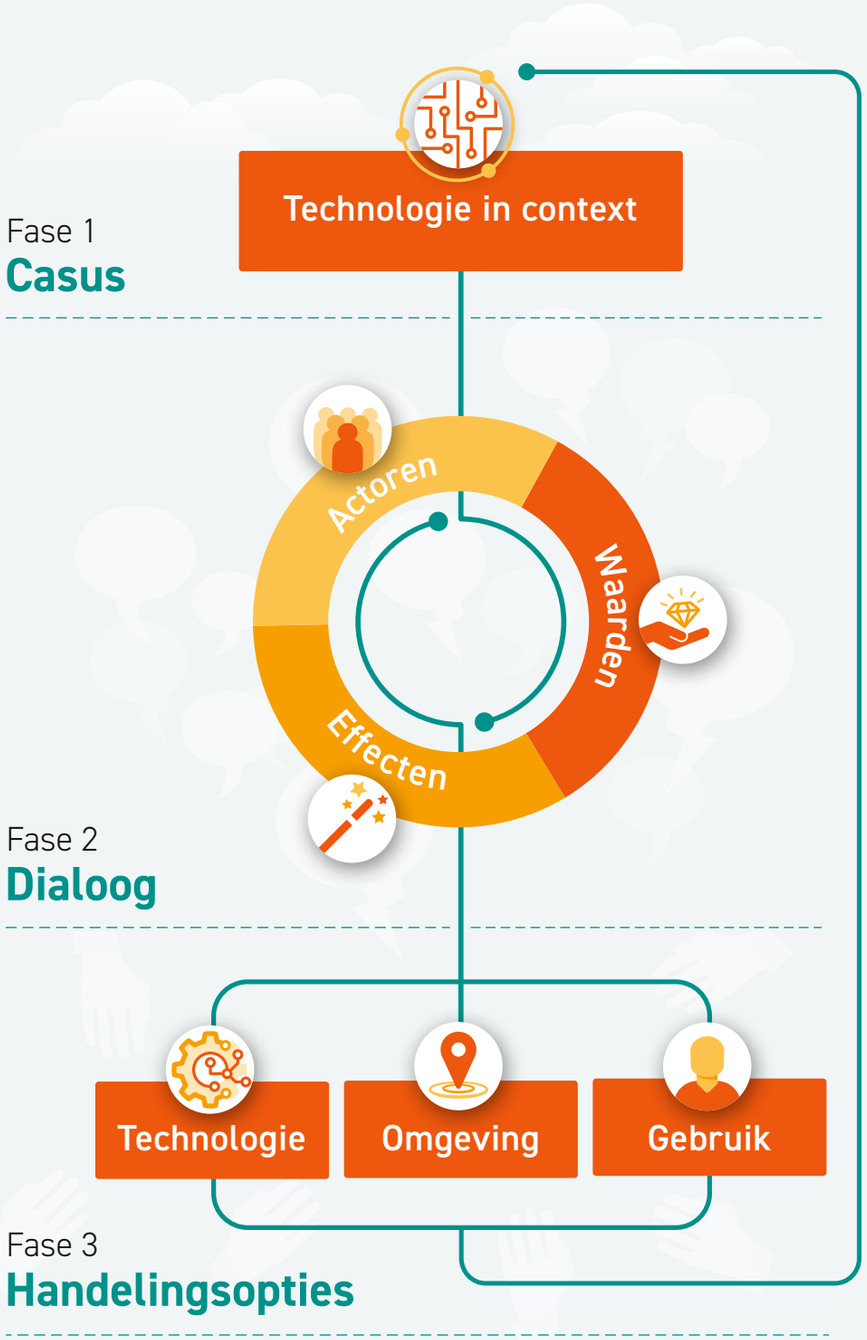
Mensen kunnen ten aanzien van het gebruik van technologie meer en minder waardenvol gedrag vertonen. In het verkeer vindt er bijvoorbeeld bewustwording en training plaats via verkeerslessen op school, zijn er rijlessen die tot een rijbewijs kunnen leiden en is er bewustwording door een BOB-campagne en verkeersborden met spelende kinderen erop.

In dit hoofdstuk zijn verschillende elementen benoemd die belangrijk zijn bij een praktische vertaling van begeleidingsethiek. Deze elementen zijn de bouwstenen voor de aanpak die in het komende hoofdstuk wordt gepresenteerd.



Hoofdstuk 4
Toelichting aanpak begeleidingsethiek

In dit hoofdstuk worden de principes vertaald naar een methode: de aanpak begeleidingsethiek. De figuur op pagina 32 geeft de hoofdelementen van het model weer in drie stappen.



Fase 1 **Casus: technologie in context**

Beschrijf de technologie en de context waarin die technologie functioneert. Het gaat erom dat we goed begrijpen waar we mee te maken hebben. Een deel van de discussies over technologie gaat eerder over (positieve of negatieve) beelden van die technologie dan over hoe zij daadwerkelijk functioneert en wat het voor mensen betekent. Omdat we kiezen voor een concrete technologie in een concrete context hebben we de mogelijkheid hier tot een vrij precieze beschrijving te komen. Zowel van de technologie, als van wat het gebruik van de technologie betekent in de context. Het gaat daarbij om een heldere begrijpelijke beschrijving, zonder te veel jargon of technische details. De beschrijving moet voor een geïnteresseerde buitenstaander te begrijpen zijn.

Fase 2 **Dialoog: actoren, effecten, waarden**

In deze stap wordt de casus verder uit gewerkt. We kennen de technologie en context, maar willen weten wat de mogelijke effecten zijn van de inzet van een technologie in die context. We willen weten wie erbij betrokken zijn en welke waarden een rol spelen.

Actoren

In context wordt meestal snel helder wie de actoren zijn. Eventueel kan ook aan de partijen worden gevraagd wie er verder nog betrokken zou moeten worden. Generiek gaat het vaak om cliënten/burgers, professionele gebruikers, beleidsmedewerkers/bestuur, ontwerpers. Bij de zorgcases zijn bijvoorbeeld bijna altijd patiënten, zorgprofessionals en mantelzorgers betrokken.

Idealiter hebben de echte betrokkenen inbreng in de dialoog. Dat hoeven niet altijd mensen te zijn die actief betrokken zijn bij het gebruik van de technologie. Het gebruik van een technologie kan ook op niet-gebruikers een enorme impact hebben, denk maar aan de invloed van auto's op voetgangers. Als het niet kan om alle betrokkenen deel te laten nemen, zijn er wellicht mensen die hen vertegenwoordigen of bereid en in staat

zijn vanuit hun perspectief te denken. Ook academici of andere experts die zich met het onderwerp bezighouden, kunnen betrokken worden voor het brede maatschappelijke perspectief.

Effecten

Het gebruik van de technologie heeft allerlei effecten. Sommige effecten kunnen onmiddellijk duidelijk zijn, waar andere zich mogelijk voordoen, bijvoorbeeld in de toekomst of onder specifieke omstandigheden. De eerste stap is de effecten zo open mogelijk te verzamelen, zonder te beoordelen op wenselijkheid of waarschijnlijkheid. Daarna is het goed en pragmatisch om te benoemen welke effecten het meest relevant zijn.

De dialoog blijft uitgaan van de concrete technologie en context en van daaruit komen verschillende mogelijke effecten aan bod. Het onderscheiden van verschillende effecten kan helpen bij het verkrijgen van een rijk en reëel beeld:

- positieve en negatieve effecten,
- bekende en te verwachten effecten,
- directe en indirecte effecten,
- effecten voor verschillende actoren,
- effecten op verschillende niveaus: individueel (micro), sociaal (meso) en maatschappelijk (macro).

Waarden

Er spelen altijd verschillende waarden rondom technologie. Denk aan rechtvaardigheid, toepasbaarheid, betrouwbaarheid, solidariteit, respect, autonomie. Vaak blijven ze in discussies impliciet, omdat kritiek op technologie meer concreet wordt geformuleerd. Als iemand bij de AI-toepassing in de GGZ-instelling het gevoel heeft te worden gecontroleerd, kan de onderliggende waarde autonomie, of privacy zijn.

Bij de meeste cases zijn verschillende waarden belangrijk. Net als bij de effecten gaat het eerst om een open inventarisatie, vervolgens welke waarden het meest relevant worden geacht. Daarbij blijft het belangrijk de ‘minder relevante’ waarden in het oog te houden.

Op dit moment worden er in vele sectoren, in bedrijven en door de overheid verschillende ethische codes of richtlijnen geschreven. Een kleine greep op het gebied van AI alleen leidt tot de volgende voorbeelden: Google AI Principles, Microsoft AI Principles, UK initial code of conduct for data driven health and care technology, European Commission High Level Expert Group on AI, Smart Dubai AI Principles, Nesta principles for public sector use of AI, Ethische Code AI Nederland ICT, AI Impact Assessment (ECP), Ethisch verantwoord innoveren: 7 principes voor de overheid (BZK). Als voorbeeld staat in het kader de lijst die door de High Level Expert Group on AI van de EU is vastgesteld. Voor de aanpak begeleidingsethiek zijn dit inspiratiebronnen om te komen tot de waarden die een rol spelen bij de technologie in context.

European Commission High Level Expert Group on AI

1. Accountability
2. Data Governance
3. Design for all (by all -include diversity)
4. Governance of AI Autonomy (Human oversight)
5. NonDiscrimination
6. Respect for Human Autonomy
7. Respect for Privacy
8. Robustness
9. Safety
10. Transparency

https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai

Dialoog

Deze fase wordt 'dialoog' genoemd. In een open uitwisseling tussen de betrokken actoren wordt duidelijk wat de mogelijke effecten en belangrijke waarden rondom het gebruik van een technologie zijn. Een dialoog is een belangrijk onderdeel van de aanpak en vindt vaak plaats in een workshopsetting, zeker als de betrokkenen rondom een technologie een gezamenlijke verantwoordelijkheid voelen en zien dat ze elkaar nodig hebben om vervolgstappen te zetten. Delen van de dialoog vinden uiteraard ook plaats buiten de workshop. Het gesprek gaat door.

Aanvullend en ter voorbereiding op het directe gesprek kunnen ook andere middelen worden ingezet. Interviews leveren op een andere manier inzicht op in de effecten en waarden volgens de gesprekspartners. Literatuuronderzoek kan vaak een waardevolle aanvulling zijn, zeker indien gevolgd door een goede analyse.

Een goede dialoogfase heeft enkele opbrengsten. Ten eerste worden er vaak onduidelijkheden weggenomen. De inbreng van verschillende soorten kennis geeft iedereen een beter beeld en dus een beter zicht op de mogelijke effecten. Door het doordenken van de effecten wordt duidelijk waar de verwachtingen en angsten liggen en mogelijk hoe reëel die zijn.

Ten tweede brengen de actoren ook verschillende perspectieven in en daarmee komen de belangrijkste waarden naar boven die spelen rondom deze technologie binnen deze context. Dit leidt vaak tot begrip, omdat mensen dan beter in staat zijn zich te verplaatsen in het perspectief van de ander. Ze hoeven het niet met het belang van de waarde van de andere eens te zijn, maar kunnen begrip opbrengen voor het belang dat de ander eraan hecht.

Er zijn ook andere methoden die een dergelijke dialoogfase ondersteunen, bijvoorbeeld een moreel beraad, een socratische gesprek e.d. Dat laat al zien dat deze fase op zich waardevol is. Bij de aanpak begeleidingsethiek is deze fase een onderdeel, ingebed in ideeën over techniekontwikkeling, van de eerste fase en vooral als input voor de derde fase; het komen tot handelingsopties.

Fase 3 **Handelingsopties**

De kern van de aanpak begeleidingsethiek is om technologie op een waardenvolle/ethische wijze in de samenleving te begeleiden en tevens de samenleving te begeleiden bij het waardenvol/ethisch gebruik van nieuwe technologie. Dat vraagt om handelen. Daarom ligt de nadruk niet alleen op het komen tot een goed gesprek of het vaststellen van een ethische code, maar op het komen tot handelingsopties. Er worden drie soorten handelingsopties onderscheiden: vanuit de technologie, vanuit de context en vanuit de gebruiker. Na een korte uitleg volgt een kader met daarin voorbeelden van handelingsopties van de vier cases.

Technologie, ethics by design

Ethics by design staat al langer in de belangstelling. Inmiddels is het een best practice om momenten van ethische reflectie in het ontwerptraject van een technologie op te nemen, zodat ethische waarden daadwerkelijk belanden in het ontwerp van technologie. Ethiek is niet alleen een zaak van mensen maar ook van technologieën. Elke technologie heeft invloed op de keuzes en het gedrag van mensen. Een 'ethics by design'-aanpak geeft doelbewust vorm aan die invloed, vanuit expliciete ethische reflectie.

Ethics by design gebeurt soms ook zonder het te benoemen. In een goed functionerende markt geven klanten aan wat ze verwachten van een product en zal de producent zich daaraan proberen aan te passen, inclusief de waarden die daarbij horen, zoals veiligheid, duurzaamheid, schoonheid, toepasbaarheid etc.

Veel markten zijn echter imperfect. In de zorg hebben klanten bijvoorbeeld weinig inkoopmacht, omdat die aan de verzekeraar is overgegeven en bij platforms ontstaan vaak monopoloïde situaties. Daar heeft de overheid natuurlijk een rol: zij kan randvoorwaarden opleggen door wet- en regelgeving. Dat is echter een vrij traag en weinig precies instrument. Bovendien laten niet alle waarden zich optimaal vertalen in marktwerking. Daarom is het interessant om ook op andere wijze te proberen de ontwerpers in een vroeg stadium ethische waarden mee te laten nemen in hun werk.

Handelingsopties: ontwerp, ethics by design

AI in psychiatrische zorg:

- Professioneel oordeel verplicht: een angst is dat psychiaters blind op de AI-toepassing gaan vertrouwen. Een technologische oplossing is dat het AI-oordeel pas toegankelijk wordt nadat ze hun eigen analyse hebben ingevuld.
- Meer transparantie: het algoritme doet voorspellingen die wel goed, maar weinig transparant zijn. Mogelijk kan er transparantie aan het systeem worden toegevoegd, zodat het begrijpelijker en beter uit te leggen wordt op basis waarvan het algoritme tot zijn aanbevelingen komt.

Newsfeed Facebook:

- Nep-accounts weghalen: veel 'vervuiling' op sociale media komt van nepaccounts. Hiertegen is 'real name policy' van Facebook een middel.
- Fakenieuws herkennen en lager ranken: clickbait en sensationalisme leiden tot vergelijkbare patronen (weglatingen in titel, teleurstelling bij gebruiker) en zijn daarmee herkenbaar door AI-algoritmes. Deze kunnen dit soort content vervolgens lager in tijdslijnen zetten.

Eetrobot:

- Beter afwasbaar: het blijkt dat de robotarm waarmee men eet niet alleen gebruikt wordt om te eten, maar bijvoorbeeld ook om mee op je hoofd te krabben (net als mensenhanden).
- Praten is niet eten: kan er AI in de eetrobot worden ingebouwd, zodat het eten pas naar de mond wordt gebracht als de gebruiker niet praat?

Zicht op ondermijning

- In verband met privacy is de keuze gemaakt voor representatie op wijkniveau, niet op persoonsniveau.
- Er is de keuze gemaakt om geen blackbox te creëren. Dat beperkt het aantal mogelijke AI-toepassingen.

Omgeving, ethics in context

Bij de ethische dialoog over technologie ligt de focus al snel op die technologie, terwijl de omgeving of de context waarin die technologie functioneert zeker zo belangrijk is. Die context is vaak ook verschillend. Het maakt uit of Facebook wordt gebruikt door iemand van 13, 35 of 80 jaar, of het binnen werkcontext gebruikt wordt of privé. Het maakt uit of een AI-toepassing in de zorg of in de bouw wordt gebruikt.

Het is dus belangrijk meer aandacht te schenken aan die omgeving, ook omdat daar een deel van de oplossingen ligt. Als je technologie ontwikkelt, zul je als ontwerper en als gebruiker na moeten denken wat dat betekent voor het systeem waarin die technologie terecht komt. Een organisatie of een systeem zal veranderen na het invoeren van een technologie. Hoe wordt die verandering vormgegeven? Kan dat op een manier waarbij rekening wordt gehouden met de waarden?

Die aanpassing van de omgeving kan overigens zowel fysiek als sociaal of juridisch zijn. Denk bij de invoering van de destijds nieuwe technologie auto aan een stoep, een rotonde of een verkeerslicht als fysieke aanpassingen, aan onderlinge afspraken in het verkeer als sociaal, en de afspraken die kunnen worden vastgelegd in een verkeerswet als juridisch.

Waar bij 'ethics by design', ontwerpers en (tech-)bedrijven vaak de sleutel in handen hebben, zijn bij het aanpassen van de omgeving juist vooral organisaties (meso) en de overheid (macro) aan zet. Een bedrijf dat robots invoert, zal aanpassingen doen voor de veiligheid van werknemers, de opleidingen, de nieuwe processen etc. Voorbeelden daarvan zijn de aanleg van wegen, invoering van de algemene verordening gegevensbescherming (AVG) en het stellen van randvoorwaarden aan een persoonlijke gezondheidsomgeving, via het vaststellen van (MedMij) standaarden.

Handelingsopties: omgeving

AI in de psychiatrische zorg

- Inrichten 'rode knop'/meldpunt waar een patiënt of betrokkene het oordeel van de AI-toepassing kan laten negeren als hij denkt dat het echt niet klopt. Dan gaat de 'oude' manier van werken weer in gang.
- Afspraak dat het altijd gaat om een dubbel oordeel, van de professional en van de AI-toepassing.
- Afspreken met familie, cliënt en collega's wanneer je AI inzet en wanneer niet. Eventueel met de mogelijkheid er geen gebruik van te maken.

Newsfeed Facebook

- Samenwerking met externe factcheckers die disclaimers kunnen bieden bij berichten.
- Voorbereiding wetgeving over transparantie van politieke advertenties.

Eetrobot

- Er zijn afspraken rondom veiligheid nodig. Wie is verantwoordelijk als er iets misgaat, als door gebruik van de eetrobot iemand verwond wordt.
- Vergoeding van technologie (eetrobot, robotarm) die zelfstandig eten en drinken bevordert, is wenselijk. Daarvoor zijn duidelijke vergoedingsregels noodzakelijk.
- Afspraken en werkwijzen van de zorgprofessionals moeten worden aangepast. Bijvoorbeeld over hygiëne.

Zicht op ondermijning

- Alle analyses op de data vindt plaats op basis van een vier ogen principe om te voorkomen dat er bias ontstaat.
- CBS hanteert een output controle waar bij automatisch èn handmatig gecontroleerd wordt of er geen tot-persoon-herleidbare informatie uit de analyse van de omgeving wordt gehaald.

- Er is sprake van logging van alle analyses door de analisten waarbij “onrechtmatige” onderzoeken (die buiten de onderzoeksvraag liggen) gesignaleerd worden.
- Analisten mogen in de afgelopen twee jaar geen werkzaamheden hebben verricht in de opsporing.
- Analisten moeten een uitgebreide geheimhoudingsverklaring tekenen.

Goed gebruik, ethics in use

Er ligt ook ruimte bij de mens voor het ethisch gebruik van technologie. Mensen kunnen voorzichtig of roekeloos, goed of slecht getraind omgaan met technologie. De eerste stap is bewustwording. Wat doet of kan een technologie en wat kan ik, als gebruiker dan doen? Handelingsopties kunnen de vorm hebben van voorlichtings- of bewustwordingscampagnes, maar ook via school, via mond-op-mond, via nieuwsdragers kun je je bewust worden van een nieuwe technologie en de implicaties ervan.

De tweede stap is daadwerkelijk gedragsverandering. Vaak betekent dat training en oefening. Dat kan variëren van het lezen van een gebruiksaanwijzing tot het doen van een studie. Soms is gedrag niet moeilijk om uit te voeren, maar wel moeilijk te veranderen. Het vraagt een andere gewoonte of discipline, zoals bijvoorbeeld niet met alcohol op achter het stuur. Iedereen kan het, maar het gaat om de acceptatie dat dit inderdaad gevaarlijk is en dat een drankje laten staan niet uncool is. Het verkrijgen van een rijbewijs is een voorbeeld van iets dat voor de meeste mensen juist wel veel training vergt, voordat het gebruik van de technologie veilig onder de knie is.

Goed gebruik

AI in de psychiatrische zorg

- Als zorgprofessional kunnen uitleggen wat je doet en waarom, op dezelfde manier als dat bij een MRI-scan gebeurt. Hier moeten zij training voor krijgen.
- Als patiënt weten dat AI wordt ingezet en de mogelijkheid kennen van een second opinion.

Newsfeed Facebook

- Training mediawijsheid: weten dat nepnieuws bestaat, manieren om het te herkennen en te testen.
- De redactierol van kranten en uitgevers populariseren zodat meer mensen in staat zijn zelf de poortwachtersrol te vervullen.
- Het inrichten van een plek waar mensen correcte en betrouwbare informatie kunnen vinden over vaccinaties, voeding of gezondheid.

Eetrobot

- Zorgmedewerkers hebben ook hulp en training nodig. Zij moeten weten wat er met de technologie kan en wat veilig is.
- Zorgmedewerkers krijgen door de eetrobot een andere rol. Zij hoeven minder vaak te zorgen of op een andere manier. Hiervoor kan een training nodig zijn.
- Er kan van alles met de robotarm, meer dan alleen eten. Voor gebruikers kan het behulpzaam zijn samen uit te vinden hoe de robotarm gebruikt/bediend kan worden. Een suggestie is om een middag te ‘freestylen’ met enkele supergebruikers. Om zo te leren wat er kan en wat jou past.

Zicht op ondermijning

- Trainen van beleidsmedewerkers en andere gebruikers van de uitkomsten van de analyses over de waarde en beperkingen van data-analyses.
- Een zeer beperkt aantal personen heeft toegang tot de micro-data van het CBS en er zijn hoge eisen aan het opleidingsniveau. Statistische wetenschappelijke analyses zijn altijd goed gedocumenteerd gecontroleerd en op herhaalbaarheid getoetst.

Wat opvalt is dat het mogelijk is voor elk van de cases te komen tot handelingsopties in alle drie de categorieën.

De gebruikers in meer rollen

We zien dat de gebruiker bij alle drie de handelingsopties in verschillende rollen opduikt. In discussies worden die rollen vaak verward, dus we willen ze hier nog een keer benoemen.

- Bij ethics by design gaat het erom dat de gebruiker een stem heeft bij het ontwerp van technologie.
- Bij context en omgeving gaat het erom dat de gebruiker mee kan praten over hoe die omgeving wordt aangepast. Er zijn voorbeelden te over waar de invoering van een computersysteem zonder overleg met de medewerker is gebeurd, met talloze frustraties van dien.
- Bij goed gebruik gaat het om de omgang met technologie en context door de gebruiker. De gebruiker kan beter of slechter gedrag vertonen.

Bij bijvoorbeeld internetbankieren zien we die rollen terug. Technologisch wordt betalen via internet telkens makkelijker gemaakt. Dat gaat van een betere interface via de website, tot de introductie van een app op de telefoon de mogelijkheid niet alleen te betalen, maar ook betalingsverzoeken (tikkies) te sturen.

Een voorbeeld van het inrichten van de omgeving, was toen een bank de data van klanten geanonimiseerd wilde doorverkopen. Dat leverde veel oppositie op. De technologie maakte iets mogelijk, maar hier werden door klanten randvoorwaarden aan gesteld. Ze wilden niet dat hun data gedeeld werden, hoe geanonimiseerd ook.

Het gebruik van internetbankieren is steeds makkelijker geworden, maar veel (oudere) mensen moesten het leren en kregen daarbij ondersteuning van bijvoorbeeld ouderenbonden. Ook proberen de banken 'goed gedrag' te stimuleren, bijvoorbeeld door de campagne waar mensen erop worden gewezen te kijken naar het 'slotje' bij de internetpagina, om te checken of ze op een veilige site zitten.

Hoofdstuk 5
Gebruik van de aanpak begeleidingsethiek

In het vorige hoofdstuk hebben we de aanpak begeleidingsethiek beschreven. In dit hoofdstuk gaan we in op de vraag wanneer en hoe je die aanpak in kunt zetten.

Het is duidelijk dat de aanpak begeleidingsethiek een rol kan spelen wanneer er ethische vragen ontstaan rondom een technologie. Dat kan zijn bij de introductie van een nieuwe technologie, maar het kan ook juist wanneer die technologie zich breder of op een andere manier in de samenleving manifesteert. In dit hoofdstuk wordt ingegaan op de potentie van de methode (tussen klein en groot), op het gebruik binnen een organisatie of in de publieke ruimte en gaan we in op hoe een workshop kan worden voorbereid.

Binnen een organisatie of in een maatschappelijk speelveld

De aanpak begeleidingsethiek is zowel te gebruiken binnen één organisatie als binnen een maatschappelijk speelveld met meerdere partijen.

Organisatie

Binnen een organisatie is de eerste vraag wie de probleemeigenaar is. Ethiek behoort meestal niet tot het primaire proces dus het zal waarschijnlijk moeite kosten er aandacht voor te krijgen. Ethiek wordt vaak gezien als extra, als showstopper of iets van enkele liefhebbers. De aanpak begeleidingsethiek komt in haar aard aan een aantal bezwaren tegemoet en opent perspectieven om op een opbouwende manier bezig te zijn met ethiek. De kracht van de aanpak binnen een organisatie is dat handelingsopties relatief snel omgezet kunnen worden in acties.

Blijft de vraag op welke manier, op welke moment en in welk proces het gebruik van de aanpak past. Sommige organisaties hebben ethische commissies, met als taak aandacht te vragen voor ethische dilemma's, bij anderen past het eerder bij afdelingen HRM, scholing, cliëntenraad/klantenpanel of is het juist een zaak van de boardroom.

Wanneer de beslissing genomen is om de aanpak begeleidingsethiek op te pakken gaat het er natuurlijk om de juiste betrokken aan tafel te krijgen. Hier zullen bijna altijd ook mensen van buiten de organisatie (gebruikers, beleidsmakers) bij betrokken zijn.

Grote bedrijven hebben meestal wel een business code of conduct, design principles en soms ook development patterns. Code of conduct zijn gedragsregels voor het hele bedrijf, het eigen gedrag. Design principles gaan echt over hoe we ontwerpen en development patterns gelden voor implementatie, de manier waarop oplossingen bij klanten worden toegepast. De aanpak begeleidingsethiek kan daar een belangrijke aanvulling zijn. Het verbindt binnen en buiten, niet alleen door de gebruiker een betere rol in het ontwerpproces te geven, maar ook door het bedrijf te betrekken bij de context waarin de oplossing gaat functioneren en mee te denken over wat dit voor de vaardigheden van de gebruiker betekent.

Maatschappelijke speelveld

Veel ethische debatten ontstijgen het niveau van één organisatie. Het zijn onderwerpen die velen in de maatschappij aanspreken en die plaats moeten vinden in de publieke ruimte. Partijen die daar voortrekker in kunnen zijn, zijn Ministeries, gemeenten, adviesraden, maatschappelijke groeperingen, actiegroepen, debatcentra, et cetera.

De aanpak begeleidingsethiek is juist geschikt om daar vorm aan te geven. De verschillende actoren worden uitgenodigd vanuit hun perspectief mee te doen aan het debat én worden uitgenodigd te denken in handelingsopties. De aanpak begeleidingsethiek vraagt serieuze interesse in de problematiek van zowel de technologie als de context en daagt uit te komen met suggesties voor handelingsopties. Als het goed is zitten ook de verschillende partijen die iets kunnen doen aan tafel. Op die manier kunnen stappen vooruit worden gezet en wordt de wereld een beetje *waardenvoller.*

Workshop: voorbereiding, vervolg

Groot of klein, publiek of privaat, bijna bij alle dialogen wordt er een workshop gehouden. In deze paragraaf volgen enkele suggesties voor het inrichten van een workshop bij de aanpak begeleidingsethiek. We volgen de drie fases.

Technologie in context:

- **Scherp formuleren technologie**
  - Is iedereen voldoende op de hoogte van de technologie?
  - Moeten er vooraf stukken worden aangeleverd, op basis van interviews en literatuur?
  - Moet er een expert aanwezig zijn om ter plekke een toelichting te geven en vragen te beantwoorden? Is de expert een buitenstaander of een van de deelnemers?

- **Scherp formuleren omgeving**
  - Kent iedereen de omgeving waarin de technologie wordt ingezet?
  - Moeten er vooraf stukken worden aangeleverd, op basis van interviews en literatuur?
  - Moet er een expert aanwezig zijn om ter plekke een toelichting te geven en vragen te beantwoorden? Is de expert een buitenstaander of een van de deelnemers?

Dialoog

- **Deelnemers**
  - Zoveel mogelijk betrokken actoren aan tafel, liefst met handelingsmacht.
  - Evenwichtig samenstellen van de groep waarbij verschillende perspectieven voldoende zijn vertegenwoordigd.
  - Ter aanvulling eventueel een aantal experts (academici) voor generieke inbreng en duiding of heel specialistische kennis.
  - De groepsgrootte moet zodanig zijn dat er een dialoog gevoerd kan worden. Afhankelijk van het ontwerp van de workshop tussen de 10-30 mensen, in één groep is 15 het maximum.
  - Gevoel voor verhoudingen tussen de deelnemers: zijn er spanningen of juist sterke verbindingen tussen personen.
  - Optie: vooraf deelnemers interviewen levert vaak scherpere focus van de bijeenkomst en heeft als bijkomend voordeel dat deelnemers zich ook al hebben voorbereid.

- **Mogelijk vooraf of tijdens de workshop te brengen informatie:**
  - Geen: tijdens de bijeenkomst wordt iedereen voldoende geïnformeerd en wordt gekozen voor de creativiteit van de groep.
  - De technologie en de context.
  - Lijst met relevante waarden.
  - Lijst met mogelijke effecten.

- **Aandachtspunten workshopopzet:**
  - Goede moderator.
  - Heldere, gedeelde doelstelling.
  - Onderscheid maken tussen het creatief proces en het keuzeproces.
  - Eventueel subgroepen om te verdiepen en te versnellen.
  - Zorgen voor een gemeenschappelijke afsluiting met helderheid over het vervolg.
  - Goede verslaglegging.

Handelingsopties

- Deze fase kan per soort handelingsopties besproken worden: technologie, omgeving, gedrag.
- Deze fase kan ook vanuit verschillend actorperspectief besproken worden: dus bij de eetrobot: welke handelingsopties zijn er vanuit de patiënt gezien, welke vanuit de zorgverlener?
- Voor handelingsopties geldt: hoe concreter hoe beter. Vraag bij vage suggesties door, maar verlies je niet in details, die komen later.
- Het is makkelijk handelingsopties bij de ander neer te leggen. Probeer te kijken of handelingsopties geadresseerd kunnen worden bij verschillende personen aan tafel.
- Probeer een sfeer te creëren waarin mensen elkaar helpen hun verantwoordelijkheden te nemen en handelingspotentie te gebruiken.

Belangrijk is natuurlijk wat er gaat gebeuren na de workshop. Heel concreet zouden een aantal handelingsopties in handelingen kunnen worden omgezet. Een deel van de opties zal een proces van jaren vergen, bijvoorbeeld een nieuw ontwerp of de invoering van een wet. Dan kunnen er natuurlijk wel eerste stappen gezet worden. Een ander deel zal 'laaghangend fruit' zijn; opties die direct kunnen worden omgezet in handelingen. Zo kan er een freestyle middag georganiseerd worden voor de gebruikers van een eetrobot of kan er een overleg worden ingericht om over de veiligheid ervan te beslissen.

Het kan ook zijn dat het onderdeel dialoog juist vraagt om vervolg. Dat het gesprek dat in de workshop gevoerd is, vraagt om vervolg. Dat kan door meerdere workshops, door een publicatie, door lezingen etc.

Uiteindelijk is de aanpak begeleidingsethiek interactief. We geloven niet in eindoplossingen, maar wel in verbeterstappen.

Hoofdstuk 6
Vervolg

De aanpak begeleidingsethiek zoals beschreven in het voorgaande hoofdstuk is een product van de ECP-werkgroep met mensen uit het bedrijfsleven, de overheid, maatschappelijke sectoren en ethici/techniekfilosofen. Het is gelukt, met veel dank aan de inbrengers van de cases, het gedachtegoed vanuit de filosofie praktisch bruikbaar te maken. Dit is wat ons betreft echter slechts de eerste stap, een uitnodiging zelfs, ook aan u als lezer, om vervolgstappen te zetten.

Verspreiding en verbetering

Het concept begeleidingsethiek geeft de mogelijkheid om op een andere manier over technologie en menselijke waarden te denken. In discussies, in krantenartikelen en gesprekken over innovaties merken we dat bijna altijd nog uitgegaan wordt van een fundamentele scheiding tussen mens en technologie, van 'top-down' denken vanuit 'de' technologie, in plaats van 'bottom-up' denken van vele technologieën in een context. Ook de methodes die gebruikt worden om met ethische dilemma's om te gaan hebben veelal het klassieke schema als achtergrond.

We merken dat het concept van begeleidingsethiek aanslaat en herkend wordt, maar ook dat het een stevige dialoog vraagt om zich deze denkwijze eigen te maken. Het gesprek daarover zal nog lang en breed gevoerd moeten worden.

Dat gesprek over begeleidingsethiek is niet alleen een academische discussie, hij is uiterst relevant voor de praktijk. Dit is het moment waarop digitaliseringsontwikkelingen hun plek vinden in de maatschappij, in alle domeinen. Artificiële intelligentie (AI) was bijvoorbeeld lang een belofte en begint die nu waar te maken, met alle ethische vragen van dien. En willen we blijven innoveren, dan hebben we een methode nodig die ons daarbij begeleidt, die ethisch mee-innoveert. En dat is de aanpak begeleidingsethiek.

Breed gebruik van de aanpak begeleidingsethiek

We hebben de aanpak begeleidingsethiek ontwikkeld, op basis van theorie en gebruik van cases. We geloven dat de aanpak stevig staat én weten dat hij zich door moet ontwikkelen. Dat kan alleen door hem te gebruiken in de praktijk. Er zijn nog veel aspecten waarover veel geleerd kan worden. Bijvoorbeeld over het gebruik van de workshop binnen organisaties, de doorwerking van workshops, de begeleiding van de workshops, de afbakening van de case, de manier waarop de resultaten in een vergelijkbare situatie kunnen worden gebruikt, de kracht van de aanpak voor een maatschappelijke discussie.

We denken dat de aanpak rijp genoeg is om voordeel op te leveren voor organisaties of trajecten, die hem willen gebruiken. Tegelijkertijd willen we de aanpak blijven verbeteren. We hopen dit te kunnen begeleiden en een netwerk te bouwen dat dit kan ondersteunen.

Ethiekfabriek met handelingsopties

Als de aanpak begeleidingsethiek vaak wordt uitgevoerd, komen er veel handelingsopties: handelingsopties voor technisch ontwerp, voor de aanpassing van de omgeving en voor beter gebruik. Deze handelingsopties zijn waarschijnlijk breder bruikbaar dan alleen voor de case. We hebben dan ook de ambitie de handelingsopties te verzamelen, opdat ook anderen er gebruik van kunnen maken. Daarvoor bedachten we de term ethiekfabriek.

Het gaat om het beeld, dat de handelingsopties breder bruikbaar zijn en ook voor andere partijen beschikbaar komen en dat er veel mogelijkheden zijn. Verspreiding kan natuurlijk ook gericht gebeuren. Stel bijvoorbeeld dat er verschillende handelingsopties voor de eetrobot bij een zorginstelling worden benoemd. Dan is het interessant die opties te gebruiken en verder aan te vullen bij andere instellingen, maar ook bij andere ontwikkelaars van de eetrobot en andere gebruikers, zodat de gevonden oplossingen breed en gedeeld ingang vinden.

Gebruik de aanpak begeleidingsethiek, werk mee aan de ethiekfabriek

B

Bijlage 1

Deelnemers werkgroep digitalisering en ethiek

Bijlage 1 – **Deelnemers werkgroep digitalisering en ethiek**

Pallas Agterberg	Alliander
Brigitte Boon	Siza, Academy Het Dorp
Hans Bos	Microsoft Nederland B.V.
Roxane Daniels	Vereniging Nederlandse Gemeenten (VNG)
Valerie Frissen	SIDN fonds
Edo Haveman	Facebook
Jos Huigen	KPN
Esther Keymolen	Universiteit Leiden
Lana Klok	ECP | Platform voor de InformatieSamenleving
Natalja Krijgsman	ECP | Platform voor de InformatieSamenleving
Anja Lelieveld	Ministerie van BZK (2019)
Roelof Meijer	SIDN
Rob Nijman	IBM Nederland BV
Marc Noordhoek	Ministerie van Binnenlandse Zaken en Koninkrijksrelaties
Bettine Pluut	Erasmus Universiteit Rotterdam
John Post	Topsector Energie
Sabine Roeser	Technische Universiteit Delft
Daniel Tijink	ECP | Platform voor de InformatieSamenleving
Arie van Bellen	ECP | Platform voor de InformatieSamenleving
Erik van de Poel	Dienst Koninklijk Huis
Jeroen van der Ham	Nationaal Cyber Security Centrum (NCSC)
Theo van der Plas	Nationale Politie
Sandra van der Weide	Ministerie van Economische Zaken en Klimaat
Angeline van Doveren	Rijnstate Ziekenhuis
Peter-Paul Verbeek	Universiteit Twente
Focco Vijselaar	Ministerie van Economische Zaken en Klimaat

Bijlage 2
Meer lezen

Bijlage 2 – **Meer lezen**

ECP | Platform voor de InformatieSamenleving, *Het verhaal van digitaal, samen vormgeven aan onze digitale samenleving*, ECP, 2018

ECP | Platform voor de InformatieSamenleving, *Artificial intelligence impact assessment*, ECP 2018

Esther Keymolen en Simone van 't Hof, *Can I still trust you, my dear doll? A philosophical and legal exploration of smart toys and trust*, Journal of Cybersecurity, issue2, vol 4, 2019, p. 143-159

Sabine Roeser, *Risk, Technology and moral emotions*, Routledge 2018

F.E. Scheepers, V. Menger, K. Hagoort, *Datascience in de psychiatrie*, Tijdschrift voor psychiatrie 60(2018)3, 205-209

Peter-Paul Verbeek, *Op de Vleugels van Icarus, hoe techniek en moraal met elkaar meebewegen*, Lemniscaat, 2014

Peter-Paul Verbeek, *De grens van de mens, over techniek, ethiek en de menselijke natuur*, Lemniscaat, 2011

Jan Van Zanen e.a., *Verlenging CITY DEAL Zicht op Ondermijning (nr 48699)*, Staatscourant, 14 februari 2019