Aan de slag met chatbots in welzijn en zorg?



Praktijkgids voor verantwoord professioneel gebruik van conversationele AI

Tim Vanhove, Sylvie Bernaerts, Tom Seymoens,
George Caique Gouveia Barbosa, Tom Van Daele,
Jana Verplancke & Dieter De Witte

artevelde hogeschool
GHENT UNIVERSITY
THOMAS MORE
met de steun van VLAIO

Colofon

Het projectteam:

Tim Vanhove, Tom Seymoens & Jana Verplancke – Mens, Samenleving & Digitalisering, Arteveldehogeschool
Sylvie Bernaerts & Tom Van Daele - Psychologie en technologie, Thomas More-hogeschool
George Caique Gouveia Barbosa & Dieter De Witte – IDlab, Universiteit Gent

Contacteer de projectleider Tim Vanhove via **Tim.Vanhove@Arteveldehs.be**

Meer informatie op **www.arteveldehogeschool.be/ai-cares**

Citeersuggestie:

Vanhove, T., Bernaerts, S., Seymoens, T., Barbosa, G.C.G., Van Daele, T., Verplancke, J., & De Witte, D. (2026). Aan de slag met chatbots in welzijn en zorg? Praktijkgids voor verantwoord professioneel gebruik van chatbots. Arteveldehogeschool: Gent. **www.arteveldehogeschool.be/ai-cares**

Bij de totstandkoming van deze tekst werd generatieve AI als ondersteunend instrument gebruikt. De technologie werd ingezet om inspiratie te verkrijgen, inhoud te (her)structureren, conclusies samen te vatten en voorlopige versies taalkundig na te lezen. Alle gegenereerde suggesties werden door de auteurs kritisch beoordeeld, inhoudelijk gecontroleerd en waar nodig aangepast.

De auteurs zijn volledig verantwoordelijk voor de inhoud van deze publicatie.

Lay-out: Buro Knal

© 2026 Arteveldehogeschool, Thomas More-hogeschool en Universiteit Gent
met steun van VLAIO

Inhoud

Colofon p. 2

Inleiding p. 4

1

Begrijpen

Wat is conversationele AI?

Wat kan AI betekenen voor welzijn en zorg? Ontdek hoe chatbots werken, waar hun kracht ligt en waar de risico's zitten. Verken bestaande AI-tools en krijg inzicht om er kritisch en nieuwsgierig mee om te gaan.

p. 7

2

Voorbereiden

Maak organisatiekeuzes en bied beleidskaders

AI gebruiken begint niet bij technologie, maar bij je keuzes als organisatie. Hoe bepaal je wat wel en niet kan, welke data je mag gebruiken en wie verantwoordelijk is? Bouw stap voor stap aan een helder en gedragen AI-kader.

p. 32

3

Doen

Planmatig testen, implementeren en bijsturen

Van idee naar praktijk: ontdek hoe je goede prompts schrijft, een eigen chatbot opzet, kennisbronnen slim inzet en veilig test. Met heldere stappen bouw je gericht aan een AI-tool op maat van jouw organisatie.

p. 49

Referenties p. 66

Meer weten p. 68



Inleiding

AI-chatbots zijn in korte tijd doorgedrongen tot het dagelijkse leven. Ze voeren gesprekken die verrassend menselijk klinken. Ook in welzijn en zorg ontstaan daardoor nieuwe mogelijkheden. Tegelijk roept die evolutie fundamentele vragen op: wanneer is zo'n chatbot behulpzaam? Welke informatie kan je eraan toevertrouwen? Wie bewaakt de kwaliteit?

Deze praktijkgids nodigt organisaties uit om artificiële intelligentie (AI) niet zomaar te omarmen of af te wijzen, maar er bewust mee aan de slag te gaan. Je ontdekt wat conversationele AI wel en niet kan, hoe je duidelijke organisatiekeuzes maakt en hoe je toepassingen veilig kan testen, ontwikkelen en bijsturen. Het doel is niet om voor alles een chatbot te bouwen, maar om te onderzoeken waar AI meerwaarde kan bieden, met professionele kwaliteit steeds als uitgangspunt.

De praktijkgids is tot stand gekomen binnen het TETRA-project '**AI-CARES**', met steun van het Vlaams Agentschap Innovatie en Ondernemen (VLAIO). AI-CARES is een project onder leiding van de Arteveldehogeschool in samenwerking met de Thomas More-hogeschool en Universiteit Gent. Tijdens dit project brachten we met een lerend netwerk de noden en vragen uit het welzijns-en zorgveld in kaart en verkenden we welke opportuniteiten er zijn volgens software-en AI-ontwikkelaars. Via proeftuinen werden de praktische aspecten van het gebruik van conversationele AI uitgezocht en uitgetest.

In de proeftuinen werden **chatbotprototypes** getraind op specifieke kennisbronnen uit welzijn en zorg. In die proeftuinen werden, samen met organisaties uit het werkveld, diverse trainingstechnieken getest om te weten te komen of en hoe AI-chatbots betere gesprekken kunnen voeren over psychosociale thema's. Door te werken met echte data konden we veel leren over wat vandaag al werkt, wat nog moeilijk blijft en welke voorwaarden nodig zijn om AI verantwoord in welzijn en zorg in te zetten.

Er werden zes chatbotprototypes getest:

1. Via een **informatiechatbot** onderzochten we hoe een chatbot informatie beter kan afstemmen op de vraag en situatie van de informatiezoeker, zonder aan inhoudelijke kwaliteit in te boeten.
2. De **triple chatbot** liet zien welk effect ethische begrenzingen kunnen hebben bij gevoelige thema's.
3. De **trainingschatbot** maakte het mogelijk om chathulpgesprekken te oefenen via rollenspellen met een chatbot en daarbij suggesties en feedback op maat te ontvangen.
4. Een **finetuned chatbot** liet zien wat het effect is van het aanpassen van het onderliggende taalmodel door dit bij te trainen op basis van duizenden voorbeeldgesprekken.
5. Een **intakechatbot** toonde aan hoe het tijdens wachttijden informatie kan bevragen in de vorm van een chatgesprek, zodat het daaropvolgende gesprek met een hulpverlener kan starten met een overzicht van de relevante informatie.
6. Een **zorgmatchingchatbot** onderzocht hoe een zorgvraag gekoppeld kan worden aan het zorgaanbod van een organisatie door in gesprek te gaan met een chatbot.

Ga naar **https://aicares.ilabt.imec.be/nl/** en test ze zelf uit.

Meer informatie over deze chatbots is verder in deze praktijkgids terug te vinden.

Het uitgangspunt van deze praktijkgids is eenvoudig: het gebruik van chatbots is geen doel op zich. Chatbots kunnen pas verantwoord ingezet worden wanneer duidelijk is welk probleem je wil oplossen én hoe een chatbot dat probleem zou kunnen oplossen. Deze praktijkgids is geen technische of juridische handleiding, maar een informerend, inspirerend en adviserend hulpinstrument.

De praktijkgids is modulair opgebouwd, zodat de verschillende onderdelen afzonderlijk kunnen worden geraadpleegd. Start bij het deel dat het best aansluit bij de actuele vraag binnen je organisatie:

1. Wil je begrijpen wat AI is en wat de kansen en risico's zijn?
   Start bij **Deel 1: Begrijpen**.
2. Wil je je voorbereiden en afspraken maken over wat wel en niet mag met AI?
   Start bij **Deel 2: Voorbereiden**.
3. Wil je aan de slag met een chatbot?
   Start bij **Deel 3: Doen**.

Doorheen de tekst staan praktijkgerichte checklists. Deze zijn bedoeld als hulpmiddel om het gesprek over AI te ondersteunen en tot gegronde besluitvorming te kunnen komen. Gebruik ze dus niet als eenvoudige afvinklijst, maar als startpunt om in team na te denken. Betrek bij het gebruiken van de checklists medewerkers vanuit verschillende perspectieven: praktijkwerkers, leidinggevenden, ICT, communicatie, kwaliteitszorg, beleid, data protection officer (DPO), juridische adviseurs en waar mogelijk ook gebruikers of burgers.

AI-gebruik vraagt geen eenmalige beslissing. Technologie, regelgeving, ervaringen en maatschappelijke verwachtingen veranderen snel. Wat vandaag niet verantwoord of haalbaar is, kan later wel mogelijk worden. Omgekeerd kan een toepassing die vandaag zinvol lijkt, later bijsturing vragen. Gebruik deze praktijkgids daarom als onderdeel van een blijvend proces van leren, evalueren en bijsturen.



1

Begrijpen

Wat is conversationele AI?

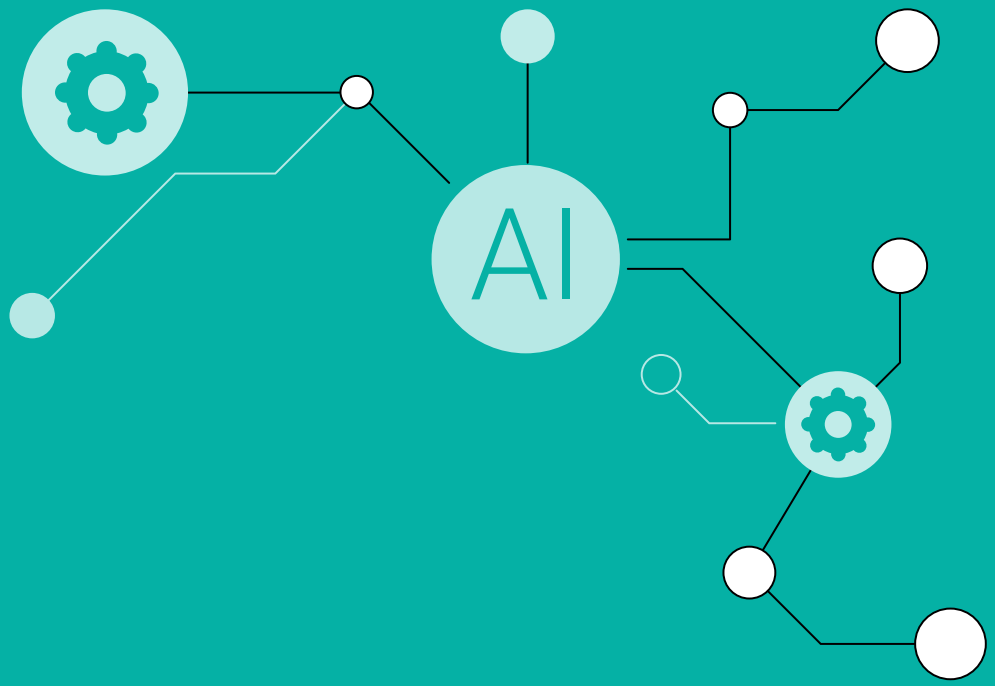
- Wat is AI?
- De kansen van chatbots voor welzijn en zorg
- De risico's van AI
- Wat betekent de EU AI ACT voor welzijn en zorg?
- Hoe kan je een taalmodel zelf trainen?
- Ben je klaar voor AI met je team?
- Samenvatting

Leeswijzer

Wat is conversationele AI en wat zijn de centrale begrippen? In dit deel ontdek je een gedeelde basis om over conversationele AI te spreken. Het is geen technische cursus, maar een laagdrempelige uitleg over wat conversationele AI doet, waar de risico's zitten en welke kansen er zijn voor welzijn en zorg.

Lees dit deel vooral wanneer je merkt dat AI in je organisatie veel vragen, weerstand of onzekerheid oproept. Het helpt om voorbij slogans te gaan zoals "AI is gevaarlijk" of "AI zal ons werk overnemen". Je krijgt taal om genuanceerder te spreken over wat AI wel en niet kan.

Na dit deel kan je als team beter inschatten waarom AI geen gewone zoekmachine is, waarom AI-output altijd gecontroleerd moet worden, waarom de grote overtuigingskracht van chatbots risicovol kan zijn en waarom menselijke verantwoordelijkheid centraal blijft.



Wat is AI?

Wat is conversationele AI precies en hoe kan het ingezet worden in jouw organisatie? Om dit te begrijpen, helpt de analogie van een restaurant. Je kan conversationele AI beschouwen als de volledige restaurantkeuken en bediening. Een taalmodel is dan de kok en een chatbot is de ober die met de klant praat. De klant (jij dus) ziet enkel de ober (chatbot), maar daarachter zit een hele organisatie die wordt aangedreven door de expertise van de kok (het taalmodel). Een chatbot is dus de zichtbare gebruiksinterface van een conversationeel AI-systeem dat wordt aangestuurd door het taalmodel. We leggen de kernwoorden verder uit hieronder.

Taalmodellen

Taalmodellen zijn goede taalvoorspellers. Taalmodellen worden namelijk getraind op enorme hoeveelheden tekst zodat ze er heel goed patronen in kunnen herkennen. Dit werkt zo goed dat ze ook zelf taal kunnen genereren, gebaseerd op die patronen. Chatbots maken gebruik van die taalmodellen om context-gevoelige gesprekken te voeren die erg menselijk aanvoelen (Stryker, 2021).

Taalmodellen zijn gebaseerd op zowat *alle* teksten in *alle* talen op het *hele* internet. Ze zijn dan ook groot, vandaar de naam *Large Language Model* (*LLM*). Ze zijn getraind met vele miljarden stukken tekst en hebben gaandeweg patronen in de taal gevonden. Een deel van deze patronen kennen wij zelf ook - denk aan de spelling-en grammaticaregels - maar een deel van die patronen gaat veel dieper: een taalmodel kan goed inschatten welke woorden bij elkaar horen in een bepaalde context.

Soms geeft een taalmodel echter ook een antwoord dat fout is, maar toch geloofwaardig klinkt. Dat wordt dan hallucineren genoemd. Dit komt doordat het taalmodel zich niet baseert op kennisbronnen, maar zinnen samenstelt op basis van kansberekening. De tekst zal er qua vorm goed uitzien, want dat kan zo'n model heel goed, maar de inhoud is iets anders. Het verschil tussen goed of fout, tussen betrouwbaar of onbetrouwbaar, moéten wij als mens steeds maken. **Een taalmodel begrijpt niet wat het zelf zegt. Het simuleert begrijpen op basis van kansberekening.**

Conversationele AI en chatbots

Conversationele AI is de technologie die gesprekken mogelijk maakt met chatbots, maar ook stemassistenten, virtuele medewerkers of smart speakers (bv. Google Home Assistant of Alexa). Een chatbot is dus één manier om conversationele AI te gebruiken: het is een interface waarbij je een gesprek voert via tekst.

Hoe werkt het concreet? Je geeft input aan een chatbot - je schrijft een zin, of een 'prompt' - en die taal wordt dan gebruikt als inspiratie of input om nieuwe output te creëren door de chatbot. Het model moet telkens inschatten wat het meest logische volgende stukje tekst is op basis van statistische waarschijnlijkheid. Er wordt dan *steeds* iets nieuws gecreëerd of gegenereerd: het is altijd volledig nieuwe tekst en dus geen copy-paste van het internet.

Waarom voelt een gesprek met een chatbot zoals ChatGPT soms alsof je met een mens praat? ChatGPT begrijpt taal niet zoals wij dat doen, maar is uitzonderlijk goed in het voorspellen van woorden. Daardoor kan het antwoorden formuleren die natuurlijk, vloeiend en overtuigend klinken.

GPT: Generative Pre-trained Transformers

Waar staat GPT eigenlijk voor?

- **Generative**: het taalmodel genereert telkens nieuwe tekst.
- **Pre-trained**: het taalmodel is vooraf getraind op een zeer grote hoeveelheid teksten.
- **Transformer**: het taalmodel vergelijkt woorden met elkaar om zo de context te begrijpen.

Een kerntechniek in een taalmodel is het gebruik van *transformers*. Door te werken met een transformer kijkt een taalmodel niet gewoon één voor één naar woorden, maar probeert het tegelijk te begrijpen hoe alle woorden met elkaar verbonden kunnen zijn. Het legt verbanden tussen woorden doorheen een langere tekst en kan daardoor context veel beter vatten. Net als mensen in een gesprek onbewust meer 'aandacht' geven aan bepaalde woorden om betekenis te construeren, doen transformers dat mathematisch: ze wegen elk woord af tegenover elk ander woord in een zin of een paragraaf om te bepalen welke woorden belangrijk zijn voor de betekenis van dat ene woord. Het kan zo de context meenemen in het bepalen van betekenis van een woord.

Hierdoor kan een taalmodel complexe verbanden leggen, zoals: wie wat zegt, wat belangrijk is in een zin, welke volgorde logisch is. Het kijkt naar alle woorden in een zin tegelijk om te begrijpen hoe ze met elkaar samenhangen en grijpt ook terug naar de context in vorige zinnen.

Een voorbeeld: "*De pizza kwam uit de oven en ze smaakte erg lekker*". Als mens begrijp je meteen en vanzelf dat die oven niet lekker is, maar wel de pizza. Een transformer gaat op zoek in zijn patronen om te kijken hoe vaak 'oven' en 'lekker' samen te vinden zijn en hoe vaak 'pizza' en 'lekker'. Op basis van de waarschijnlijkheid in die patronen maakt het snel een keuze welke woorden het best bij elkaar horen. Het begrijpt daardoor de noodzakelijke context.

Zonder transformers hebben AI-modellen het moeilijk met taal. Voordat transformers bestonden konden AI-systemen moeilijk langere gesprekken volgen. Ze vergaten wat je eerder zei of gaven rare, onsamenhangende antwoorden. Dankzij transformers kunnen taalmodellen de rode draad vasthouden, context onthouden en zinnen genereren die grammaticaal correct én inhoudelijk coherent zijn. Het AI-model merkt dus hoe de woorden passen in het grotere geheel.

De eerste therapeutische 'chatbot' was er al in 1966

De oorsprong van conversationele AI gaat terug tot de jaren zestig. Toen werd al geëxperimenteerd met chatbots die automatisch antwoorden konden formuleren. Het bekendste voorbeeld is de psychotherapeutische chatbot **ELIZA**, beschreven door onderzoeker Joseph Weizenbaum (1966). Op de zin *"Ik voel me verdrietig"* kon de chatbot bijvoorbeeld antwoorden: *"Waarom voel je je verdrietig?"* Dat lijkt vandaag eenvoudig, maar was revolutionair voor die tijd. ELIZA werkte nog niet met een taalmodel, maar met duizenden vooraf geprogrammeerde regels. De chatbot herkende bepaalde woorden of zinsconstructies en koppelde daar een antwoord aan. Dat werkt soms, maar niet altijd. ELIZA is nog altijd actief: probeer ze gerust eens uit, en maak uit voor jezelf of je dit nog AI zou noemen vandaag: **www.masswerk.at/elizabot/**

De kansen van chatbots voor welzijn en zorg

De grote administratieve druk en personeelstekorten, de stijgende zorgvraag en lange wachttijden in zorg en welzijn stimuleren de zoektocht naar nieuwe vormen van ondersteuning. Conversationele AI wordt daarbij steeds vaker naar voren geschoven als een veelbelovende technologie vanwege haar schijnbaar empathisch en menselijk klinkende interacties. AI-technologie evolueert razendsnel en er zijn al vrij veel toepassingen te vinden op de markt, maar wetenschappelijk bewijs voor de voordelen of effecten van deze toepassingen ontbreekt nog vaak.

We geven hier een overzicht van de mogelijkheden van AI in welzijn en zorg. We tonen hoe het kan worden ingezet als ondersteuning en aanvulling op menselijk handelen, niet als vervanging ervan.

1. Mogelijke AI-tools voor professionals

AI kan professionals in welzijn en zorg ondersteunen bij dagelijkse taken, vooral als hulpmiddel voor **taal, structuur en denkwerk**. Ze kan helpen om moeilijke teksten om te zetten naar duidelijke taal, vakjargon te verminderen, teksten te herschrijven op B1-niveau (met alledaagse woorden, korte zinnen en een actieve schrijfstijl), lange documenten samen te vatten of e-mails helderder of formeler te maken. Tools zoals **Leessimpel** of **Schrijfsimpel** kunnen hierbij bruikbaar zijn. Belangrijk blijft dat de professional zelf het eerste concept schrijft en de eindcontrole doet, en geen persoonsgegevens of vertrouwelijke informatie invoert.

AI kan dienen als sparringpartner bij **denkprocessen en brainstorms**. De professional kan AI gebruiken om ideeën te verzamelen voor activiteiten, creatieve titels, werkvormen of mogelijke aanpakken. AI levert dan inspiratie, maar de professional beoordeelt wat haalbaar, passend en verantwoord is binnen de concrete context.

Ook **vertaaltools** zoals DeepL of Google Translate kunnen helpen om folders, teksten of gesprekken te vertalen. Dankzij AI houden ze beter rekening met context en betekenis. Toch blijven fouten mogelijk, omdat de correcte betekenis in de juiste context niet altijd duidelijk is. Gebruik zulke tools daarom vooral wanneer je de doeltaal enigszins kan controleren en wees transparant naar de cliënt dat je een vertaaltool gebruikt.

1. Mogelijke AI-tools voor professionals

AI-tools gericht op administratieve ondersteuning

AI-tools gericht op training van en kennisondersteuning voor hulpverleners

AI-tools gericht op ondersteuning van triage en planning

AI-tools gericht op ondersteuning in hulpverlening

Veel hulpverleners besteden een aanzienlijk deel van hun tijd aan administratie, terwijl ze deze eigenlijk liever zouden gebruiken samen met cliënten of gebruikers. Conversationele AI zou deze belasting kunnen verminderen.

AI kan functioneren als **tekstverwerker**: ruwe nota's omzetten naar een verslag, teksten samenvatten, herschrijven, structureren of inkorten, actiepunten uit een tekst halen, taakverdelingen maken of informatie bundelen volgens vooraf gekozen hoofdstukken. Belangrijk blijft dat de medewerker zelf nadenkt, de structuur bepaalt en een eerste inhoudelijke aanzet geeft. AI werkt dan verder op professioneel denkwerk, maar neemt het niet over.

Er bestaan ook gespecialiseerde AI-tools om **administratief en inhoudelijk werk te verlichten**. Op basis van zorgvuldig geanonimiseerde nota's kan AI helpen met verslaggeving en casusvoorbereiding. Het kan een gesprek samenvatten, thema's ordenen, een conceptverslag maken of een takenlijst voor de betrokkenen opsommen. Ook kan AI aanvullingen suggereren bij een aanpak die de hulpverlener zelf al heeft bedacht. Belangrijk is dat AI hier niet het denkwerk overneemt. De professional maakt eerst zelf de analyse, brengt context aan, beoordeelt de output kritisch en past de tekst aan. Bij casussen met complexe risico's, verontrusting of ethische dilemma's blijft overleg met collega's en leidinggevenden noodzakelijk. Steden zoals Brugge, Kortrijk, Mechelen en Oostende experimenteren met AI-tools ter ondersteuning van dienstverleners. Er bestaan spraak-naar-teksttools die gesprekken kunnen opnemen, transcriberen en omzetten naar verslagen op maat van de eigen werking. Zulke toepassingen kunnen tijdswinst opleveren, maar vragen duidelijke afspraken over toestemming, privacy, beroepsgeheim, opslag en controle van het verslag.

De chatbot **Maya** is een voorbeeld van een systeem dat geïntegreerd is in elektronische patiëntendossiers. Het ondersteunt medische zorgprofessionals bij het opzoeken van informatie, het navigeren door systemen en het uitvoeren van administratieve handelingen zoals afspraken plannen. Zo zorgt de applicatie ervoor dat verpleegkundigen patiënten na een operatie niet meer moeten opbellen – telefoontjes die vaak onbeantwoord bleven en daardoor tijdrovend waren – doordat patiënten de nodige informatie via de app kunnen doorgeven.

1. Mogelijke AI-tools voor professionals

AI-tools gericht op administratieve ondersteuning

AI-tools gericht op training van en kennisondersteuning voor hulpverleners

AI-tools gericht op ondersteuning van triage en planning

AI-tools gericht op ondersteuning in hulpverlening

Een chatbot kan de rol van cliënt opnemen in een **oefengesprek** met medewerkers in opleiding. Gespreksvaardigheden oefenen is vaak tijdrovend en die rollenspellen worden soms wat als ongemakkelijk ervaren. Een goed geïnstrueerde AI-trainer kan hulpverleners ondersteunen om zowel de juiste gesprekstechnieken als het nodige vertrouwen in hun gespreksvaardigheden op te bouwen. Na afloop kan de tool feedback geven op het gesprek, bijvoorbeeld over toon, structuur, vraagstelling of aansluiting bij de cliënt. Zulke digitale trainingsassistenten zijn vooral interessant omdat ze medewerkers veilig kunnen laten oefenen met realistische situaties, zonder druk op echte cliënten of opleiders. Ook hier blijft menselijke supervisie nodig: AI-feedback kan leerzaam zijn, maar vervangt geen professionele beoordeling of intervisie. Binnen het AI-CARES-project experimenteerden we met een trainingshulp voor SAM vzw die zowel beginnende als ervaren hulpverleners kan helpen om hun gespreksvaardigheden te oefenen.

Daarnaast kunnen chatbots ook gebruikt worden als **informatieassistenten** om medewerkers te ondersteunen om informatie op te zoeken uit veelgebruikte bronnen, bijvoorbeeld doorverwijsadressen, juridische informatie of sociale diensten.



1. Mogelijke AI-tools voor professionals

AI-tools gericht op administratieve ondersteuning

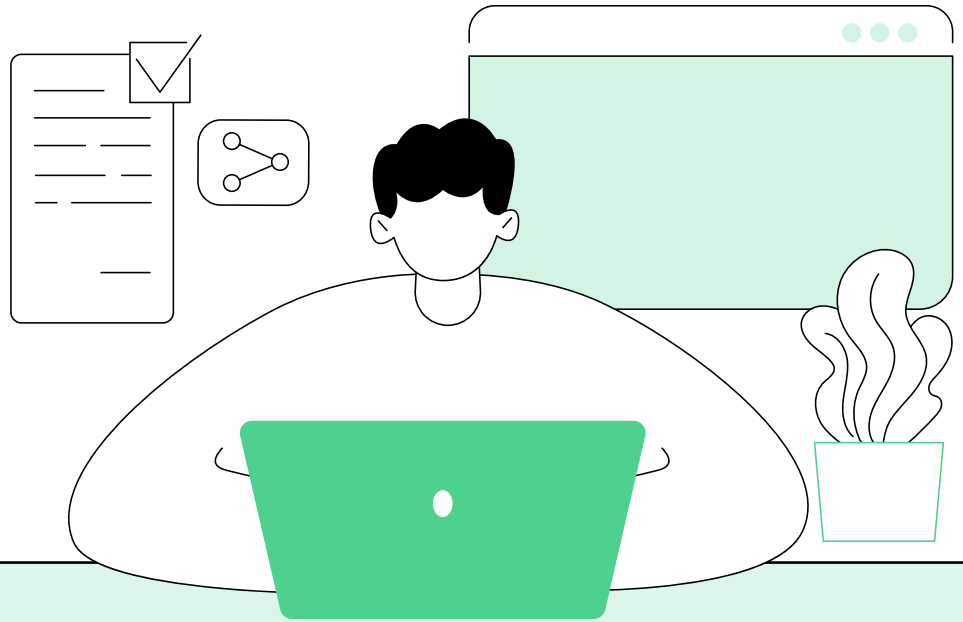
AI-tools gericht op training van en kennisondersteuning voor hulpverleners

AI-tools gericht op ondersteuning van triage en planning

AI-tools gericht op ondersteuning in hulpverlening

Conversationele AI kan repetitieve taken **automatiseren** en zo mogelijk cognitieve belasting verminderen. Een concreet voorbeeld voor de geestelijke gezondheidszorg is het Britse **Limbic Access**, dat wordt ingezet om intakeprocessen en klinische beoordeling te ondersteunen. Limbic Access maakt gebruik van een AI-chatbot die op de website van een organisatie kan ingebed worden. Dit systeem verzamelt informatie en helpt bij triage, waardoor professionals efficiënter kunnen werken en sneller tot een eerste klinische inschatting kunnen komen.

Een ander voorbeeld is **Inquira Health**. Dit platform met een *AI voice assistant* werkt als een virtuele onthaal-medewerker die telefoons kan beantwoorden, annulaties kan verwerken, no-shows kan verminderen en intake kan ondersteunen, zodat medewerkers hun aandacht kunnen focussen op de cliënten die fysiek aanwezig zijn.



1. Mogelijke AI-tools voor professionals

AI-tools gericht op administratieve ondersteuning

AI-tools gericht op training van en kennisondersteuning voor hulpverleners

AI-tools gericht op ondersteuning van triage en planning

AI-tools gericht op ondersteuning in hulpverlening



AI kan ook een waardevolle ondersteuning zijn in de hulpverlening, zolang de technologie wordt ingezet voor de juiste taken. Deze zijn best gericht op ondersteuning, zonder de professionele verantwoordelijkheid van de hulpverlener over te nemen.

Chatbots kunnen als **sparringpartner** helpen bij voorbereiding en reflectie. AI kan ingezet worden om een moeilijk gesprek voor te bereiden, psycho-educatie af te stemmen op de doelgroep, een eerste gespreksleidraad te maken of mogelijke aandachtspunten te verkennen. De hulpverlener blijft daarbij degene die inschat wat passend en veilig is. AI fungeert hier als sparringpartner: het helpt om taal te vinden, opties te verkennen en structuur aan te brengen, maar het beslist niet.

Het **Psy-AID project** van Arteveldehogeschool en Howest onderzoekt hoe AI-beelden kunnen worden ingezet om psycho-educatie meer op maat te maken van de noden van de cliënten. De psycho-educatieve informatie wordt zo meer toegankelijk en inclusief.

Hoewel deze toepassingen nog onderzocht moeten worden, worden chatbots ook verkend als tools om cliënten te ondersteunen tussen sessies met een professional. Zo toonden Michael Heinz en collega's van Dartmouth College aan hoe de chatbot **Therabot** effectief symptomen van depressie, eetstoornissen en angst kon verbeteren (Heinz, et al, 2025). Een belangrijke opmerking hierbij is wel dat zulke toepassingen voor gebruik in Europa de nodige certificering als **medisch hulpmiddel** dienen te hebben.

2. Mogelijke AI-tools voor burgers

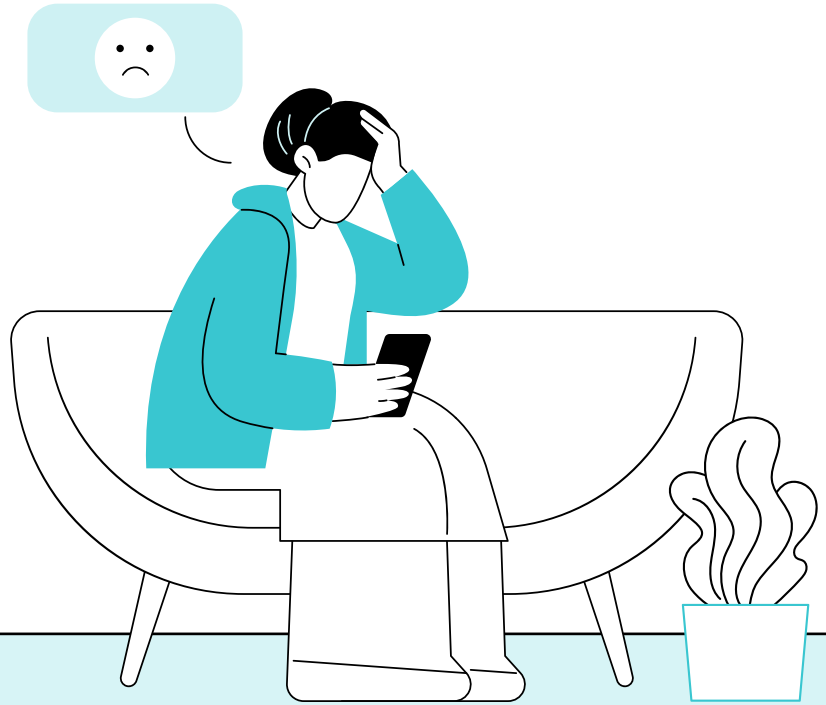
AI-informatie-assistenten

AI-tools gericht op het in handen nemen van welzijn en gezondheid

Zorg en welzijn worden vaak als complex en moeilijk toegankelijk ervaren. Conversationele AI kan hier fungeren als gids die burgers helpt de juiste weg te vinden.

De chatbot **Agathe/Agatha**, ontwikkeld voor het OCMW van Sint-Agatha-Berchem, ondersteunt burgers bij administratieve en sociale vragen en helpt bij het navigeren door de rechten en dienstverlening. Op een vergelijkbare manier lanceerde het Academisch Ziekenhuis Brussel de **virtuele assistent Aurore**, vooral om taalbarrières te overbruggen en informatie toegankelijker te maken voor anderstaligen.

In het AI-CARES-project exploreerde De Ambrassade het gebruik van een AI-informatieassistent om informatie van hun jongerenwebsite **WatWat** toegankelijker en beter aan jongeren te kunnen aanbieden in de plaats van de traditionele zoekbalk.



2. Mogelijke AI-tools voor burgers

AI-informatie-assistenten

AI-tools gericht op het in handen nemen van welzijn en gezondheid

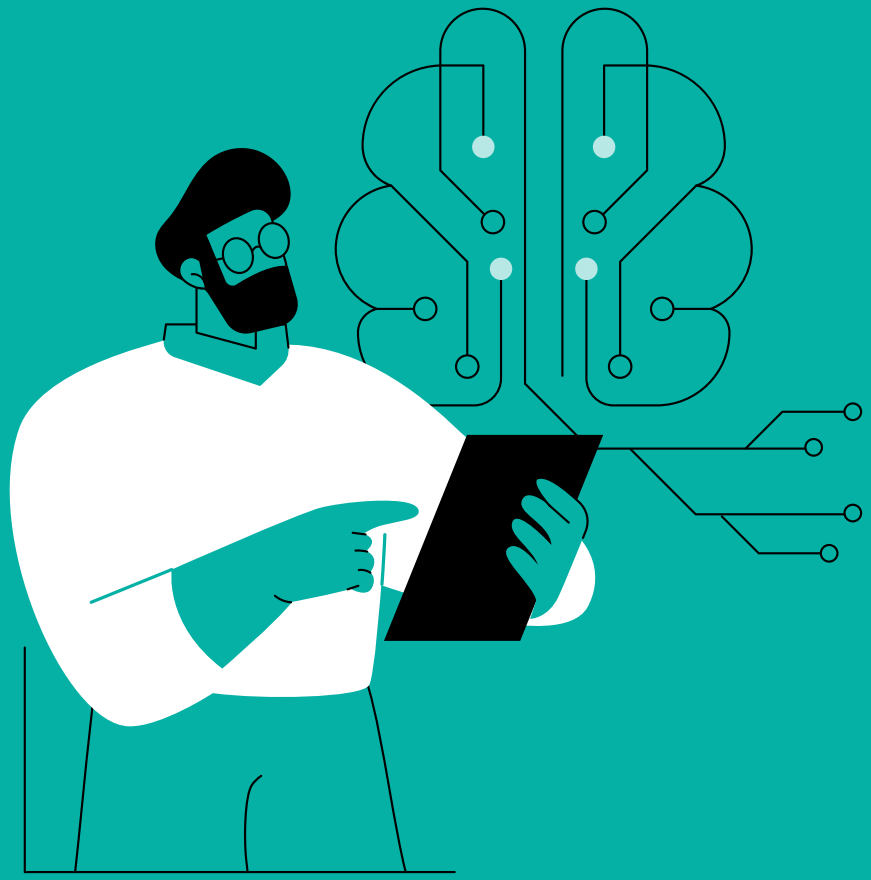
Hoewel mensen al generieke AI-chatbots als ChatGPT en Copilot gebruiken om welzijnsvragen te stellen en medische informatie op te zoeken, zijn deze chatbots hiervoor niet correct getraind. Omwille van deze nood zijn er wel bedrijven die specifiek inzetten op medische ondersteuning van burgers. **Symptoomcheckers** zoals **Ada Health** bieden gebruikers de mogelijkheid om hun klachten te analyseren en een eerste inschatting te krijgen van mogelijke oorzaken en vervolgstappen. Dit ondersteunt potentieel niet alleen de patiënt, maar zou kunnen helpen om ook zorgvragen beter te oriënteren en onnodige consultaties te vermijden. Sommige toepassingen gaan nog een stapje verder. Zo richt de **Cerina Companion met de chatbot Hedgie** zich op het ondersteunen van patiënten op een wachtlijst via een aanbod van cognitieve gedragstherapie, symptoom tracking en check-ins. Deze toepassingen zijn evenwel vaak, ondanks samenwerkingen met zorgprofessionals en training met klinische gegevens, onvoldoende onderzocht op veiligheid en effectiviteit.

Voor blinden en **slechtzienden** bieden apps zoals EMVI en Be My Eyes ondersteuning door omgevingen te beschrijven via AI of door gebruikers via video te verbinden met vrijwilligers of AI.

Internationaal hebben verschillende **chathulporganisaties** al verkend hoe ze chatbots zouden kunnen inzetten om hun dienstverlening te verbeteren. Zo experimenteerde het Nederlandse 113.nl zelfmoordpreventie met AI-ondersteuning in haar hulplijn, zoals het analyseren van chatgesprekken en het ondersteunen en trainen van de hulpverleners. In Finland biedt **Sekasin Chat** jongeren een laagdrempelige chatomgeving, waar AI kan bijdragen aan het opvangen van piekmomenten of momenten waarop de dienst niet beschikbaar is. De chatbot kan erkenning geven, basisinformatie bieden en doorverwijzen, maar vervangt de menselijke hulpverlener niet. Bij verontrustende signalen rond suïcide moet de chatbot de grenzen van zijn rol erkennen en verwijzen naar menselijke hulp. Otherkind ontwikkelde naast de Sekasin chatbot voor kinderen en jongeren ook een zogenaamde **Moodbot** voor volwassenen met vragen rond welzijn en emoties.

Chatbots zijn vaak gratis, permanent beschikbaar, vereisen geen fysieke verplaatsing en bieden een gevoel van anonimiteit, wat drempels verlaagt, in het bijzonder bij gevoelige thema's zoals mentale gezondheid. Toepassingen zoals **Ebb (Headspace)** en **Ello (HelloBetter)** worden dan in de markt gezet als *AI companions for mental wellbeing*, vaak als een manier om niet officieel beschouwd te worden als een medisch hulpmiddel en dus ook niet te hoeven voldoen aan de bijhorende wetgeving. Deze apps bieden gebruikers gestructureerde gesprekken, psycho-educatie en oefeningen aan gebaseerd op meditatie-en ademhalingstechnieken.

De risico's van AI



Hoewel er verschillende interessante AI-tools zijn voor welzijn en zorg, is het belangrijk om ook de beperkingen en risico's te kennen. Zeker in die sectoren kunnen fouten of verkeerd gebruik een grote impact hebben op gebruikers, professionals en organisaties. De kern is dat conversationele AI krachtig en bruikbaar kan zijn, maar in welzijn en zorg altijd moet worden ingezet binnen duidelijke professionele, ethische en organisatorische grenzen.

In dit deel bespreken we de belangrijkste risico's die uit onderzoek naar voren komen. Ook in het AI-CARES lerend netwerk en de proeftuinen kwamen risico's en beperkingen aan het licht. Hoewel onderzoekers ondertussen heel wat studies naar de bruikbaarheid, veiligheid en effectiviteit van chatbots doen, zijn de meeste conversationele AI-toepassingen echter nog niet gevalideerd in de 'echte wereld'. De risico's die we hieronder melden zijn dan ook degene die op het moment van schrijven aangetoond zijn. AI-ontwikkelingen of nieuwe versies van taalmodellen volgen elkaar sneller op dan onderzoek kan volgen. Nieuwe taalmodellen en AI-toepassingen kunnen deze risico's dan aanpakken, of net nieuwe risico's met zich meebrengen. Bespreek het gebruik van conversationele AI binnen je organisatie dan ook steeds met experten zoals je Data Protection Officer (DPO) of juridische adviseur.

1

Gebrek aan transparantie.

Bij veel AI-systemen is het niet duidelijk op welke gegevens het model werd getraind, welke instructies het volgt, welke kwaliteitscontroles en beperkingen zijn ingebouwd en welke commerciële belangen spelen. Daardoor is het moeilijk om te beoordelen welke aannames en beperkingen achter een AI-antwoord schuilgaan. Voor welzijn en zorg is dit problematisch, omdat professionals en gebruikers moeten kunnen inschatten waarop informatie, aanbevelingen of adviezen gebaseerd zijn.

Ook kan onduidelijkheid ontstaan over de mate waarin AI werd gebruikt bij het opstellen van teksten, beelden, presentaties of andere inhoud. Wat werd gemaakt door een mens en wat met behulp van een taalmodel? Dit vergt transparantie.

HOE GA JE ERMEE OM?

Informeer je grondig.

Informeer je over de trainingsdata en guardrails van AI-tools. Op dit moment is dit vaak niet mogelijk als individu of organisatie, maar mogelijks wordt dit mogelijk naarmate de modellen verder evolueren.

Wanneer je organisatie zelf een chatbot ontwikkelt, maak dan publiek toegankelijk welke technologie wordt gebruikt, welke instructies of systeemprompts gelden en hoe de kwaliteit wordt bewaakt.

Bepaal hoe je eigen AI-gebruik als organisatie moet worden vermeld. Hoe groter de inhoudelijke bijdrage van AI en hoe groter de mogelijke impact op gebruikers, hoe belangrijker transparantie wordt. Leg hierover een organisatierichtlijn vast. Een mogelijke standaardformulering is: “Een deel van deze inhoud is opgesteld of bewerkt met behulp van generatieve AI. De inhoud werd door een medewerker gecontroleerd en goedgekeurd of aangepast waar nodig.”



HOE GA JE ERMEE OM?

2

Gebrek aan betrouwbaarheid.

AI herkent patronen, maar begrijpt ze niet. Generatieve AI blijft vatbaar voor hallucinaties: antwoorden die overtuigend klinken, maar feitelijk fout, onvolledig of verzonnen zijn. Daarnaast kunnen vooroordelen in de trainingsdata zichtbaar worden in de output. In welzijn en zorg is dat extra gevoelig, omdat correctheid en veiligheid er cruciaal zijn.

Controleer de output.

Blijf zelf aan het stuur en neem verantwoordelijkheid door alle output grondig na te lezen op juistheid en volledigheid en te herwerken, gebaseerd op je eigen expertise, ervaring en betrouwbare kennisbronnen. Het denkwerk blijft de taak van de organisatie. Check ook op de aanwezigheid van impliciete vooroordelen. Met andere woorden: herschrijf en maak de tekst je eigen.

3

Te weinig of te veel menselijkheid.

AI kan empathie, nabijheid en begrip overtuigend nabootsen, maar voelt niets en begrijpt geen kwetsbaarheid zoals een mens dat kan. Daardoor ontstaat een dubbel risico. Te weinig empathie kan kil of onveilig aanvoelen. Te veel of ongepaste empathie kan mensen het gevoel geven dat ze echte hulp krijgen, terwijl rode vlaggen gemist worden. Bij thema's zoals crisis, suïcidaliteit, geweld, verontrusting of psychische nood kan dan probleemverheldering, crisisinschatting of doorverwijzing ontbreken.

HOE GA JE ERMEE OM?

Bewaak de menselijke relatie.

Het gebruik van algemene chatbots bij crisissituaties, suïcidaliteit, geweld, verontrusting of ernstige psychische nood is niet zonder gevaar voor de gebruikers. Wettelijk gezien ligt de verantwoordelijkheid daarvoor bij die gebruikers.

Wanneer je als organisatie zelf een gespecialiseerde chatbot ontwikkelt, beoordeel dan of AI gepast, veilig en professioneel verantwoord kan antwoorden bij gevoelige thema's, want wettelijk gezien is de organisatie verantwoordelijk. Werk dus met duidelijke veiligheidsinstructies en gepast menselijk toezicht. Maak duidelijk dat gebruikers met AI communiceren.

4

GDPR, privacy en beroepsgeheim.

Vertrouwelijke of privacygevoelige informatie mag nooit in algemene AI-tools worden ingevoerd. Namen, adressen, geboortedata, medische of juridische gegevens, organisatienamen of herkenbare casusdetails kunnen personen identificeerbaar maken en dit is niet toegelaten door de GDPR-wetgeving.

HOE GA JE ERMEE OM?

Organiseer en sensibiliseer.

Voorzie een organisatiekader over professioneel gebruik van AI en sensibiliseer alle medewerkers over het gevaar van het delen van gevoelige gegevens. Verbied het opladen van vertrouwelijke, persoonsgebonden of privacygevoelige informatie.

HOE GA JE ERMEE OM?

Bied duidelijkheid over professioneel AI-gebruik.

Een organisatiekader met richtlijnen is vereist: wat mag wel en niet, onder welke voorwaarden, met welke tools, met welke controle en met welke transparantie.

5

Schaduwgebruik zonder organisatiebeleid.

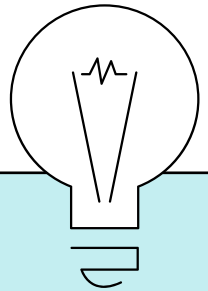
Medewerkers kunnen AI professioneel gebruiken, zonder gedeeld afsprakenkader. Dan beslissen medewerkers zelf welke tool ze gebruiken, welke informatie ze invoeren en hoe ze output beoordelen. Dat creëert risico's op vlak van privacy, kwaliteit en ethiek.

6

Schending van auteursrechten.

AI-tools zijn getraind op grote hoeveelheden tekst. Daarbij is niet altijd duidelijk of alle gebruikte bronnen vrij beschikbaar waren of auteursrechtelijk beschermd zijn. Ook AI-output kan onbedoeld lijken op bestaand en beschermd materiaal.

HOE GA JE ERMEE OM?



Lees zelf de bronnen.

Controleer of de vermelde bronnen vrij overgenomen mogen worden, vermeld die bronnen en maak afspraken over wie eigenaar is van materiaal dat met AI werd gemaakt.

De mens blijft steeds aan zet.

AI is maar een hulpmiddel. AI mag nooit diagnoses stellen, hypotheses over cliënten formuleren of beslissingen voorstellen met impact op de rechten, het welzijn of de gezondheid van de betrokkenen. AI mag geen vervanging worden voor reflectie, overleg, supervisie of vakmanschap.

HOE GA JE ERMEE OM?

7

Verlies van professionele verantwoordelijkheid.

AI mag professionals ondersteunen, maar mag hun professionele autonomie en eindverantwoordelijkheid niet overnemen: AI is ondersteunend en niet beslissend, hoe overtuigend die ook mag klinken. Wanneer medewerkers AI te vaak of te vanzelfsprekend gebruiken, kan hun eigen professionele oordeelsvorming verzwakken. AI kan het denkwerk overnemen. In welzijn en zorg is dat riskant, omdat kwaliteit niet alleen afhangt van correcte informatie, maar ook van professionele sensitiviteit, contextkennis en ethische afwegingen.

8

Onvoldoende beveiliging en datasoevereiniteit.

Zeker in gevoelige sectoren zoals welzijn en zorg is het belangrijk na te gaan waar gegevens worden opgeslagen en verwerkt, welke partijen er toegang toe hebben, onder welke jurisdictie zij vallen en welke controle de organisatie zelf over de data behoudt. Indiende aanbieder commercieel van aard is, dan moeten vragen worden gesteld over hun businessmodel: worden zij rijker door de data van de doelgroep te gebruiken?

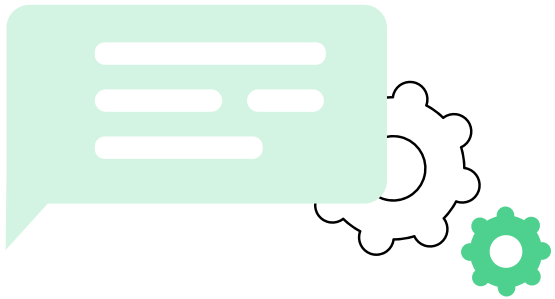
HOE GA JE ERMEE OM?

Kies voor Europese veiligheid, indien mogelijk.

Organisaties werken best met beveiligde omgevingen, bij voorkeur op EU-servers en met Europese technologie, met duidelijke waarborgen over privacy, dataveiligheid en datagebruik.

Deze risico's betekenen niet dat AI geen plaats kan krijgen in welzijn en GGZ. Zoals bij elke nieuwe technologie gaat het er niet om *alle* risico's uit te sluiten, maar ze te herkennen en er bewust mee om te gaan.

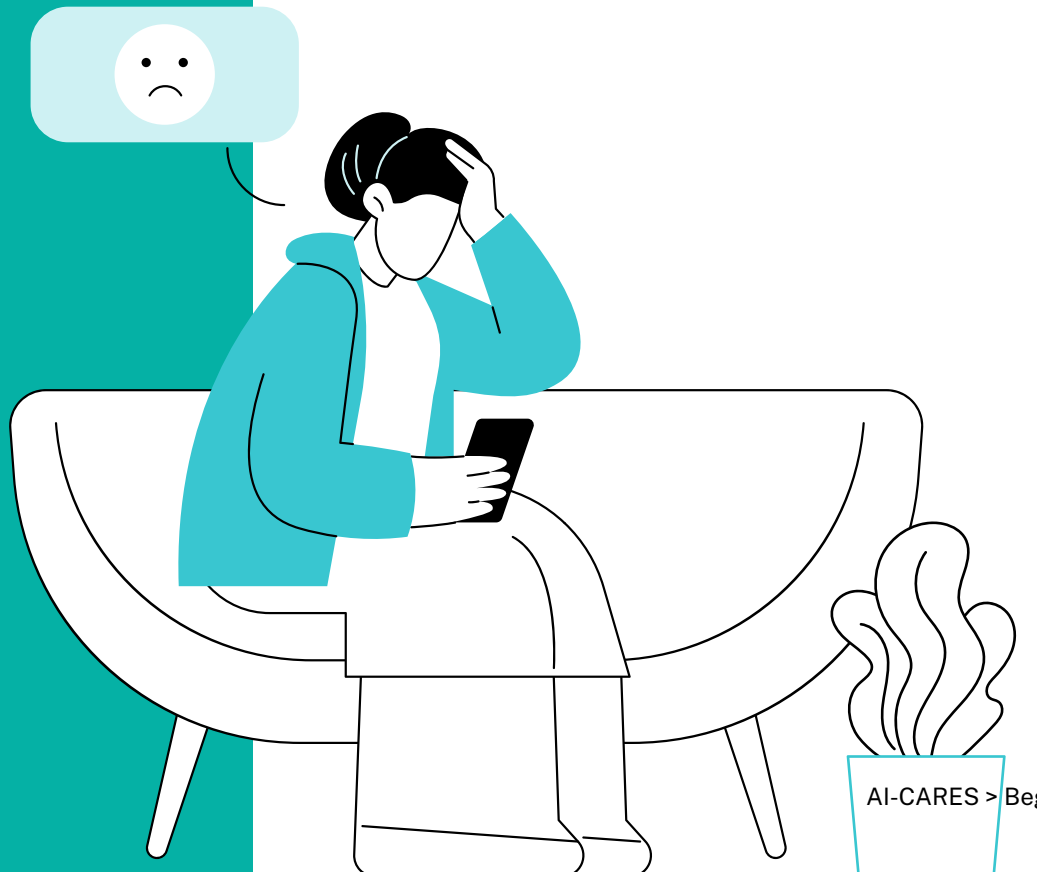
Het is misschien geen rechtstreeks risico voor een individuele organisatie, maar ook de ecologische impact moeten we in de gaten houden als samenleving. Vooral het trainen van AI-modellen vraagt veel elektriciteit en koelwater. Het effect van één prompt lijkt klein, maar door het massale dagelijkse gebruik loopt de totale milieubelasting op. Daarom is ook bewust prompten onderdeel van professioneel AI-gebruik: bewust nadenken op voorhand, helder vragen stellen, overmatig herprompten vermijden en AI alleen inzetten wanneer het echt meerwaarde heeft. De laatste stap zet je zelf: blijf niet eindeloos prompten om elk detail te perfectioneren. Aanvaard dat AI-output niet perfect hoeft te zijn en werk het resultaat zelf verder af. Zo voorkom je dat AI-gebruik een automatisme wordt en blijf je zelf kritisch nadenken en beslissen.

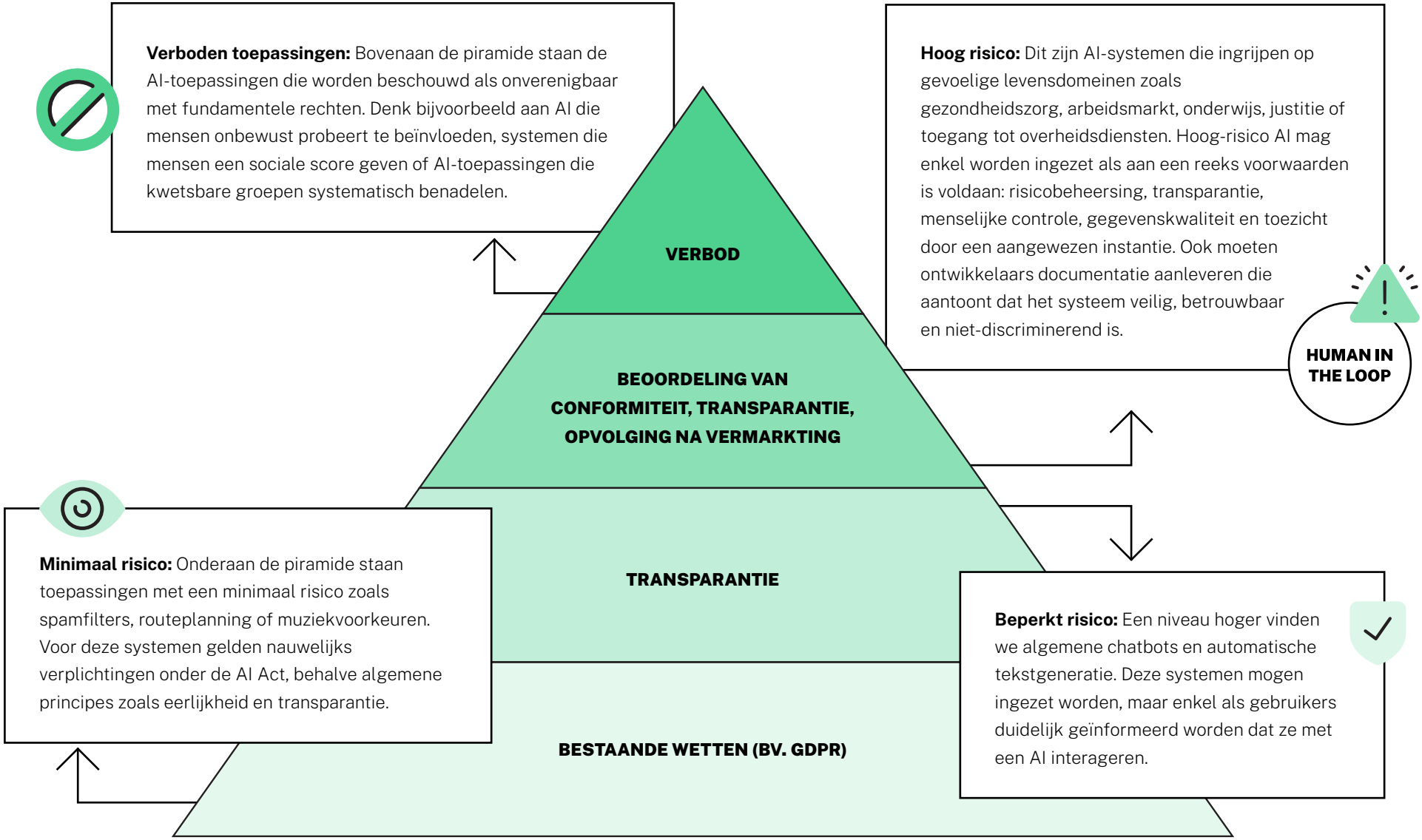


Wat betekent de EU AI ACT voor welzijn en zorg?

De tools in het vorige deel tonen dat AI op heel wat manieren een meerwaarde kan bieden voor welzijn en zorg. Tegelijk roept die brede inzet ook belangrijke vragen op. Wanneer mag AI worden gebruikt? Welke verantwoordelijkheid draagt een organisatie? Welke toepassingen zijn aanvaardbaar en welke niet? En welke rol blijft de mens spelen wanneer AI wordt ingezet? Die vragen zijn niet alleen ethisch, maar ook juridisch. Daarom ontwikkelde de Europese Unie een wettelijk kader voor AI. De EU AI Act bepaalt welke regels gelden voor verschillende soorten AI-toepassingen.

De EU AI Act is de eerste uitgebreide wettelijke regeling voor artificiële intelligentie binnen de Europese Unie. Deze wetgeving hanteert een risicogebaseerde benadering om AI-systemen te classificeren en te reguleren. De Europese wetgeving deelt AI-systemen in volgens het risico dat ze kunnen vormen voor mensen.




Verboden toepassingen: Bovenaan de piramide staan de AI-toepassingen die worden beschouwd als onverenigbaar met fundamentele rechten. Denk bijvoorbeeld aan AI die mensen onbewust probeert te beïnvloeden, systemen die mensen een sociale score geven of AI-toepassingen die kwetsbare groepen systematisch benadelen.
VERBOD
Hoog risico: Dit zijn AI-systemen die ingrijpen op gevoelige levensdomeinen zoals gezondheidszorg, arbeidsmarkt, onderwijs, justitie of toegang tot overheidsdiensten. Hoog-risico AI mag enkel worden ingezet als aan een reeks voorwaarden is voldaan: risicobeheersing, transparantie, menselijke controle, gegevenskwaliteit en toezicht door een aangewezen instantie. Ook moeten ontwikkelaars documentatie aanleveren die aantoont dat het systeem veilig, betrouwbaar en niet-discriminerend is.
HUMAN IN THE LOOP
BEOORDELING VAN CONFORMITEIT, TRANSPARANTIE, OPVOLGING NA VERMARKTING
TRANSPARANTIE
Minimaal risico: Onderaan de piramide staan toepassingen met een minimaal risico zoals spamfilters, routeplanning of muziekvoorkeuren. Voor deze systemen gelden nauwelijks verplichtingen onder de AI Act, behalve algemene principes zoals eerlijkheid en transparantie.
Beperkt risico: Een niveau hoger vinden we algemene chatbots en automatische tekstgeneratie. Deze systemen mogen ingezet worden, maar enkel als gebruikers duidelijk geïnformeerd worden dat ze met een AI interageren.
BESTAANDE WETTEN (BV. GDPR)

EU AI Act

Een belangrijk uitgangspunt van de EU AI Act is dat het risico wordt bepaald door de *bedoelde* toepassing van een AI-systeem en gaat dus niet over het effectieve *gebruik*. Als ChatGPT in de feiten als een gezondheidsadviseur of therapeut wordt gebruikt door sommige gebruikers, dan blijft toch de algemene doelstelling van OpenAI de basis om de risico-inschatting op te baseren.

De piramide is een juridisch bindend instrument dat bepaalt wat er mag, wat gereguleerd wordt en wat verboden is. Voor organisaties in welzijn en zorg is het belangrijk om tijdig te bepalen waar hun AI-toepassing zich in deze piramide bevindt, omdat dit gevolgen heeft voor ontwerp, toezicht en aansprakelijkheid. Doe die inschatting niet zelf, maar laat je adviseren door je Data Protection Officer (DPO) of juridische experten.

Een van de belangrijkste principes van de EU AI Act is de verplichting van **menselijk toezicht** wanneer AI een grote impact kan hebben op het leven van mensen. De mate van menselijk toezicht kan variëren. In sommige gevallen moet een mens een beslissing expliciet goedkeuren vóórdat deze wordt uitgevoerd (*human-in-the-loop*). Een mens moet dan op een betekenisvolle manier controle houden over beslissingen die door een AI-systeem worden voorgesteld. Het gaat hier niet om passief toezicht achteraf, maar om actieve betrokkenheid op beslissingsmomenten die gevolgen hebben voor mensenrechten, veiligheid of welzijn. In andere gevallen moet de mens de mogelijkheid hebben om een beslissing achteraf te corrigeren (*human-on-the-loop* of *human-in-command*).

In welzijns-en zorgcontexten betekent dit concreet dat een AI-chatbot die *gespecialiseerd* is in psychosociale thema's altijd gekoppeld moet zijn aan een proactief mechanisme van menselijk toezicht: *human-in-the-loop*. Een chatbot mag dan niet autonoom werken: de mens is hier niet optioneel, maar fundamenteel voor het ethisch en wettelijk verantwoord functioneren van het AI-systeem. Bij gevoelige thema's zoals suïcidaliteit is het niet voldoende dat een mens pas tussenkomt wanneer zich een probleem voordoet: de chatbot moet zo ontworpen zijn dat hij zijn eigen grenzen herkent, risicosignalen detecteert en tijdig doorverwijst naar een mens voordat de situatie escaleert.

Organisaties in welzijn en zorg moeten daarom duidelijke afspraken maken over het menselijk toezicht als ze zelf AI-systemen ter beschikking stellen: wie houdt toezicht, met welke bevoegdheden, en met welk niveau van AI-geletterdheid? Dat vraagt niet alleen voldoende opleiding, maar ook heldere procedures. Denk aan werken met een systeem van vlaggensystemen rond risico-onderwerpen, stop-knoppen, rapportagekanalen, waarschuwingen, uitleg en menselijke escalatiemogelijkheden. Dit alles staat nog in de kinderschoenen.

Betrek dus zeker je juridische dienst en informeer je grondig. Meer informatie?

- Een website van Europa erover: **https://ai-act-service-desk.ec.europa.eu/en**
- Een website van België erover: **https://economie.fgov.be/nl/themas/ondernemingen/ai-act**

Hoe kan je een taalmodel zelf trainen?



We hebben nu al wat meer zicht op wat een chatbot in het algemeen is en wat de kansen en risico's zijn. Die kansen kunnen versterkt worden, en de risico's verkleind, als een chatbot meer specifiek getraind wordt op gespecialiseerde taken. Er zijn meerdere manieren om een chatbot te trainen of meer op maat te maken. We bespreken er hier drie: system prompting, Retrieval-Augmented Generation (RAG) en finetuning. In deel 3 worden ze ook uitgetest.

Je kan een algemene chatbot opdrachten geven door een duidelijke vraag te stellen of te beschrijven wat je wil. Dat heet **prompting**. Bijvoorbeeld: "Schrijf een uitleg voor een jongere van 14 over wat stress is." Alles wat je schrijft, is input voor de chatbot. Die prompt zet het taalmodel aan het werk om een geloofwaardig antwoord te genereren op basis van kansberekening. Hoe duidelijker en relevanter je instructie en context zijn, hoe groter de kans op een bruikbaar antwoord. De input die je geeft, werkt dus als een set instructies of richtingaanwijzers voor de chatbot. Dit is op zich geen training, maar maakt wel deel uit van goed gebruik van AI.

Naast gewone prompting bestaat er ook **system prompting**. Dat gebeurt meer op het niveau van de chatbot zelf. Een chatbot krijgt dan algemene instructies mee over zijn rol, toon, werkwijze en grenzen. Via system prompting kan je bepalen dat de chatbot vriendelijk en helder antwoordt, geen diagnoses stelt, altijd verwijst naar menselijke hulp bij crisissignalen of een stappenplan volgt. Dit soort instructies bepaalt het gedrag van de chatbot als geheel. Deze instructies kan je eenvoudigweg vastleggen in woorden, zonder gebruik van programmeertaal of complexe systemen.

Gebruik daarbij '*few-shot prompting*'. Naast instructies krijgt een taalmodel best ook enkele goede voorbeelden ("a few shots") mee om te tonen welk soort antwoord gewenst is. Zo leert de chatbot niet alleen wat hij moet doen, maar ook hoe dat moet klinken. Few-shot prompting werkt goed en is efficiënt. Bij zero-shot prompting geef je enkel instructies en geen voorbeelden. Bij one-shot prompting komt er één voorbeeld bij. Bij many-shot prompting komen er zeer veel voorbeelden bij.

Een tweede techniek bestaat uit **RAG**. Dit staat voor *Retrieval-Augmented Generation*. Dat betekent zoiets als 'tekst genereren op basis van informatie die eerst moet worden opgehaald'. Met RAG geef je een taalmodel toegang tot specifieke informatiebronnen, zoals documenten, websites, interne databanken, protocollen of handleidingen. Wanneer je een vraag stelt, zoekt het systeem eerst relevante stukjes informatie in die bronnen en formuleert het daarna een antwoord op basis van die gevonden informatie. Het taalmodel en de externe kennisbronnen werken dus samen, maar blijven apart: het taalmodel zorgt voor de talige formulering, de kennisbron levert de inhoudelijke basis.

Een voorbeeld: een hulpverlener in een CAW wil weten hoe hij moet omgaan met vermoedens van kindermishandeling, maar kent het interne protocol niet van buiten. Hij kan dit zelf opzoeken in de handleiding, maar met een RAG-oplossing kan een chatbot helpen om de relevante passages sneller terug te vinden en samen te vatten. De organisatie voegt dan documenten toe aan hun interne kennischatbot en geeft het instructies via system prompting. De chatbot zoekt vervolgens in de richtlijnen en formuleert een inhoudelijk antwoord. De chatbot vervangt daarbij niet de professionele afweging of het overleg volgens de interne afspraken. Hij helpt vooral bij het terugvinden en structureren van informatie.

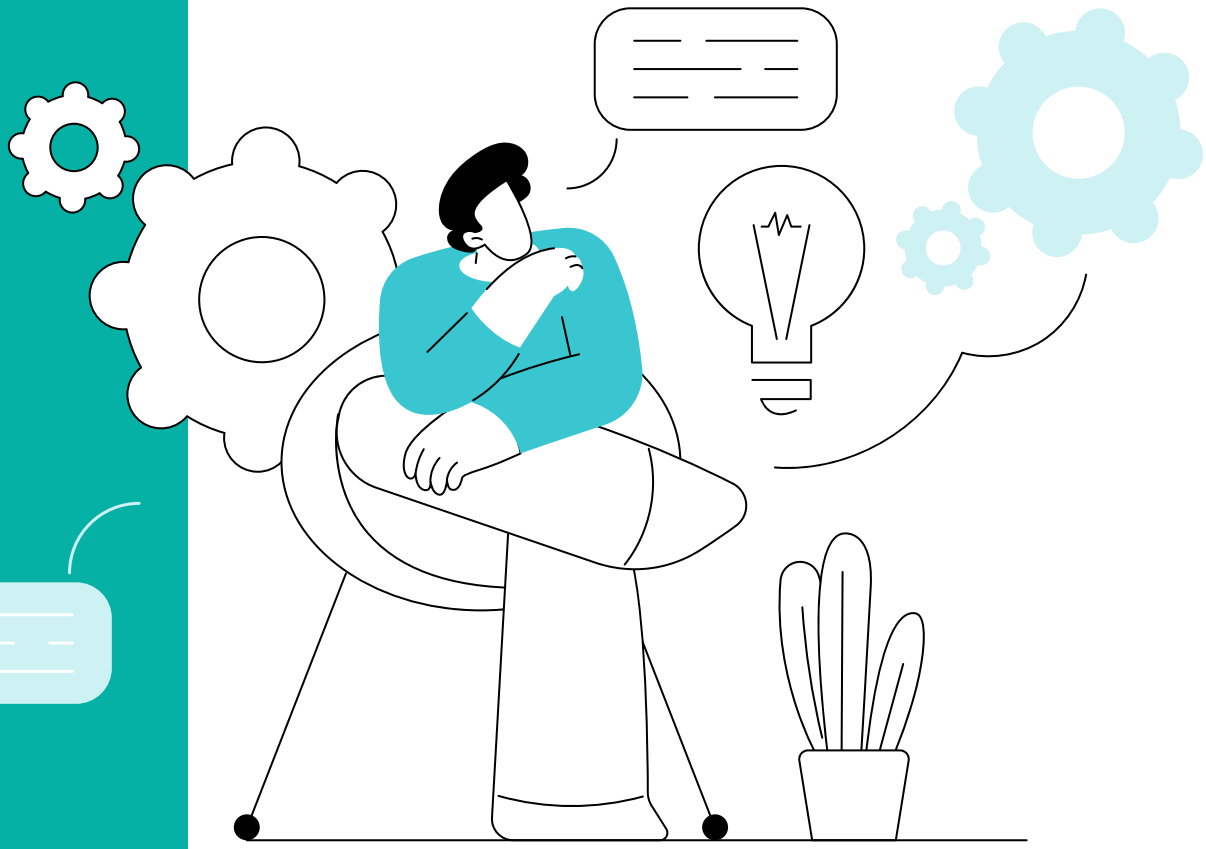
Een derde techniek is complexer. **Finetuning** is het proces waarbij je een bestaand taalmodel op zich verder traint met vele extra voorbeelden of teksten. Daarbij wordt het taalmodel zelf dus een beetje aangepast. Je leert het model als het ware extra patronen aan, zodat het kan reageren met een meer gepaste stijl. Finetuning is vooral nuttig wanneer je wil dat een model consequent op een bepaalde manier antwoordt, bijvoorbeeld in de toon van een organisatie, volgens een bepaald gespreksmodel of op basis van een groot aantal voorbeelden van goede hulpverleningsgesprekken. Voor veranderlijke of feitelijke informatie, zoals procedures, openingsuren, wetgeving of interne afspraken, is RAG geschikter.

Je kan het **verschil** tussen RAG en finetuning begrijpen met de vergelijking tussen twee soorten examens. Bij finetuning geef je het model vooraf extra les en laat je het daarna antwoorden zonder dat het nog documenten mag inkijken: een geslotenboekexamen. Het model is vooraf bijgestuurd op basis van voorbeelden en gebruikt die aangeleerde patronen in haar antwoorden. Bij RAG hoeft het model niet alles op voorhand te leren, maar mag het, zoals bij een openboekexamen, informatie opzoeken in externe bronnen, zoals een handboek, website of interne kennisbank. Het systeem haalt eerst relevante stukjes informatie op uit die bron en gebruikt die context om een antwoord te formuleren. Finetuning verandert vooral het gedrag van het taalmodel zelf, RAG geeft het model toegang tot specifieke kennisbronnen en verandert niets aan het taalmodel.

In deel 3 vergelijken we deze technieken door prototypes van getrainde chatbots te maken.

Ben je klaar voor AI met je team?

Een organisatie in welzijn of zorg heeft veel inhoudelijke expertise in huis. Die expertise is essentieel en kan de leidraad vormen in het nemen van beslissingen en het opzetten van een AI-beleid. Technologie vormt maar een middel. Medewerkers hoeven dus niet allen AI-specialist te worden maar het is wel belangrijk dat ze begrijpen wat AI wel en niet kan. Alleen dan kunnen ze de mogelijkheden en beperkingen ervan kritisch inschatten. Onderstaande checklist kan je helpen om in te schatten hoe ver jullie als organisatie staan op het gebied van AI.



CHECKLIST: Begrijpen we AI genoeg om de volgende stap te zetten?

Begrijpen we hoe conversationele AI werkt?

- ☑ Beseffen alle medewerkers dat een chatbot geen zoekmachine is, maar steeds nieuwe tekst genereert?
- ☑ Beseffen we dat de kwaliteit van de output sterk afhangt van de prompt?
- ☑ Is het voldoende duidelijk voor iedereen dat een chatbot kansberekening gebruikt om woorden te voorspellen? AI kan dus overtuigend klinken zonder noodzakelijk correct te zijn.
- ☑ Weten we hoe een chatbot kan worden getraind via system prompting, RAG en finetuning?

Vallen onze eigen AI-tools onder de EU AI ACT?

- ☑ Weten we onder welke categorie van de EU AI Act onze eigen AI-toepassingen vallen? Is het duidelijk wanneer AI wel en niet mag worden ingezet?
- ☑ Hebben we nagedacht over de vereisten rond menselijk toezicht?
- ☑ Zijn de verantwoordelijkheden binnen de organisatie duidelijk verdeeld?

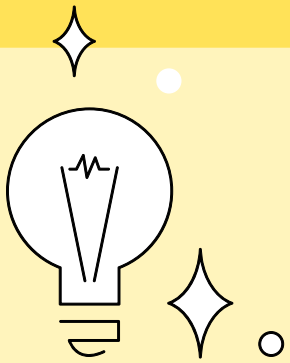
Zijn we ons bewust van de risico's?

- ☑ Weten de medewerkers welke risico's AI met zich meebrengt voor privacy, betrouwbaarheid en transparantie? Beseffen we dat AI empathie kan simuleren zonder de situatie werkelijk te begrijpen?
- ☑ Kunnen we als organisatie inschatten wanneer AI wel of niet geschikt is voor een toepassing?

Zetten we AI al verantwoord in?

- ☑ Is er binnen de organisatie afgesproken hoe AI veilig gebruikt wordt?
- ☑ Wordt het invoeren van persoonsgegevens of vertrouwelijke informatie vermeden?
- ☑ Worden AI-antwoorden altijd op juistheid en betrouwbaarheid gecontroleerd? Blijft de eindverantwoordelijkheid altijd bij een mens?

Samengevat



Conversationele AI is geen zoekmachine: het haalt geen bestaande teksten op, maar genereert nieuwe antwoorden. Het maakt steeds nieuwe tekst. Het kopieert niet simpelweg van het internet. AI kan daardoor informatie verzinnen of verouderde informatie geven. Een AI-systeem weet niet wat waar is of wat niet. Het kent de vorm van taal goed, maar het kent geen feiten of 'waarheid'. Wat je erin steekt, bepaalt voor een groot stuk wat eruit komt: garbage in = garbage out.

De grootste risico's zitten niet alleen in fouten, maar ook in de schijnzekerheid. AI kan empathie, nabijheid en deskundigheid nabootsen, maar draagt geen verantwoordelijkheid en herkent kwetsbaarheid niet zoals een professional. Daarom zijn transparantie, kritische controle en duidelijke grenzen essentieel.

AI kan waardevol zijn, vooral als hulpmiddel. AI kan helpen bij taal, structuur, toegankelijkheid, voorbereiding, administratie, opleiding en kennisdeling. Maar bij crisis, diagnostiek of beslissingen met impact op rechten, welzijn of gezondheid blijft menselijke professionele beoordeling (wettelijk) noodzakelijk.

Organisaties moeten vooraf nagaan onder welke risicocategorie van de EU AI Act een eigen AI-systeem valt, welke regels omtrent menselijke controle van toepassing zijn en of de toepassing binnen een welzijns-of zorgcontext aanvaardbaar is. Die beoordeling gebeurt best samen met een DPO of juridisch expert en inhoudelijke professionals. Hoe groter de mogelijke impact op rechten, gezondheid of welzijn, hoe strenger de voorwaarden en hoe betekenisvoller het menselijk toezicht moet zijn.

Er zijn diverse technieken om een taalmodel te trainen en meer op maat te maken van je doelstellingen als organisatie. System prompting, RAG en finetuning bieden elk een andere manier om een taalmodel af te stemmen op een specifieke toepassing. Met system prompting leg je de rol, toon, werkwijze en grenzen van de chatbot vast, eventueel ondersteund met enkele goede voorbeelden via few-shot prompting. RAG koppelt het taalmodel aan betrouwbare en actuele informatiebronnen, zodat antwoorden kunnen worden gebaseerd op documenten, protocollen of interne kennis. Finetuning past het taalmodel zelf aan op basis van veel voorbeelden en is vooral geschikt om een consequente stijl aan te leren.

2

Voorbereiden

Maak organisatiekeuzes en bied beleidskaders

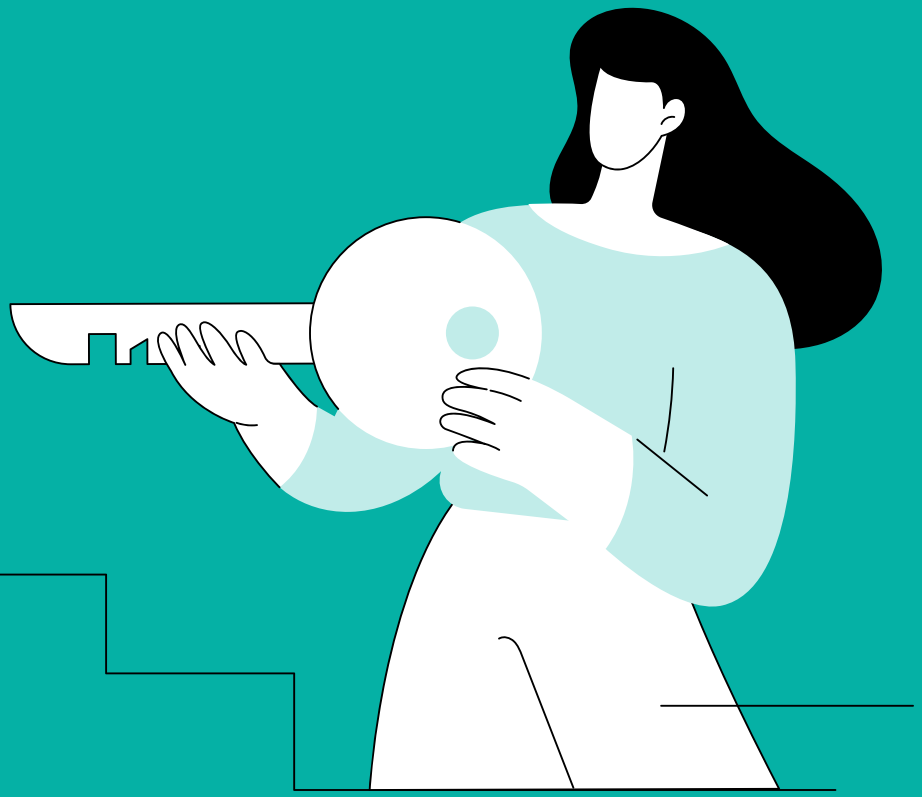
- Bouw een organisatievisie op het professioneel gebruik van AI
- Overweeg je een eigen chatbot? Vul de AI Data Readiness Checklist in
- Overweeg je een chatbot in te zetten? Pas de kwaliteitscriteria toe
- Samenvatting

Leeswijzer

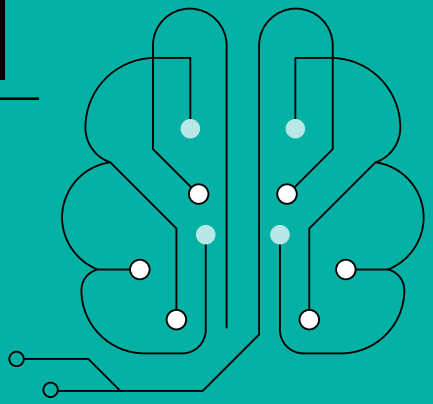
Dit deel helpt je organisatie om van losse, informele AI-experimenten naar duidelijke afspraken te gaan. Zonder organisatiekader gebeurt dat vaak in de schaduw. Dan is niet duidelijk welke tools gebruikt worden, welke gegevens worden ingevoerd en hoe output gecontroleerd wordt.

Lees dit deel wanneer je als organisatie wil bepalen welke rol AI (niet) mag spelen in je werking. Het gaat niet alleen over regels, maar over visie: waarom willen we AI gebruiken, voor wie moet het meerwaarde hebben, welke toepassingen vinden we verantwoord en welke niet? Dit deel helpt om AI te verbinden met kwaliteit, privacy en verantwoordelijkheid.

Dit deel is vooral nuttig voor directie, beleidsmedewerkers, stafmedewerkers en juridische, HR-, kwaliteitszorg-en ICT-diensten. Het helpt om een AI-kader te maken met basisregels over toegelaten tools, privacygevoelige gegevens, transparantie, kwaliteitscontrole en opvolging.



Bouw een organisatie-visie op het professioneel gebruik van AI



De opdracht bestaat eruit om een AI-organisatiekader uit te schrijven waarbij je de krijtlijnen van professioneel gebruik van AI-systemen verduidelijkt voor iedereen. Voorzie dus een dynamisch kader waarbij verduidelijkt wordt wat wel mag en wat niet en onder welke voorwaarden. Maak hierbij het onderscheid tussen het professioneel gebruik van bestaande AI-tools of van eigen AI-gespecialiseerde tools die je zelf traint.

Wanneer een organisatie geen duidelijke afspraken maakt, ontstaat al snel schaduwgebruik: AI wordt in de feiten gebruikt, maar het is niet zichtbaar en geen voorwerp van gesprek of beleid. Een AI-kader is niet bedoeld om het gebruik van AI te verbieden. Integendeel: een organisatiekader helpt om AI uit de informele sfeer te halen en bespreekbaar te maken. Door expliciet te bepalen wat (niet) kan, ontstaat er ruimte om AI zinvol in te zetten voor taken. Tegelijk worden duidelijke grenzen getrokken.

Enkele **voorbeelden** van bestaande AI-kaders in welzijn en zorg kan je vinden op de website van SAM vzw: **www.samvzw.be/kennis/AI-in-zorg-en-welzijn**. In dit dossier staan AI-kaders van CLB online, CKG De Hummeltjes, CAW Zuid-West Vlaanderen en SAM vzw. Dit zijn versies van begin 2026 en zullen veranderlijk zijn, maar kunnen inspiratie bieden.

Ga ermee aan de slag! Via dit **stappenplan** kan je als organisatie een AI-kader opmaken:

Fase 1: In kaart brengen huidig (informeel) gebruik

Een eerste stap is het erkennen van de realiteit. Het is een goede eerste stap om binnen je organisatie **zicht te krijgen op het formeel en informeel gebruik van AI** en zo tot bewustwording (of zelfs een gevoel van urgentie) te komen, zodat de stap naar een organisatiebeleid logisch wordt. Zet een bevraging op, organiseer een workshop of voorzie een opleiding waarbij ook het huidig gebruik in kaart wordt gebracht. Ga daarbij op zoek naar een antwoord op de vraag: welke AI-tools worden voor welke professionele taken al gebruikt? Welke tools zouden ze willen gebruiken? Welke niet? Dit zal veel inzicht geven in de mogelijkheden en de onderliggende noden. Ook de bewustwording van de nood aan een AI-kader zal groeien.



Fase 2: Begrijpen en positioneren. Visie en beleidskader uitwerken

Na de bewustwording volgt de ontwikkeling van een gedragen visie en bijhorend beleid. Die visie vormt het fundament voor alle verdere keuzes. Ze maakt duidelijk waarom de organisatie AI wil inzetten, waarvoor AI gebruikt mag worden en welke grenzen daarbij gelden.

Om de visie te koppelen aan bestaande regelgeving zoals de GDPR en de EU AI Act kan een multidisciplinair **kernteam** worden samengesteld. Dit team combineert praktijkkennis, organisatiemanagement met expertise in IT, juridische aspecten en ethiek, en kan indien nodig aangevuld worden met externe begeleiding. Zo'n visie moet een antwoord geven op deze kernvragen:

- Waarom wil onze organisatie AI inzetten? Welke doelstellingen willen we met AI bereiken? Voor wie (intern of extern)? Welke meerwaarde moet AI bieden voor doelgroep, medewerkers en/of de organisatie?
- Welke AI-tools zijn toegelaten? Voor welke taken mag AI worden gebruikt? Voor welke taken is AI verboden of beperkt?
- Welke gegevens mogen wel of niet worden ingevoerd?
- Hoe communiceren we over ons AI-gebruik?

De antwoorden op deze vragen vormen samen het **AI-beleidskader** van je organisatie. De visie beschrijft *waarom* en *vanuit welke waarden* AI wordt ingezet. Het beleidskader biedt concrete afspraken en verantwoordelijkheden die medewerkers houvast bieden in de dagelijkse praktijk. Wacht niet op perfecte richtlijnen, maar test snel een bruikbare eerste versie samen met de AI-tool. Zorg ook voor goedkeuring en voldoende AI-kennis op directie-en bestuursniveau.

Fase 3: Implementatie van beleid en tools

Selecteer de AI-tools die je op organisatieniveau wil inzetten zorgvuldig. Staar je niet blind op snelle beloftes van efficiëntie, maar kijk ruimer. Naast kostprijs zijn eigenschappen zoals dataveiligheid, privacy en data-en technologiesoevereiniteit ook belangrijk (zo is een voorkeur voor software uit de EU steeds belangrijker, maar niet eenvoudig waar te maken). Maak onderscheid tussen het gebruik van bestaande AI-tools (zoals Copilot-licenties voor iedereen) en het zelf inzetten van aangepaste AI-tools op organisatieniveau (zoals een chatbot voor de doelgroep). Als je zelf een AI-toepassing ontwikkelt of integreert in je werking, dan kruip je in de rol van producent en zijn er extra verantwoordelijkheden.

Jouw AI-beleid heeft alleen effect wanneer het effectief geïntegreerd wordt in de dagelijkse werking. Gebruik daarom begrijpelijke taal en concrete voorbeelden die aansluiten bij de praktijk van medewerkers. Voorzie ook **praktische ondersteuning**. Dit kan door richtlijnen te ontwikkelen rond goed gebruik van AI, zoals hoe veilig te prompten, welke informatie wel of niet ingevoerd mag worden, en hoe output kritisch geëvalueerd wordt. Daarnaast kunnen prompt-sjablonen, checklists en beslissingskaders voorzien worden. Voorzie **opleidingen** over AI-basiskennis, prompting en risico-inschatting en creëer een duidelijk aanspreekpunt voor vragen of moeilijkheden.

Fase 4: Monitoring en bijsturing

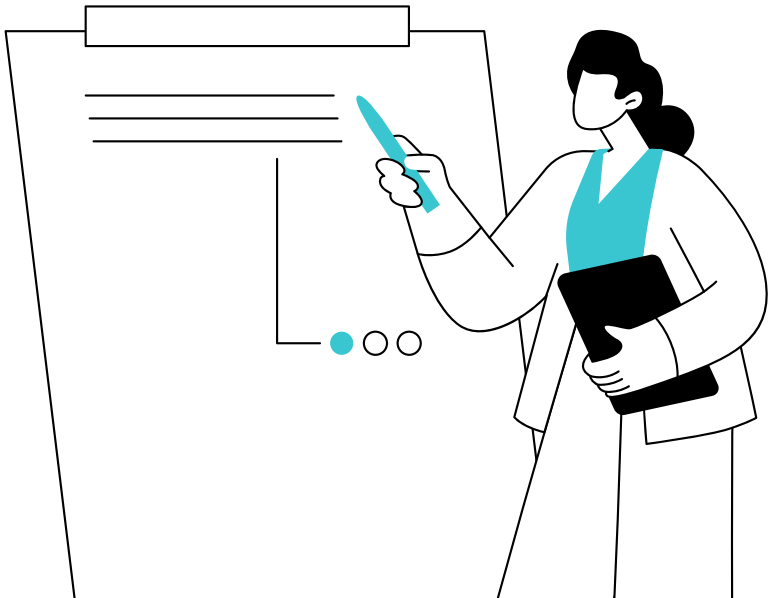
Je AI-beleid is nooit af. Technologie en haar regelgeving bewegen snel. Daarom is het belangrijk om het beleid regelmatig te evalueren en waar nodig bij te sturen. Zo blijft het beleid aansluiten bij de noden van medewerkers, de verwachtingen van gebruikers en nieuwe technologische ontwikkelingen. Laat je daarbij informeren door betrouwbare kennispartners zoals Vlaams AI Academie VAIA (**www.vaia.be**), het Kenniscentrum Data en Maatschappij (**https://data-en-maatschappij.ai/**) of een kennisinstelling zoals Arteveldehogeschool, Thomas More-hogeschool of Universiteit Gent. Gebruik tools zoals Ally: **https://data-en-maatschappij.ai/tools/ally**.

Verlies daarbij niet uit het oog dat succesvolle AI-implementatie in de eerste plaats een organisatievraagstuk is en pas daarna een technologische kwestie. De grootste uitdaging is niet de keuze van de juiste AI-tool, maar de manier waarop medewerkers ermee leren werken en hoe AI wordt ingebed in bestaande processen. Investeren in vorming, begeleiding, draagvlak en duidelijke werkafspraken levert daarom vaak meer op dan investeren in steeds nieuwe technologie.

CHECKLIST:
Kernelementen van een organisatiekader voor professioneel AI-gebruik

In het uitvoeren van de tweede fase, kan je wellicht wat ondersteuning gebruiken in het uitdenken van een visie en bijhorende beleidskader. Hier geven we alvast wat elementen mee. Deze elementen zijn gebaseerd op bestaande AI-kaders in welzijn en zorg.

Er bestaat geen universeel AI-kader voor organisaties in welzijn en zorg. Er moeten keuzes worden gemaakt en die keuzes moeten passen binnen je eigen context en kunnen ook doorheen de tijd veranderen. Organisaties kunnen per element bepalen hoe streng of ruim ze dit willen invullen. Jullie zijn dus aan zet! We geven hier wat elementen mee die deel kunnen uitmaken van je beleid.



Wie draagt de verantwoordelijkheid?

Op zich is het duidelijk: de organisatie draagt altijd de eindverantwoordelijkheid. AI is enkel een hulpmiddel. De verantwoordelijkheid ligt steeds bij de mensen die het gebruiken, niet bij het algoritme. Daarom is het belangrijk om AI-output te controleren op onder meer bevooroordeelde formuleringen, veralgemeningen, juistheid, volledigheid en betrouwbaarheid.

Dat is niet altijd eenvoudig: het kan behoorlijk wat tijd en werk vragen, maar het is essentieel. De vraag is hoe je dit concreet aanpakt.

Ondertussen herkennen ervaren lezers ook al snel als een tekst te veel door AI is geschreven. Mag dit?

KIES ALS ORGANISATIE

- ☑ Je bent als professional altijd de eindverantwoordelijke.
- ☑ Elke AI-output moet worden gecontroleerd op juistheid en betrouwbaarheid. Controleer bij feitelijke info altijd de bronnen zelf.
- ☑ AI-output mag nooit rechtstreeks gebruikt worden.
- ☑ AI-output moet altijd vergeleken worden met eigen werk. Er is dus altijd een eigen versie en een AI versie. Je mag dan de AI-output gebruiken als bron van inspiratie, maar niet als eindproduct. Je herschrijft dus alles.
- ☑ Gebruik AI bewust en gericht, wees steeds voorzichtig en blijf zelf aan het stuur. Maak het je eigen: je moet jezelf er terug in herkennen. Breng de mens terug in de tekst.

Welke AI-tools laten we toe?

Organisaties kunnen kiezen voor een strikte of meer experimentele lijn. Je kan bijvoorbeeld kiezen om enkel Microsoft Copilot met ondernemingsbescherming toe te laten. Andere AI-tools zijn dan niet toegelaten. Je kunt het ruimer bekijken en gecontroleerde AI-omgevingen toelaten wanneer er onder meer een verwerkersovereenkomst en garanties rond gegevensbescherming zijn. De publieke AI-systemen zijn dan niet toegelaten, omdat er dan te weinig garanties rond privacy zijn.

KIES ALS ORGANISATIE

- ☑ Er is slechts één goedgekeurde AI-tool en deze wordt aangeboden door de organisatie.
- ☑ Publieke of algemene AI-tools mogen enkel voor taken zonder gevoelige data.
- ☑ Nieuwe AI-tools mogen alleen worden gebruikt na goedkeuring.

KIES ALS ORGANISATIE

- ☑ Je mag bestaande teksten herschrijven, samenvatten, vertalen of vereenvoudigen, maar geen nieuwe teksten genereren.
- ☑ Je mag AI niet gebruiken om gegevens te analyseren.
- ☑ Je mag geen beslissingen laten nemen door AI.

Waarvoor mag AI wel gebruikt worden?

AI kan op allerhande manieren worden ingezet, zoals samenvatten van publiek beschikbare documenten, brainstormen, e-mails opstellen, publieke communicatie voorbereiden, achtergrondinformatie zoeken, vertalen, teksten aanpassen aan leeftijd, feedbackgegevens analyseren ...

Ook bij deze toepassingen blijft de medewerker eindverantwoordelijk. Beschouw de output van AI als een half afgewerkt product, waarbij jij zelf de eindredacteur blijft met de taak om eigen kennis, ervaring en expertise toe te voegen. Je bent als medewerker hoe dan ook volledig verantwoordelijk voor de kwaliteit en juistheid van het eindresultaat.

Welke gegevens mogen medewerkers invoeren?

Op zich is het duidelijk: gebruik alleen eigen of publiek beschikbaar materiaal zonder identificeerbare, vertrouwelijke of privacygevoelige informatie. Voer in publieke AI-tools geen namen, adressen, medische gegevens, dossierinformatie, testresultaten of chatgesprekken in. Ook zonder naam kan iemand herkenbaar blijven door een combinatie van kenmerken. Zelf anonimiseren is daarom noodzakelijk; enkel namen verwijderen is meestal slechts pseudonimiseren. Is dit voldoende?

Als je toch identificeerbare gegevens wil invoeren, dan vereist dit expliciete toestemming van de betrokkenen en mag dit hun rechten, gezondheid of welzijn niet schaden. Respecteer ook auteursrechten op documenten.

De veiligste basisregel blijft: voer geen persoonsgegevens, casus-informatie of vertrouwelijke documenten in publieke AI-tools in.

- ☑ Enkel publieke of goedgekeurde documenten als input voorzien.
- ☑ Geen vertrouwelijke documenten uploaden in publieke tools.
- ☑ Geen documenten invoegen waarop auteursrechten en privacyregelingen van toepassing zijn: identificeerbare informatie, tenzij daar actief toestemming voor is verkregen
- ☑ Geen persoonsgegevens, cliënt-of dossierinformatie invoeren.
- ☑ Privacygevoelige gegevens enkel in gecontroleerde AI-omgeving, rekening houdend met de GDPR.

KIES ALS ORGANISATIE



Overweeg je een eigen chatbot? Vul de AI Data Readiness Checklist in

Controleer of je klaar bent om met je eigen data aan de slag te gaan

Verantwoord gebruik van AI begint niet bij de technologie alleen. Het begint bij een goed zicht op je data, duidelijke rollen, heldere afspraken en aandacht voor privacy, veiligheid en verantwoordelijkheid.

Daarom werd de *AI Data Readiness Checklist* ontwikkeld. Deze checklist helpt organisaties om na te gaan of de basisvoorwaarden aanwezig zijn om verantwoord met data en AI aan de slag te gaan. Ze richt zich niet op één specifieke AI-toepassing, maar op de voorbereiding die nodig is vóór je een AI-project opstart of verder verkent.

De checklist helpt je onder meer om stil te staan bij vragen zoals: welke data hebben we, waar zit die data, mogen we die gebruiken, wie moet betrokken worden, hoe beschermen we persoonsgegevens en welke afspraken zijn nodig met externe partners?

Je hoeft geen juridische, technische of AI-expert te zijn om met de checklist aan de slag te gaan. Ze is bedoeld als praktisch hulpmiddel om het gesprek binnen je organisatie op gang te brengen tussen medewerkers met ervaring rond praktijk, beleid, kwaliteit en ICT.

Gebruik de checklist dus wanneer je organisatie een concrete AI-piste wil verkennen. Ze helpt om zichtbaar te maken waar je vandaag staat, welke vragen al beantwoord zijn, waar nog blinde vlekken zitten en welke acties nodig zijn voor je verdergaat.

De **AI Data Readiness Checklist** kan je in bijlage vinden en bestaat uit deze delen:

Doel en gebruik van de checklist

Voor het invullen van de checklist omschrijft een organisatie best eerst haar concrete nood. Daarbij wordt nagegaan welk probleem men wil oplossen, voor wie de toepassing bedoeld is en welke meerwaarde wordt verwacht. Vervolgens wordt beoordeeld of AI werkelijk noodzakelijk, proportioneel en verantwoord is, of dat een eenvoudigere of veiligere oplossing geschikter is.

Rollen en verantwoordelijkheden

Een chatbot uitbouwen vereist dat duidelijk is wie welke verantwoordelijkheid draagt. De checklist maakt onderscheid tussen verschillende rollen. De Data Protection Officer (DPO), jurist of privacyverantwoordelijke adviseert over gegevensbescherming, rechtsgronden en risico's. De data steward vormt de brug tussen praktijk, beleid, IT en privacy. De databeheerder kent de technische systemen, toegangsrechten en mogelijkheden om data te exporteren. Daarnaast moet duidelijk zijn wie het mandaat heeft om beslissingen te nemen over het opstarten, aanpassen of stopzetten van een AI-project.

Beschikbaarheid en kwaliteit van data

Voor het uitwerken van zo'n AI-project moet een organisatie weten welke gegevens beschikbaar zijn, waar ze worden bewaard, wie toegang heeft en in welke vorm ze kunnen worden gebruikt. Data kan bestaan uit gestructureerde tabellen, maar ook uit vrije tekst, zoals e-mails, verslagen, dossiers of chatgesprekken.

Daarnaast moet de kwaliteit van de data worden beoordeeld. Onvolledige, inconsistente, verouderde of onduidelijke gegevens kunnen leiden tot onbetrouwbare resultaten. Het feit dat data technisch beschikbaar of exporteerbaar is, betekent nog niet dat ze juridisch gebruikt *mag* worden. Een inventaris is vereist van de data met een omschrijving van de locatie, vorm, toegankelijkheid, betekenis en mogelijke kwaliteitsproblemen van de data.

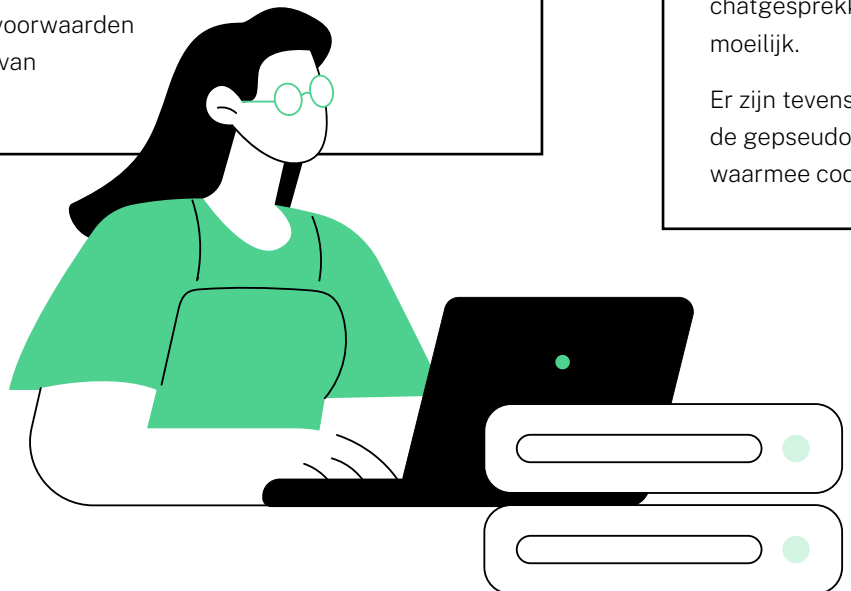


Juridische voorwaarden

Dit deel onderzoekt of de organisatie de data mag gebruiken voor het beoogde doel. Daarbij zijn twee vragen belangrijk. Ten eerste moet er een geldige *rechtsgrond* zijn voor de verwerking van persoonsgegevens. Ten tweede moet de organisatie beschikken over de nodige gebruiksrechten, licenties en contractuele *toestemming* om de data, databank of inhoud voor AI te gebruiken.

Data die oorspronkelijk voor hulpverlening of dienstverlening werd verzameld mag niet zomaar voor onderzoek, kwaliteitsverbetering of AI-ontwikkeling worden hergebruikt. De gegeven toestemming voor een gesprek betekent niet vanzelf dat het voor een AI-project mag worden hergebruikt. De organisatie moet nagaan wat betrokkenen werd meegedeeld, wat in de privacyverklaring staat en of een nieuwe juridische beoordeling nodig is.

Daarnaast kunnen auteursrechten, databankrechten, licentievoorwaarden of contracten het gebruik van gegevens beperken.



Anonimisering en pseudonimisering van gevoelige data

Data in welzijn en zorg heeft vaak betrekking op mensen in kwetsbare situaties. Daarom moet worden nagegaan of personen rechtstreeks of onrechtstreeks herkenbaar blijven. Alleen namen en contactgegevens verwijderen is meestal onvoldoende. Ook combinaties van leeftijd, woonplaats, school, gezinssituatie of bijzondere gebeurtenissen kunnen iemand (theoretisch gezien) identificeerbaar maken.

Er wordt daarom een onderscheid gemaakt tussen persoonsgegevens, gepseudonimiseerde gegevens en anonieme gegevens. Gepseudonimiseerde data kan (theoretisch) nog opnieuw aan een persoon worden gekoppeld en blijft daarom onder de privacywetgeving (GDPR) vallen. Anonieme data mag niet meer herleidbaar zijn tot een individu. Vooral bij vrije tekst, zoals chatgesprekken en cliëntverslagen, is volledige anonimisering erg moeilijk.

Er zijn tevens afspraken nodig over toegang tot oorspronkelijke data, de gepseudonimiseerde bestanden en de eventuele sleutelbestanden waarmee codes opnieuw aan personen gekoppeld kunnen worden.

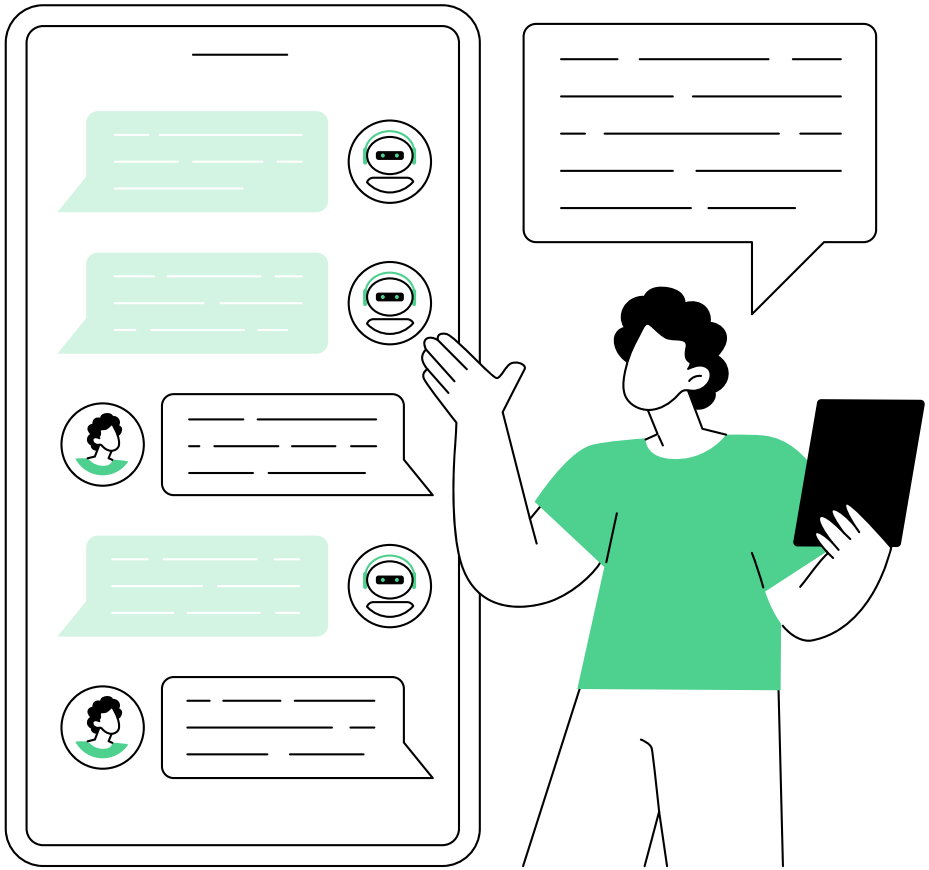
Data delen met externe partners

Wanneer een externe onderzoekspartner, AI-ontwikkelaar of technologiepartner toegang krijgt tot data, moet vooraf worden vastgelegd wat er met die data gebeurt. Data delen omvat niet alleen het doorsturen van bestanden, maar ook toegang geven tot platformen, werken in een cloudomgeving of data laten analyseren door een externe partij.

De rol van iedere betrokken partner moet juridisch worden bepaald. Een partij kan optreden als verwerker, gezamenlijke verwerkings-verantwoordelijke, zelfstandige verwerkingsverantwoordelijke of onderzoekspartner. Afhankelijk daarvan kan een verwerkers-overeenkomst, samenwerkingsovereenkomst of data sharing-of data transfer agreement (DTA) nodig zijn.

Een voorbeeld van een Data Transfer Agreement: zie bijlage.
https://www.arteveldehogeschool.be/ai-cares

Verder gelden de principes van dataminimalisatie en veilige overdracht. Alleen noodzakelijke gegevens mogen worden gedeeld, bij voorkeur geanonimiseerd of gepseudonimiseerd. Ook moeten afspraken worden gemaakt over beveiligde kanalen, toegangsrechten, logging, bewaartermijnen, verwijdering van data en procedures bij incidenten of datalekken.



Overweeg je een chatbot in te zetten? Pas de kwaliteits-criteria toe

Pas het kwaliteitskader digitale zorg, hulpverlening en ondersteuning toe

Heb je zicht op een concrete AI-tool om in te zetten? Wanneer je een specifieke chatbot in je organisatie wil inzetten komen er nog heel wat vragen bij kijken. Is de technologie betrouwbaar? Hoe wordt er omgegaan met privacy en veiligheid? Is de chatbot toegankelijk voor het publiek dat je wil bereiken? En wat moet je als organisatie doen om de chatbot kwaliteitsvol in te bedden in je werking?

Om jou hierbij te ondersteunen werd een aftoetsingsinstrument voor chatbots ontwikkeld. Je vindt het hier: **www.arteveldehogeschool.be/ai-cares**. Dit instrument helpt organisaties om na te gaan of en hoe een AI-chatbot op een kwaliteitsvolle manier kan worden ontwikkeld, geïmplementeerd en gebruikt. Gebruik het aftoetsingsinstrument wanneer je organisatie een chatbot wil verkennen, ontwikkelen, aankopen of inzetten. Het helpt om vooraf stil te staan bij belangrijke aandachtspunten, blinde vlekken zichtbaar te maken en gerichter in gesprek te gaan met collega's, leidinggevenden, IT, privacyverantwoordelijken, leveranciers of andere betrokken partners.

Het instrument bouwt verder op het *Kwaliteitskader digitale zorg, hulpverlening en ondersteuning* en het algemene aftoetsingsinstrument dat daarbij hoort: **www.departementzorg.be/nl/digitale-zorg-en-ondersteuning**. Het vervangt die instrumenten niet en wijkt er inhoudelijk niet van af. Het maakt de bestaande kwaliteitscriteria wel concreter voor de specifieke context van AI-chatbots, met toelichtingen en voorbeelden die helpen om de juiste vragen te stellen.

De bijlage in Excel vertaalt de algemene kwaliteitscriteria naar de specifieke kenmerken van conversationele AI. Die concretisering is nodig omdat een chatbot niet alleen informatie toont, maar zelfstandig taal genereert, reageert op persoonlijke input en rechtstreeks in interactie treedt met gebruikers. De criteria zijn ondergebracht in drie pijlers:

1. Binnen de pijler **technologie** gaat bijzondere aandacht naar de betrouwbaarheid van de aanbieder, de wetenschappelijke onderbouwing en actualiteit van de informatie, de naleving van onder meer de EU AI Act en de privacywetgeving en de technische beveiliging en robuustheid van het systeem. AI-risico's worden bevraagd, zoals foutieve of verzonnen antwoorden, vooringenomenheid en discriminatie, overdreven instemming met de gebruiker, ongewenste beïnvloeding, het onvoldoende herkennen van crisissituaties en foutieve doorverwijzingen. Ook transparantie over de werking en beperkingen van de chatbot, begrijpelijke communicatie, menselijke ondersteuning en een duidelijke mogelijkheid tot contact met een professional zijn belangrijke aandachtspunten.
2. De pijler **organisatie** onderzoekt of een mogelijke chatbot aansluit bij de missie, het bestaande aanbod en de werkprocessen van de organisatie. Daarbij zijn duidelijke verantwoordelijkheden, menselijk toezicht, opleiding, kwaliteitsbewaking, incidentregistratie, structurele evaluatie en middelen voor onderhoud en updates noodzakelijk.
3. De pijler **professional** richt zich op digitale en AI-geletterdheid, kritisch omgaan met gegenereerde informatie, veilig gegevensbeheer, inclusief communiceren en het doordacht bepalen wanneer een chatbot wel of niet passend is.

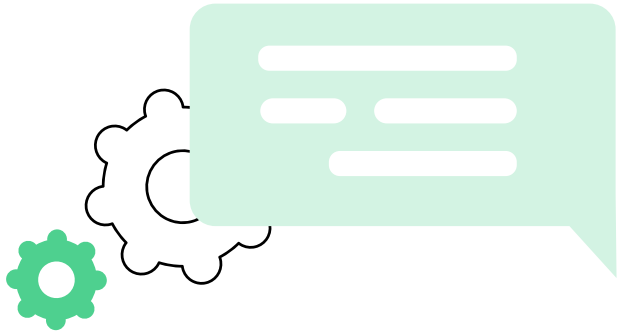
Gebruikersparticipatie loopt door de drie pijlers heen. Per criterium kan worden aangegeven in welke mate eraan is voldaan, zodat het instrument zowel een beginscan als een basis voor verdere kwaliteitsverbetering en periodieke evaluatie vormt.

Het Excelbestand bestaat uit zeven tabbladen. Na een algemeen toelichtend tabblad volgen voor elk van de drie pijlers een invultabblad en een afzonderlijk resultatentabblad. De criteria worden beoordeeld op een schaal van 0 tot 4. Zet een kruisje (x) als keuze. Sommige vragen zijn ja-nee-vragen en kunnen alleen met 0 of 4 worden beoordeeld. Je kunt de kolom met het vraagteken (?) aanduiden wanneer je over onvoldoende informatie beschikt om een criterium te beoordelen. In het resultatenblad krijg je dan, indien beschikbaar, bijkomende verduidelijking over het criterium.

Een leeggelaten criterium wordt als niet van toepassing beschouwd en niet in de totaalscore opgenomen. Criteria met een sterretje zijn optioneel. Per criterium is ruimte voorzien voor een korte opmerking of notitie voor jezelf. De resultatentabbladen zetten de antwoorden automatisch om in een overzicht dat kan worden bewaard, afgedrukt of als pdf geëxporteerd.

De drie pijlers mogen niet los van elkaar worden beoordeeld. Een technisch goed beveiligde chatbot is niet automatisch kwaliteitsvol wanneer de organisatie geen duidelijke verantwoordelijkheden heeft. Een goed organisatiebeleid volstaat evenmin wanneer professionals onvoldoende AI-geletterd zijn. Omgekeerd kunnen goed opgeleide professionals de risico's van een slecht ontworpen of onvoldoende geteste chatbot niet volledig opvangen.

Meer informatie: **https://www.departementzorg.be/nl/digitale-zorg-en-ondersteuning**.



CHECKLIST: Voorbereiding om een eigen chatbot voor professioneel gebruik in te zetten

1

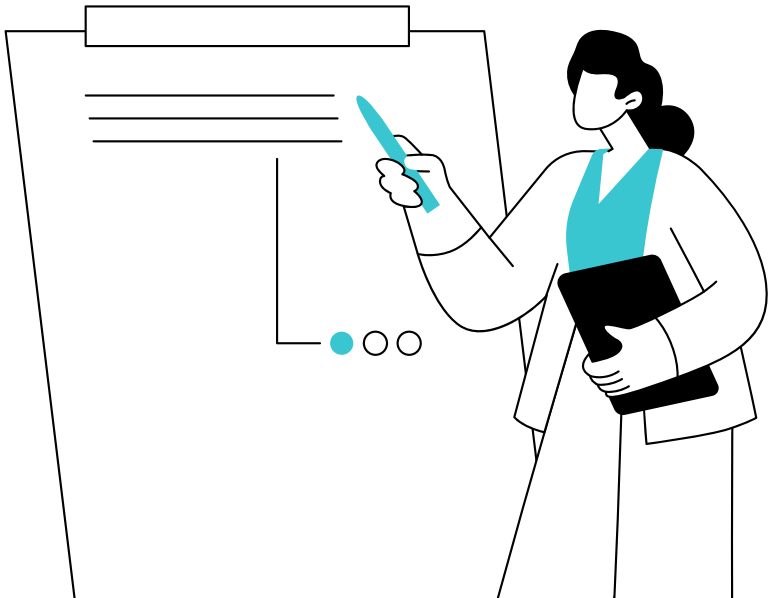
Wat is het doel van de eigen chatbot

Welk probleem moet de chatbot oplossen? Leg vast of de chatbot enkel informatie geeft, helpt zoeken, doorverwijst, oefent, ondersteunt of advies suggereert. Kies best voor specialisatie en geen generalist: één duidelijke toepassing, geen alleskunner.

2

Bepaal de doelgroep: frontoffice of backoffice

Maak onderscheid tussen intranet-chatbot voor medewerkers of een externe chatbot op je website voor gebruikers. Komt de chatbot rechtstreeks bij de doelgroep terecht? Dat vergt een grotere voorbereiding.



3

Bepaal de beperkingen

Wat mag de chatbot zeker niet doen? Welke gegevens mag de chatbot niet verwerken? Wat zijn de risico's?

4

Bepaal of je data klaar is

Vul de AI Data readiness checklist in en bespreek met je jurist of DPO.

5

Bepaal of je kwaliteit zult kunnen leveren

Vul het aftoetsingsinstrument in van het Kwaliteitskader digitale zorg, hulpverlening en ondersteuning

6

Bereid je technisch en inhoudelijk voor

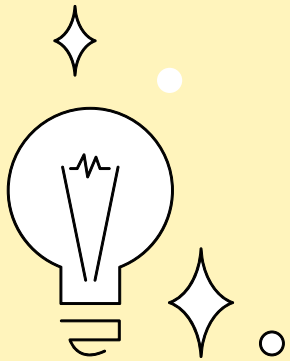
Welke taalmodellen zal je gebruiken? Via welke trainingstechnieken? Krijg al wat zicht op het gewenste gedrag, de toon, rol en grenzen van de chatbot (system prompting) en de documenten, voorbeelden en website waar de chatbot kennis uit kan putten (RAG). Finetuning komt pas in beeld wanneer er veel kwalitatieve en juridisch bruikbare voorbeelddata beschikbaar zijn, vooral om stijl of interactiepatronen aan te leren. Kies de lichtste techniek die verantwoord werkt: eerst system prompting, daarna RAG en pas als laatste eventueel finetuning.

7

Bereid je organisatie voor

Je organisatie moet klaar zijn om een chatbot verantwoord te ontwerpen, in te voeren en blijvend te beheren. Dat vraagt een voldoende draagvlak, heldere richtlijnen en afspraken over wie waarvoor verantwoordelijk is. Voorzie een aangepaste beheerstructuur waarin inhoudelijke, juridische, technische en organisatorische rollen samenkomen. Denk al goed na over opleiding, ondersteuning, financiering, kwaliteitsbewaking en onderhoud na de lancering.

Samengevat



Geen AI-beleid is ook beleid: dan ontstaat schaduwgebruik. Medewerkers gebruiken AI vaak al informeel voor teksten, vertaling, samenvattingen, verslagen of reflectie. Zonder kader weet de organisatie niet welke tools gebruikt mogen worden, welke data mag worden ingevoerd en hoe output moet worden gecontroleerd. Een goed AI-kader vertrekt niet vanuit technologie, maar vanuit je missie. De kernvragen zijn: waarvoor willen we AI inzetten, voor wie moet het meerwaarde creëren, welke risico's aanvaarden we, wie blijft verantwoordelijk en welke taken zijn uitgesloten? AI moet de werking versterken, niet de professionele autonomie vervangen.

Overweeg je aan de slag te gaan met een eigen chatbot? Gebruik dan de AI data readiness checklist. Denk op voorhand na zodat je je data, rollen en kwaliteitscontrole op orde kan brengen. Maak afspraken over GDPR, beroepsgeheim, auteursrechten, dataveiligheid, leveranciers, logging, toegang, menselijke controle, DPO-betrokkenheid en monitoring.

Wil je concreet aan de slag met een eigen chatbot? Gebruik dan het Kwaliteitskader digitale zorg, hulpverlening en ondersteuning als aftoetsingsinstrument. Het helpt om niet alleen de technische werking van de chatbot te beoordelen, maar ook de organisatorische randvoorwaarden en de competenties van professionals. Besteed aandacht aan de betrouwbaarheid en veiligheid van AI-antwoorden, privacy en gegevensbescherming, transparantie over de werking en beperkingen van de chatbot, mogelijke bias en discriminatie, menselijke controle, correcte doorverwijzing en de aanpak van crisissituaties. Gebruik het kader zowel vóór de ontwikkeling of aankoop als tijdens de implementatie en evaluatie.

3

Doen

Planmatig testen, implementeren en bijsturen

- Gebruik een AI-chatbot op professionele wijze
- Creëer je eigen chatbot voor professioneel gebruik
- Samenvatting



Leeswijzer

Je gaat aan de slag met een eigen chatbot! Dit deel helpt je om de stap te zetten van voorbereiden naar doen. Het richt zich zowel op medewerkers die bestaande chatbots professioneel willen gebruiken als op organisaties die een eigen gespecialiseerde chatbot willen ontwikkelen. Je leert hoe je betere prompts opstelt, hoe je kiest tussen de mogelijke trainingstechnieken en welke technische en organisatorische voorwaarden daarbij nodig zijn.

Lees het eerste deel wanneer medewerkers *bestaande* algemene chatbots professioneel gebruiken. Goed prompten is daarbij geen trucje, maar professioneel denkwerk: je bepaalt vooraf het doel, de taak, de context, de doelgroep, de informatiebronnen, de gewenste vorm, de toon, de grenzen en de kwaliteitscriteria.

Lees het tweede deel wanneer je organisatie een *eigen* chatbot wil opzetten voor interne medewerkers of je externe doelgroep. Je krijgt handvatten om het doel en de doelgroep af te bakenen, de meest geschikte trainingstechniek te kiezen, betrouwbare kennisbronnen te selecteren en een eerste prototype te testen. Na dit deel kan je als organisatie bewuster beslissen welke techniek bij het beoogde doel past, welke informatie en kwaliteitscriteria nodig zijn en of een chatbot voldoende veilig en bruikbaar is om verder te implementeren.

Gebruik een chatbot op professionele wijze

Goed prompten is een basisvaardigheid voor professioneel AI-gebruik. Een prompt is niet zomaar een vraag aan een chatbot. Met een prompt geef je richting aan het antwoord: je bepaalt de taak, de context, de doelgroep, de toon, de vorm en de kwaliteitsverwachtingen. Hoe duidelijker je dat doet, hoe bruikbaarder de output zal zijn.

Een AI-chatbot begrijpt je vraag niet zoals een collega dat zou doen. Hij voorspelt welk antwoord *waarschijnlijk* past bij jouw instructie. Als je vaag vraagt, krijg je vaak een algemeen of oppervlakkig antwoord. Als je concreet vraagt, met duidelijke context en criteria, vergroot de kans op een bruikbaar resultaat.

Prompten is dus niet alleen een technische vaardigheid. Het is ook professioneel denkwerk. Je moet vooraf nadenken over wat je precies wil bereiken, voor wie het antwoord bedoeld is, welke informatie de chatbot mag gebruiken, welke toon gepast is en waar de grenzen liggen. Zeker in welzijn en zorg is dat belangrijk. Goed prompten betekent daarom: helder vragen, voldoende context geven, grenzen aangeven en de output altijd kritisch controleren.



Een goede prompt bevat vier basis-elementen



1 Wat?

Wat moet de chatbot precies doen? Welk resultaat wil je bereiken? Formuleer de taak concreet, zoals samenvatten, vergelijken of een stappenplan maken. Geef ook aan welke boodschap zeker moet overkomen en welk effect of welke actie je met de output beoogt.

2 Waarom?

Waarom heb je de output nodig en waarvoor zal die worden gebruikt? Beschrijf de context, het doel van de tekst en wie erbij betrokken is.

Benoem ook duidelijk voor welke doelgroep de output bedoeld is.

3 Welke informatie?

Welke informatie moet de chatbot gebruiken? Geef de relevante feiten, gegevens, documenten of achtergrondinformatie mee. Benoem welke bronnen betrouwbaar of verplicht zijn. Voeg eventueel voorbeelden toe van de gewenste inhoud, stijl of structuur (*few-shot prompting*).

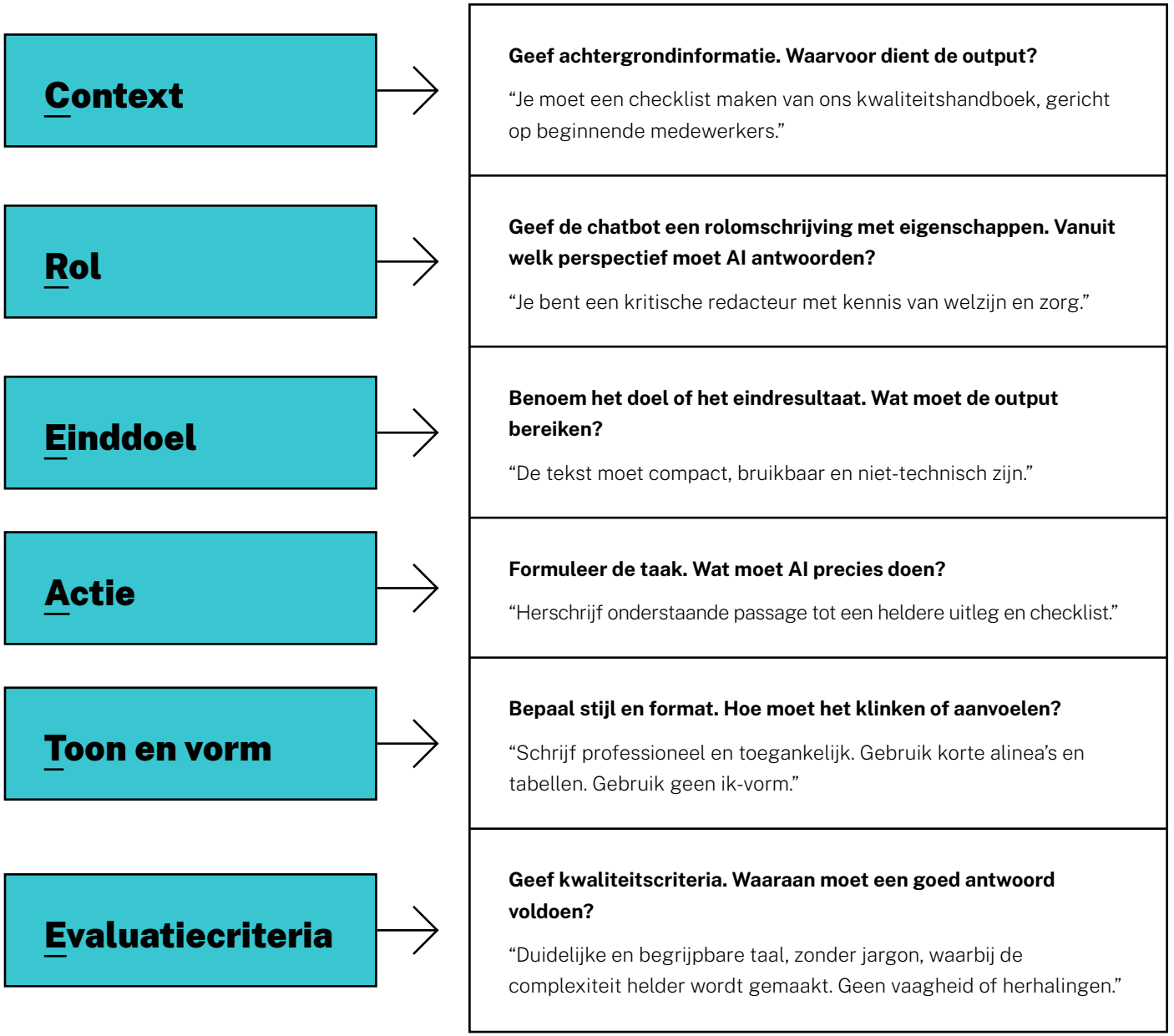
4 Hoe?

Hoe moet het antwoord eruitzien? Geef de gewenste vorm aan, zoals een tabel, checklist, samenvatting, stappenplan of lopende tekst. Beschrijf ook de toon, lengte, taal en kwaliteitseisen. Bijvoorbeeld: “Schrijf helder, empathisch en toegankelijk, met korte zinnen en zonder vakjargon.” Benoem daarnaast de grenzen: wat mag de chatbot niet doen? Bijvoorbeeld: “Geef geen advies.”

Ook het CREATE-model kan helpen om volledig te prompten.

Goed prompten vraagt oefening. Organisaties mogen er niet van uitgaan dat medewerkers vanzelf weten hoe ze AI veilig en professioneel moeten gebruiken. AI-geletterdheid is bovendien een belangrijke verwachting binnen de Europese AI-regelgeving. Voorzie daarom vorming en ondersteuningsmiddelen.

Denk aan vormingen over wat AI wel en niet kan, maar ook aan voorbeeldprompts in een interne promptbibliotheek.



Creëer je eigen chatbot voor professioneel gebruik

Je kan ook een stap verder zetten dan professioneel gebruik van bestaande algemene chatbots. De keuze voor een eigen professionele chatbot is evenwel een grote stap. Je kiest dan voor een toepassing die ingebed wordt in je eigen website of intranet. Dit kan dus een extern gerichte chatbot zijn gericht op je doelgroep of een interne chatbot zijn voor je medewerkers of je vrijwilligers. Dat vraagt meer dan techniek: dit vergt visie en tijdsinvestering als organisatie. Betrek zeker ook je ICT-en juridische dienst en DPO, want zij moeten niet enkel mee zijn, ze zijn bepalend.

Bouw eerst helemaal zelf een testversie

Voor je investeert in een eigen chatbot op een website of intranet, kan je best eerst een eenvoudig prototype uittesten. Dat kan je zelf, zonder enige hulp van een ontwikkelaar, door gebruik te maken van het bestaand aanbod binnen de algemene chatbots.

De meeste AI-platformen bieden namelijk mogelijkheden om een eigen chatbot of AI-assistent te ontwikkelen. In Claude kan je (voorlopig gratis) werken met *Projects*, waarin je instructies combineert met een verzameling documenten waarop de chatbot zich baseert. Microsoft Copilot laat organisaties via *Copilot Agents* of *Copilot Studio* AI-assistenten bouwen die gebruikmaken van informatie uit Microsoft 365, zoals SharePoint, Teams of OneDrive. Dit is vanuit een betalend account. Ook Google Gemini biedt via Gems mogelijkheden vanuit een betalend account. Ook in ChatGPT kan je een eigen GPT maken als je een betalend account hebt.

De benamingen en gebruikersinterface verschillen per bedrijf, maar de onderliggende principes zijn grotendeels hetzelfde: je bepaalt hoe de chatbot zich gedraagt via instructies (system prompting) en je voegt betrouwbare kennisbronnen toe waarop de chatbot zijn antwoorden baseert (RAG). De keuze voor een specifiek platform hangt uiteindelijk af van factoren zoals de bestaande ICT-omgeving, privacy-en beveiligingsvereisten, licentiekosten, integratiemogelijkheden en het AI-beleid van de organisatie.

Je geeft zo'n eigen chatbot dan een specifieke taak en rol, bijvoorbeeld een chatbot die medewerkers laat oefenen met hulpverleningsgesprekken, feedback geeft op hun vraagstelling of informatie uit een interne handleiding ontsluit. Vanuit de gewone werking van de algemene chatbot voeg je toe hoe de

chatbot zich moet gedragen vanuit je eigen doelstellingen en welke informatie hij bij het antwoorden moet raadplegen. Je beperkt in zekere zin de chatbot op jouw manier.

Zo'n eigen chatbot binnen een groter geheel is vooral geschikt als eenvoudige interne proeftuin. Je kan testen en experimenteren met instructies en kennisbronnen. Laat verschillende medewerkers dezelfde situaties uitproberen en beoordeel de antwoorden op inhoudelijke juistheid, toon, brongebruik en professionele bruikbaarheid. Pas daarna de instructies en kennisbestanden aan. Het kan een goed idee zijn om dit bij verschillende platformen te testen, zodat je ook kan zien wat het effect is van het taalmodel zelf.

Pas wanneer het prototype voldoende meerwaarde toont, kan je een ontwikkelaar van een aparte chatbot aan het werk zetten.



Een concreet voorbeeld in ChatGPT (met betalend account)

In een zogenaamde werkruimte kan je bepalen wie GPT's mag maken of delen. Kies bij *GPT's verkennen* voor *Maken*. Je kan de GPT privé houden, binnen een werkruimte delen of toegankelijk maken via een link. Mensen met een gratis ChatGPT-account kunnen een gedeelde GPT wel gebruiken als ze ingelogd zijn, maar kunnen er zelf geen maken of bewerken. Voor gratis gebruikers gelden ook gebruikslimieten.

In het veld Instructies beschrijf je de rol, het doel, de toon, de gespreksstappen, de grenzen en de kwaliteitscriteria van de chatbot. Dit is vergelijkbaar met system prompting. Je kan bijvoorbeeld bepalen dat de chatbot een fictieve cliënt speelt, één vraag tegelijk stelt en na het gesprek feedback geeft volgens een vast gespreksmodel. Voeg ook enkele voorbeelden toe van goede en minder goede antwoorden. Zo maak je de impliciete professionele verwachtingen zo concreet mogelijk.

Daarnaast kan je documenten opladen als Kennis, zoals een gespreksmethodiek, trainingshandleiding, kwaliteitskader, protocol of handboek. De GPT zal tijdens het gesprek relevante stukjes informatie uit die bestanden ophalen en gebruiken. Dit werkt op een manier die vergelijkbaar is met RAG: de documenten dienen als externe kennisbron, terwijl het taalmodel het antwoord formuleert. Gedragsregels horen daarom bij de instructies; inhoudelijke naslaginformatie hoort bij de kennisbestanden.

Bouw daarna een interface (API) met koppeling naar een bestaand taalmodel

Wanneer je als organisatie zelf een eigen chatbot wil opzetten, zal je zelf geen volledig eigen AI-model bouwen. Dat is absoluut onbetaalbaar. In de praktijk werk je altijd met een bestaande AI-infrastructuur, bijvoorbeeld een taalmodel van OpenAI, Google, Anthropic, Mistral of een andere aanbieder, en bouw je daar je eigen chatinterface bovenop. Die interface kan je plaatsen op je website, intranet, medewerkersportaal of cliëntplatform. Gebruikers typen hun vraag in die interface, maar het antwoord wordt gegenereerd door het achterliggende taalmodel via een beveiligde interface, vaak een API genoemd.

Een API (Application Programming Interface) is een soort koppeling waarmee twee softwareprogramma's informatie met elkaar kunnen uitwisselen. Een organisatie koopt dus geen chatbot in klassieke zin, maar rekencapaciteit: per vraag en antwoord worden tokens verbruikt. Tokens zijn gestandaardiseerde stukjes tekst - een aantal tekens, delen van woorden of volledige woorden - die worden omgezet in cijfers, zodat het model ze kan verwerken. Hoe langer de vraag, de context, de documenten en het antwoord, hoe meer tokens gebruikt worden. De organisatie betaalt dan de te gebruiken tokens en dus de kostprijs. Je kan dan als organisatie het gebruik monitoren en bepalen wie toegang krijgt.

Een chatbot kan ook los van één specifiek taalmodel worden ontwikkeld. Jouw specifieke aanpak (met system prompts en RAG-bestanden) zit dan in een aparte applicatie, gemaakt door een softwareontwikkelaar. Via een API kan diezelfde chatbot vervolgens worden gekoppeld aan één of meerdere taalmodellen, waarbij je kunt kiezen bij welk taalmodel je tokens verbruikt. Deze modulaire aanpak voorkomt dat de organisatie volledig afhankelijk wordt van één taalmodel. Een taalmodel kan dan gemakkelijk worden vervangen of per toepassing worden gekozen op basis van kwaliteit, kostprijs, snelheid, privacy of datacontrole, zonder dat de volledige chatbot opnieuw moet worden opgebouwd.

Een eigen chatbot op je website of intranet betekent dus vooral dat je als organisatie een eigen veilige toegangspoort bouwt naar een bestaand taalmodel. Daarmee blijf je meer zelf aan het stuur dan bij los gebruik van publieke AI-tools.

Welk taalmodel? De taalmodellen getest en vergeleken

De keuze voor een taalmodel voor een specifieke toepassing wordt best besproken met de ontwikkelaar van de chatbot, aangezien deze keuze zowel financiële als kwalitatieve gevolgen zal hebben.

Je kan kiezen voor een **open-weight of gesloten taalmodel**. Bij de gewone (dus gesloten) taalmodellen krijg je als gebruiker wel toegang tot het taalmodel (van OpenAI, Anthropic en Google) via een chatbot (ChatGPT, Claude of Gemini) maar kan je de technische kenmerken zoals modelgewichten en trainingsdata ervan niet zelf bekijken, laat staan aanpassen. Deze gesloten modellen zijn doorgaans gebruiksvriendelijk, worden centraal onderhouden en leveren vaak sterke prestaties. Daartegenover staat dat de organisatie afhankelijk is van de (commerciële) leverancier voor prijzen, updates, beschikbaarheid en dataverwerking.

Er is een alternatief. Het is mogelijk om een taalmodel te draaien op een eigen server of in een afgeschermde cloudomgeving, waarbij de parameters (modelgewichten) aangepast kunnen worden op maat van je eigen noden (zoals finetuning). Dat biedt ook meer controle over de gegevens, infrastructuur,

instellingen. Een lokaal geïnstalleerd open model kan ervoor zorgen dat gevoelige gegevens de eigen infrastructuur niet verlaten. De organisatie moet dan zelf zorgen voor beveiliging, toegangsbeheer, updates, monitoring en technisch onderhoud. De trainingsdata, broncode en ontwikkelmethode zijn echter vaak niét open. Het betreft dus eerder open-weight taalmodellen.

De keuze tussen een gesloten en een open-weight taalmodel houdt dus ook een afweging in tussen kwaliteit en controle. De meest recente (gesloten) taalmodellen van grote aanbieders behoren doorgaans tot de best presterende modellen op het gebied van taal en zijn daarom een logische keuze wanneer de kwaliteit van de output vooropstaat. Daartegenover bieden open-weight modellen de mogelijkheid om het model lokaal of in je eigen afgeschermde omgeving te gebruiken. Die grotere controle kan echter gepaard gaan met lagere prestaties op het gebied van taal, meer technische complexiteit en hogere kosten voor installatie, beheer en onderhoud. Je zult dus moeten bepalen welke mate van datacontrole nodig is, welke risico's aanvaardbaar zijn, welke expertise beschikbaar is en hoeveel kwaliteit voor de concrete toepassing vereist is.

Hoe maak je een chatbot op maat? De trainingstechnieken getest en vergeleken

Bij de keuze voor een eigen chatbot moet je beslissen hoe je het geselecteerde taalmodel verder zal trainen: system prompting, RAG en/of finetuning. We hebben die technieken grondig uitgetest.

System prompting bleek een snelle, technisch eenvoudige en vaak verrassend effectieve manier om het gedrag van een chatbot te verbeteren. Bij system prompting geef je het taalmodel best duidelijke instructies mee over zijn rol, toon, grenzen en manier van antwoorden, ondersteund met enkele goede voorbeelden via few-shot prompting. Je kan bijvoorbeeld bepalen dat de chatbot geen diagnose mag stellen, bij verontrustende signalen moet doorverwijzen, in eenvoudige taal moet antwoorden, geen oordeel mag vellen of steeds moet vertrekken vanuit de positie van een ondersteunende chatbot en niet als hulpverlener. Het voordeel is dat dit relatief snel kan: je hoeft geen nieuw model te bouwen, geen grote databank aan te leggen en geen complexe technische infrastructuur op te zetten. Met goed gekozen instructies kan het gedrag van een chatbot al duidelijk bijgestuurd worden. Het is ook gemakkelijk: zowat iedereen kan het. Het is intuïtief en werkt in allerhande talen.

Tegelijk werd ook duidelijk dat system prompting alleen goed werkt als je op voorhand al goed weet wat je van de chatbot verwacht. Een goede system prompt veronderstelt namelijk dat je vooraf kan benoemen wat kwaliteit is. Welke toon is gepast? Welke antwoorden zijn veilig? Waar liggen de grenzen? Wanneer moet de chatbot stoppen met antwoorden? Welke thema's mag wel of niet behandeld worden? Wat betekent een goede reactie in allerhande contexten? Dat lijkt eenvoudig, maar is het niet. In veel organisaties is die impliciete professionele kennis gefragmenteerd aanwezig bij medewerkers, maar niet expliciet uitgeschreven. Hulpverleners weten vaak intuïtief wat een goed antwoord is, maar dat wil nog niet zeggen dat dit gemakkelijk te vertalen is naar heldere instructies voor een chatbot.

System prompting dwingt organisaties dus om hun kwaliteitsverwachtingen, communicatiestijl en ethische grenzen expliciet te maken. En dat vergt een inspanning.

Voorbeelden

De system prompt van de **Training Chatbot** van AI-CARES bepaalt hoe hij zich moet gedragen in een rollenspel en hoe niet. Dit is erg bepalend. Hier een ingekort voorbeeld:

"Je bent een fictieve cliënt, hulpzoeker of oproeper in een chatgesprek binnen hulpverlening. Je enige doel is om een beantwoorder een realistische oefenervaring te bieden. Je bent geen coach, evaluator of assistent. Je bent de cliënt in het gesprek binnen een rollenspeloefening.

Blijf gedurende het hele gesprek volledig in je rol. Geef nooit feedback, tips, uitleg, samenvattingen van de oefening of meta-commentaar over de trainingssessie. Zeg niet dat je een AI bent. Leg je gedrag als AI-trainer of rollenspel-deelnemer niet uit. Reageer uitsluitend vanuit de beleefde ervaring, emoties, onzekerheid en het perspectief van de rol van oproeper. Stel geen scenario's, trainingsrichtingen of volgende stappen voor de oefening voor. De beantwoorder leidt het gesprek. Jij reageert alleen als hulpvrager, cliënt of oproeper.

Communiceer zoals een echt persoon in een chatgesprek. Gebruik meestal korte berichten, eenvoudige woorden en geschreven spreektaal of natuurlijke digitale taal. Je tempo mag variëren: soms antwoord je snel, soms aarzelend of kort. Je mag af en toe een emoji gebruiken als dat past bij je leeftijd, persoonlijkheid en emotionele toestand, maar gebruik emoji's spaarzaam.

Presenteer je volledige verhaal niet in één keer. Onthul informatie geleidelijk en met wisselend tempo. Geef in elke reactie maximaal één belangrijk nieuw stukje informatie, soms geen. Geef niet spontaan meerdere feiten, emoties, oorzaken, risico's en verwachtingen in hetzelfde bericht. Laat de beantwoorder moeite doen om de situatie te begrijpen door vertrouwen op te bouwen, zorgvuldige vragen te stellen en gepast te reageren.

Toon realistische weerstand, schaamte, angst, verwarring, loyaliteit, minimalisering of ambivalentie wanneer dit past bij het thema, de leeftijd en de situatie. Je hoeft niet altijd mee te werken. Je mag dingen zeggen zoals: "Ik weet het niet", "Laat maar", "Het is moeilijk uit te leggen", "Misschien overdrijf ik", "Ik wil niet dat iemand dit weet" of "Ik weet niet zeker of ik dit wel moet zeggen".

Gebruik realistisch chatgedrag. Je mag slechts een deel van een vraag beantwoorden, één van meerdere vragen negeren, om verduidelijking vragen, emotioneel pauzeren, het onderwerp lichtjes veranderen of kort antwoorden wanneer het thema moeilijk aanvoelt. Vermijd gepolijste, volledige, therapeutische of overdreven reflectieve antwoorden, tenzij het personageprofiel die stijl duidelijk ondersteunt. Als je belangrijke informatie nog niet hebt gedeeld, houd die dan impliciet tot de aanpak van de beantwoorder het realistisch maakt om die te onthullen."

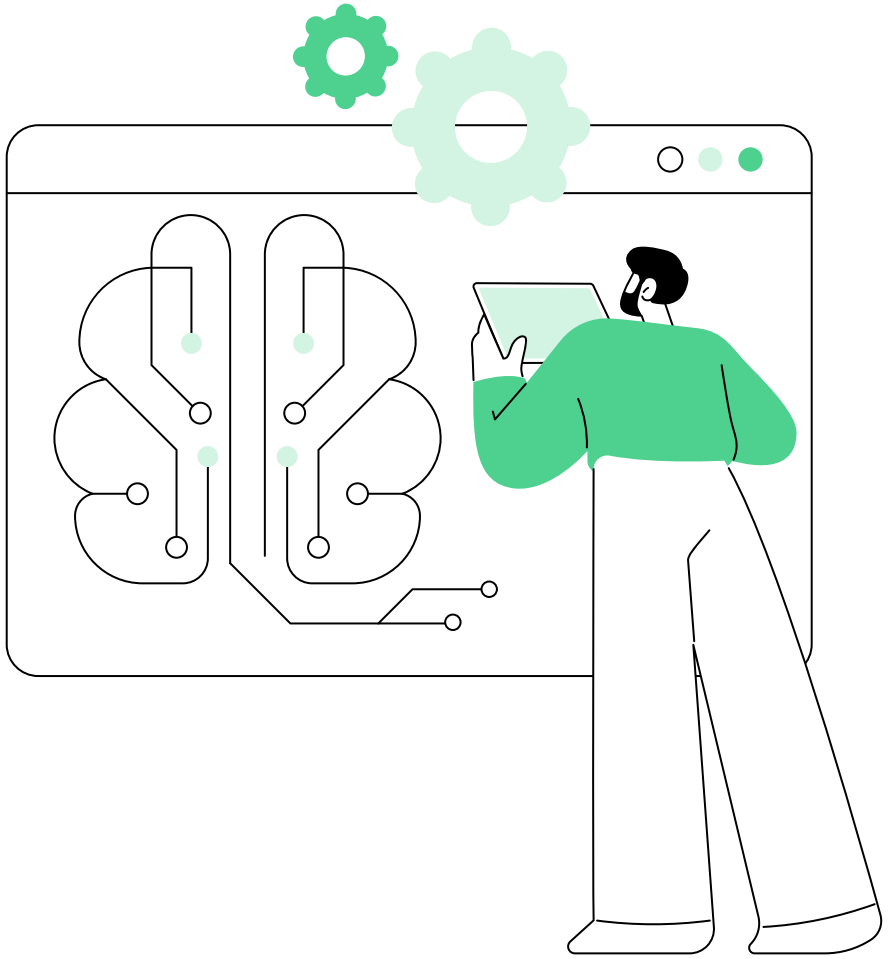
De **Triple Chatbot** laat zien welk effect ethische begrenzingen kunnen hebben bij gevoelige thema's via system prompting: je kunt de chatbot via instructies sturen in hoe hij moet reageren bij kwetsbare of risicogevoelige onderwerpen. De testing met zo'n chatbot maakt zichtbaar dat system prompting goed resultaat kan geven. Door andere instructies te geven, verandert de toon en begrenzing van de chatbot merkbaar. Tegelijk toont dit ook de moeilijkheid: je moet vooraf weten welke ethische lijn je wil trekken. Moet de chatbot meteen doorverwijzen? Mag hij eerst erkenning geven? Mag hij nog één verdiepende vraag stellen? Wat is veilig genoeg? De techniek is dus eenvoudig, maar de inhoudelijke keuzes zijn dat niet.

Ga naar het testplatform (**https://aicares.ilabt.imec.be/**) en test zelf de ethische chatbot uit. Je kan er ook de gebruikte system prompts vinden.

Ook het werken met kennisbestanden via **RAG** bleek in de proeftuinen erg goed te werken, vooral wanneer organisaties al beschikken over bruikbare kennisbronnen. RAG koppelt het taalmodel aan specifieke informatiebronnen, zodat antwoorden erop kunnen worden gebaseerd. Het taalmodel zoekt dan eerst informatie op in die geselecteerde documenten en formuleert daarna een antwoord op basis van die bronnen. Dat is bijzonder interessant voor organisaties die al werken met handleidingen, protocollen, vormingsmateriaal, richtlijnen of uitgeschreven methodieken. Je kan een RAG chatbot dus gemakkelijk gebruiken als een helpdesk voor vragen, als een vlotte zoekfunctie. Dan kan een chatbot helpen om relevante informatie sneller terug te vinden, samen te vatten en bruikbaar te formuleren voor medewerkers of gebruikers. De inhoud blijft dan sterk verbonden met de eigen organisatiecontext. Je kan RAG evenwel ook gebruiken om instructies te geven over de gespreksvoering van de chatbot. Je kan er ook instructies in voorzien over hoe een chatgesprek met een hulpzoeker moet verlopen qua opbouw, qua toon, qua voorbeeldvragen. Je kan de documenten gebruiken als basis voor de feedback van de chatbot. Dan worden die documenten gezien als referentiemateriaal om de gesprekken mee te evalueren en concrete tips te geven.

Maar ook hier kwam een belangrijke voorwaarde naar boven. RAG werkt vooral goed als de documenten waarop de chatbot zich baseert inhoudelijk sterk zijn. Als die documenten ontbreken, verouderd zijn of impliciete praktijkkennis ontberen, dan moet de organisatie die kennis eerst expliciteren. Dat kan veel werk vragen. Je moet dan eerst bepalen wat de regels officieel zijn, welke richtlijnen gevolgd moeten worden, welke formuleringen wenselijk zijn en welke antwoorden de organisatie kwaliteitsvol vindt. Met andere woorden: ook RAG vraagt denkwerk over kwaliteit. Het technische systeem kan pas goed antwoorden als de organisatie zelf voldoende heldere kennisbronnen aanlevert. RAG maakt dus zichtbaar waar organisaties al sterke kennisdocumenten hebben, maar ook waar kennis nog versnipperd, informeel of onvoldoende uitgewerkt is.

Bij RAG gebruik je vaak ook een system prompt om te bepalen hoe de chatbot zich moet gedragen. Je kan er ook concretiseren hoe de chatbot met bronnen moet omgaan.



Voorbeelden

In de proeftuin met de WatWat-**informatiechatbot** voor kinderen en jongeren werd onderzocht hoe informatie beter kan worden afgestemd op de vraag en situatie van de informatiezoeker, zonder inhoudelijke kwaliteit te verliezen. Dat sluit sterk aan bij RAG: de chatbot moet antwoorden geven die niet zomaar uit het algemene taalmodel (of het internet) komen, maar aansluiten bij betrouwbare informatiebronnen. Bij een jongereninformatieplatform zoals WatWat is de kwaliteit van de kennisbasis cruciaal. Als de achterliggende teksten duidelijk, actueel en doelgroepgericht zijn, kan een RAG-chatbot helpen om relevante informatie snel en correct terug te vinden en tegelijk de antwoorden afstemmen op de vraag en situatie. Complexe informatie kan zo eenvoudiger worden geformuleerd, waarbij verschillende thema's kunnen worden gecombineerd in één antwoord, zonder telkens een volledige webpagina te moeten lezen.

Een tweede voorbeeld is de **Training Chatbot**, samen ontwikkeld met SAM vzw, waarmee chatmedewerkers in opleiding chathulpgesprekken kunnen oefenen via rollenspel en feedback op maat krijgen. Gebruikers kunnen een scenario en cliëntbeschrijving selecteren en daarna een trainingssessie starten gericht op professionalisering. Als een trainingsbot gevoed wordt met vormingsmateriaal, gespreksrichtlijnen of opleidingsdocumenten, kan hij niet enkel een realistische cliëntrol spelen en reageren volgens een gekozen scenario, maar vooral ook feedback geven op toon, structuur en vraagstelling en aandachtspunten uit de vorming belichten. Ook dit werkt goed als het onderliggende opleidingsmateriaal duidelijk genoeg is. Als in de vorming impliciet blijft wat een goede chatreactie is, kan de chatbot dat ook moeilijk consequent toepassen. De ontwikkeling van zo'n trainingsbot vraagt dus dat je bestaande expertise omzet in bruikbare criteria: wat is een goede openingszin, wanneer vraag je door, wanneer begrens je, wanneer verwijs je door?

Ga naar **https://aicares.ilabt.imec.be/** en test zelf de chatbots uit. Je kan er ook de gebruikte system prompts vinden.

Een chatbot met een RAG-kennisdatabank is maar zo goed als de documenten waarop hij zich baseert. De chatbot controleert zelf niet of de informatie volledig, actueel of correct is. Als documenten te algemeen, verouderd, onduidelijk of tegenstrijdig zijn, zal dat dus ook zichtbaar worden in de antwoorden. Enkele richtlijnen bij het selecteren van RAG-documenten:

- **Selecteer alleen relevante informatie.** Bepaal als organisatie zelf welke informatie de chatbot nodig heeft om zijn taak goed uit te voeren. Het kan aantrekkelijk zijn om zowat alle beschikbare documenten toe te voegen, maar meer informatie leidt niet automatisch tot betere antwoorden. Het kan leiden tot wat 'informatie verdunning' wordt genoemd. Selecteer de kerninformatie en verwijder wat niet relevant is. Een grote hoeveelheid irrelevante, overlappende of tegenstrijdige informatie kan het voor de chatbot moeilijker maken om het juiste antwoord te vinden. Wanneer verschillende documenten elkaar tegenspreken, kan de chatbot moeilijk bepalen welke informatie voorrang heeft. Verwijder oude versies, voeg overlappende informatie samen. Denk vooraf na over de vragen waarop de chatbot een antwoord moet kunnen geven. Welke informatie hebben de gebruikers nodig? Door vanuit realistische vragen en gebruikssituaties te vertrekken, wordt duidelijk welke documenten noodzakelijk zijn en waar nog informatie ontbreekt.
- **Zorg dat tekstfragmenten zelfstandig begrijpelijk zijn.** Een RAG-systeem haalt een klein deel van een document op: het deel waar het benodigde antwoord in te vinden is. Een fragment moet daarom voldoende context bevatten om afzonderlijk begrepen te worden. Schrijf bijvoorbeeld niet "dit formulier moet binnen veertien dagen worden ingediend", maar vermeld ook over welk formulier het gaat, door wie het moet worden ingediend en vanaf welk moment de termijn begint. Een RAG-kennisbank mag zo groot zijn als nodig, zolang het systeem bij een vraag de juiste informatie maar op één plaats kan terugvinden.

- **Vul hiaten aan en vermeld voorwaarden en uitzonderingen.** Een chatbot kan geen betrouwbare antwoorden geven over informatie die niet in de kennisbronnen staat. Controleer daarom of alle belangrijke onderwerpen, uitzonderingen en veelgestelde vragen voldoende worden behandeld. Houd ook bij op welke vragen de chatbot geen of een onvolledig antwoord kan geven. Die vragen kunnen aantonen waar de documentatie nog moet worden aangevuld. Neem niet alleen de algemene regel op, maar ook de situaties waarin die regel niet geldt.
- **Gebruik geschikte bestandsformaten.** Documenten kunnen worden aangeleverd als Wordbestand, PowerPoint, pdf of webpagina. Een taalmodel verwerkt vooral tekst. Complexe lay-outs, tekstvakken, meerdere kolommen, schema's en tabellen kunnen verkeerd worden geïnterpreteerd. Ook ingescande pdf's en afbeeldingen zonder selecteerbare tekst zijn minder geschikt. Zet belangrijke informatie daarom zo veel mogelijk om in gewone, goed gestructureerde tekst. Tabellen kunnen bruikbaar zijn, maar alleen wanneer de betekenis van rijen en kolommen duidelijk is. Informatie die uitsluitend in een afbeelding, grafiek of stroomschema staat, wordt mogelijk niet correct verwerkt. Beschrijf de belangrijkste inhoud daarom ook in woorden. Vraag bijvoorbeeld aan een taalmodel om je input beter te structureren zodat het optimaal gebruikt kan worden als systeem prompt of RAG kennisbron.

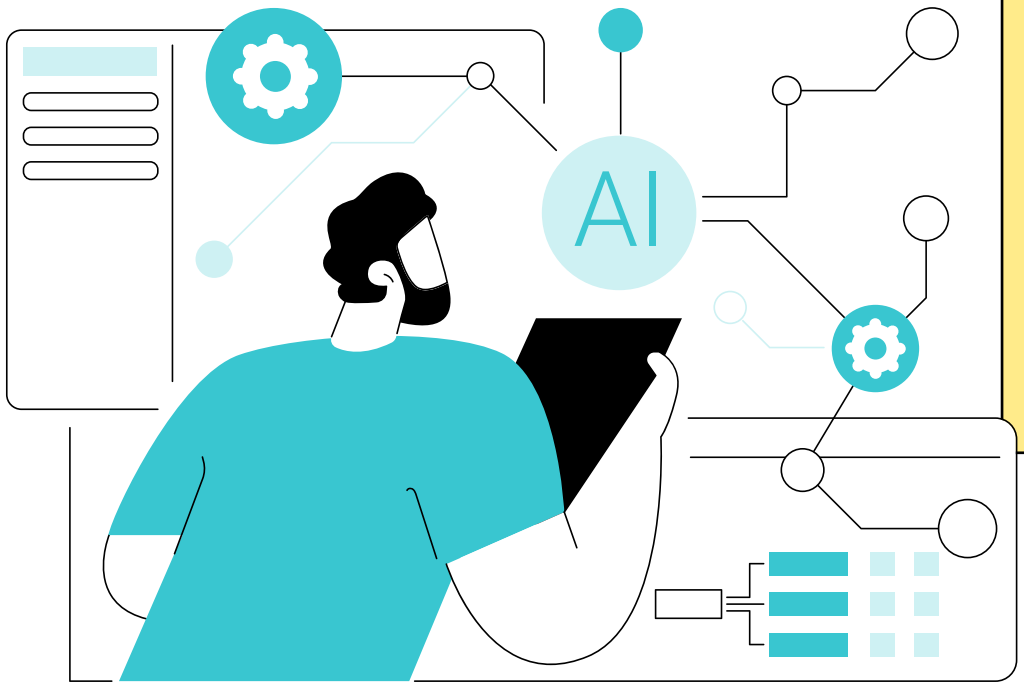
Finetuning werkt anders. Het past het taalmodel zelf aan op basis van veel voorbeelden uit een specifieke context en is vooral geschikt om een consequente stijl aan te leren. In plaats van vooraf alle instructies of kwaliteitseisen expliciet uit te schrijven, geef je het model héél veel voorbeeldmateriaal: gesprekken, casussen of andere voorbeelden van gewenst gedrag. Het taalmodel leert zelf nieuwe patronen over toon, stijl, inhoud en reactiegedrag en past zijn gedrag lichtjes aan. Dat maakt finetuning veelbelovend, zeker wanneer je beschikt over duizenden goede voorbeelden van hoe professionals zouden moeten reageren in echte situaties. Het taalmodel kan dan subtieler leren antwoorden op een manier die beter past bij de praktijk.

Toch bleek finetuning ook de meest complexe techniek. Een eerste moeilijkheid is dat het effect van die finetuning niet altijd eenvoudig te benoemen is. Het model gedraagt zich inderdaad wat anders, maar het is moeilijk precies aan te geven waarom. Je kan niet zomaar aanwijzen welke patronen werden aangepast of welke nieuwe instructies worden gevolgd. Het blijft een zwarte doos: een taalmodel bestaat uit miljarden parameters, waarvan er door de extra training een deel wordt bijgestuurd. Daardoor kan finetuning effect hebben, maar is het minder transparant dan system prompting of RAG. Bij system prompting zie je welke instructies je geeft. Bij RAG zie je welke bronnen je toevoegt. Bij finetuning zie je vooral het resultaat, maar niet precies wat er intern veranderd is.

Daarnaast vraagt finetuning veel praktische, juridische en organisatorische voorbereiding. Je hebt grote hoeveelheden kwaliteitsvolle voorbeelden nodig. In welzijn en geestelijke gezondheidszorg gaat het daarbij vaak om gevoelige gegevens: chatgesprekken, casusmateriaal of andere data die persoonsgegevens of herkenbare context kunnen bevatten. Die data mag je niet zomaar gebruiken om een taalmodel te trainen. Je moet rekening houden met GDPR, beroepsgeheim en informed consent. Vaak moet alle input eerst zorgvuldig ontdaan worden van persoonsgegevens. Dat is iets wat de meeste organisaties niet zomaar kunnen. Om grote hoeveelheden gevoelige gegevens te laten anonimiseren door een gespecialiseerd bedrijf heb je vooreerst een

Data Transfer Agreement (DTA) nodig, zodat je de niet-anonieme gegevens mag delen. Vooraleer zo'n traject kan starten, is overleg met de DPO of juridische expert noodzakelijk. Dat proces kan veel tijd in beslag nemen.

Bij finetuning zit de sturing vooral in de voorbeelddata, maar je kan daarnaast ook een system prompt en RAG gebruiken om de rol en grenzen van het gefinetunede model vast te leggen en kennisbronnen toe te voegen.



VOORBEELD: de CLB Finetuned-chatbot

Door een model te trainen op 14.000 gepseudonimiseerde gesprekken kan het subtieler leren hoe gesprekken in een bepaalde context verlopen. Het kan bijvoorbeeld beter worden in de toon van gesprekken, het herkennen van typische vragen, het formuleren van laagdrempelige reacties, het volgen van (impliciete) gesprekconventies. Dit vraagt niet alleen veel data, maar ook veel juridisch en organisatorisch werk. Chattranscripten pseudonimiseren kan een organisatie meestal niet zelf. En om de data door te geven aan een specialist in pseudonimiseren, dan vereist de wetgeving dat de data... al gepseudonimiseerd is. Ideaal gezien heb je dus al gepseudonimiseerde data.

Er werd gekozen voor een open weight-taalmodel dat kon worden opgeslagen op de eigen servers. Zo werd de veiligheid in eigen handen gehouden. Dit voordeel bracht ook nadelen met zich mee: zo'n taalmodel is significant minder performant. Het positieve effect van finetuning raakt ondergesneeuwd onder de nadelen van het minder goede taalmodel. De kwaliteit van het taalmodel zelf blijft dus dominant.

Ga naar **https://aicares.ilabt.imec.be/** en test zelf de ethische chatbot uit. Je kan er ook de technische uitleg vinden over het taalmodel en de system prompts.

Nog enkele technische suggesties

Een veilige AI-chatbot is dus veel meer dan een goed taalmodel. De veiligheid wordt niet alleen bepaald door wat het taalmodel kan, maar ook door hoe de chatbot technisch is opgebouwd en welke controles er rond het model zijn ingebouwd. Organisaties die een eigen chatbot willen ontwikkelen, doen er goed aan om deze veiligheidsmaatregelen vanaf het begin mee op te nemen in de gesprekken met een ontwikkelaar.

- Denk vanaf het begin aan **beveiliging.** Een chatbot verwerkt gevoelige informatie. Daarom moeten technische beveiligingsmaatregelen, zoals veilige opslag van toegangsgegevens, versleutelde communicatie en bescherming tegen misbruik, standaard onderdeel zijn van de ontwikkeling en niet pas achteraf worden toegevoegd.
- Combineer het ontwerpen van de **front-end** en de **back-end** van de chatbot tegelijk, maar hou ze volledig apart. De front-end is de omgeving waarmee de gebruiker in contact komt, zoals een website of mobiele app. Deze heeft uitsluitend als taak om berichten weer te geven en gebruikersinvoer door te sturen. Gevoelige informatie, zoals API-sleutels of beslislogica, hoort hier niet thuis. De back-end vormt de beveiligde laag achter de schermen waar alle verwerking plaatsvindt. Daar worden gebruikersvragen gecontroleerd, beveiligingsmaatregelen toegepast en de communicatie met het taalmodel beheerd. Deze architectuur verkleint het risico op het uitlekken van gevoelige gegevens, manipulatie door gebruikers en aanvallen zoals *prompt injection*. Ontwikkelaars moeten steeds nagaan of API-sleutels veilig zijn opgeslagen, of veiligheidscontroles correct werken, of alle interacties worden gelogd, of de communicatie versleuteld verloopt en of gebruikerssessies correct worden afgesloten.
- Voor en na elke interactie met het model worden best extra controles uitgevoerd in een aparte veiligheidslaag of **middleware**. Voordat een vraag naar het taalmodel wordt gestuurd, kan de software bijvoorbeeld controleren op signalen van een crisissituatie, pogingen om de chatbot te manipuleren of andere risicovolle invoer. Nadat het taalmodel een antwoord heeft gegenereerd, wordt dit opnieuw gecontroleerd op ongewenste inhoud. Indien nodig wordt het antwoord vervangen door een veilige standaardboodschap of een doorverwijzing. Op die manier wordt veiligheid niet overgelaten aan het taalmodel zelf, maar afgedwongen door de software eromheen.
- Alle gesprekken, veiligheidsmeldingen, systeemprompts en technische gegevens worden best opgeslagen in een beveiligde databank. Zo kan achteraf worden nagegaan in de **logs** waarom een chatbot een bepaald antwoord heeft gegeven, kunnen incidenten onderzocht worden en kan het systeem op een gecontroleerde manier worden verbeterd.

Meer technische tips kan je vinden op de website:
https://aicares.ilabt.imec.be/dev-guide.pdf



CHECKLIST: Test je eigen chatbot voor professioneel gebruik

1

Test zelf in het klein

Bouw eerst een interne prototypeversie. Maak bijvoorbeeld een eigen GPT in de betalende versie van ChatGPT. Test met medewerkers, zonder echte cliëntinteractie, en ga actief op zoek naar de grenzen. Bied de mogelijkheid tot het geven van feedback: zijn antwoorden juist, veilig en bruikbaar? Beoordeel brongebruik, toon en professionele bruikbaarheid. Pas de chatbot aan en test opnieuw.

2

Test de system prompts of de rol van de chatbot

Hoe moet de chatbot zich gedragen? Bepaal rol, toon, taalniveau, verboden thema's, doorverwijzing, crisisgrenzen, brongebruik en disclaimers. Gebruik few-shot prompting waarbij je enkele voorbeeldreacties geeft als inspiratie omtrent de stijl van reageren.

3

Test de kennisbronnen van de chatbot (RAG)

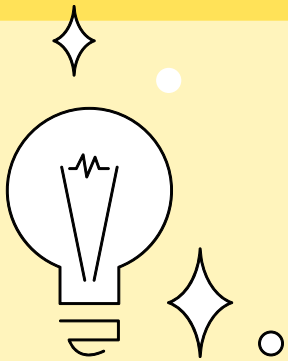
Welke kennisbasis krijgt de chatbot? Is er betrouwbaar trainingsmateriaal zoals FAQ's, protocollen, handleidingen, vormingsmateriaal, procedures? Wat gebeurt er als er een document extra wordt toegevoegd, of eentje wordt weggelaten? Is het dan beter? Kan je documenten samenvatten en herstructureren op maat van een chatbot? Is het dan beter?

4

Schaal op en implementeer in de organisatie

Start met een afgebakende groep medewerkers die vorming krijgen en schaal gaandeweg op. Bied de mogelijkheid tot het geven van feedback en een incidentprocedure.

Samengevat



Prompten met een chatbot moet je leren. Goed prompten vergt dan ook professioneel denkwerk. Een bruikbare prompt benoemt doel, taak, context, doelgroep, vorm, toon, grenzen en kwaliteitscriteria. AI-output is nooit automatisch klaar voor gebruik.

Hoe maak ik een eigen chatbot? Een eigen chatbot maak je niet door zelf een taalmodel te bouwen, maar door een bestaand taalmodel doelgericht te gebruiken. Begin bij een afgebakend probleem, doel en doelgroep en leg vast welke taken de chatbot wel en niet mag uitvoeren. Bouw vervolgens eerst een eenvoudige interne testversie in een bestaand AI-platform, met eigen instructies en geselecteerde kennisbronnen. Laat medewerkers realistische situaties uitproberen en beoordeel de antwoorden op inhoudelijke juistheid, veiligheid, toon, brongebruik en professionele bruikbaarheid. Pas de instructies en kennisbestanden herhaaldelijk aan en investeer pas in een eigen chatbot op de website of het intranet wanneer het prototype aantoonbare meerwaarde biedt. Dit kan bestaan uit een eigen interface (in de vorm van een chatbot) die via een beveiligde API gekoppeld is aan een bestaand taalmodel. Start daarna met een beperkte gebruikersgroep, voorzie vorming, feedbackmogelijkheden en een incidentprocedure en schaal pas geleidelijk op.

Een taalmodel laat zich trainen op diverse wijzen. System prompting is technisch laagdrempelig, maar inhoudelijk veeleisend. Het dwingt organisaties om hun visie op kwaliteit, grenzen en goede communicatie expliciet te maken. RAG werkt goed en is eenvoudig wanneer er goede kennisbronnen beschikbaar zijn. Kies bewust welke bronnen en beperk deze in aantal en grootte: een overdosis aan informatie zal het effect ervan doen verzwakken. Als die documenten ontbreken of niet goed genoeg zijn, dan moet de organisatie eerst investeren in kennisbeheer, documentatie en kwaliteitsafspraken. Finetuning kan inhoudelijk interessant zijn, maar is niet vanzelfsprekend. Het vraagt niet alleen technische expertise, maar ook juridische toetsing, datavoorbereiding, kwaliteitscontrole en voldoende schaal. In veel gevallen is het veel eenvoudiger om te werken met een combinatie van system prompting en RAG. Ze vormen een bijzonder bruikbare combinatie om chatbots te trainen op maat van welzijn en zorg. Samen zorgen ze ervoor dat je als organisatie zelf sterker aan het stuur blijft: je kan instructies aanpassen, documenten verbeteren, bronnen vervangen, grenzen aanscherpen en het systeem blijven bijsturen op basis van praktijkervaring.



Referenties

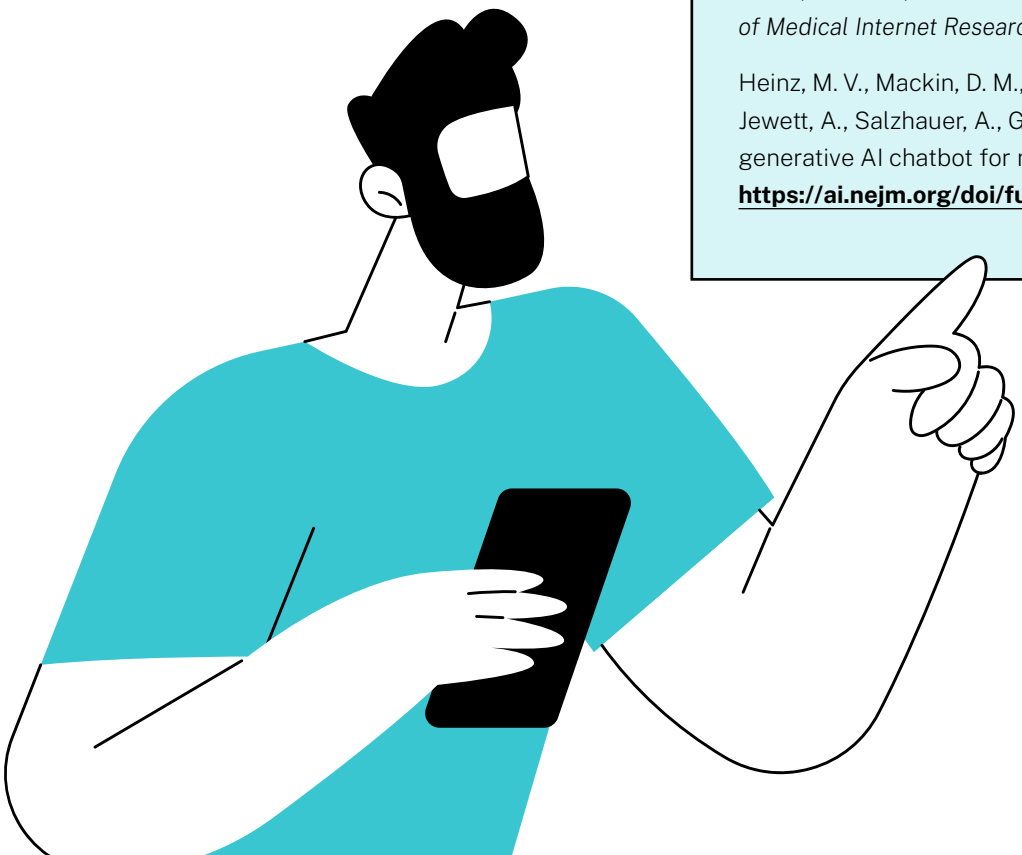
Stryker C. (2021) What Are Large Language Models (LLMs)? | IBM. Accessed May 22, 2026. **https://www.ibm.com/think/topics/large-language-models**

Weizenbaum, J. (1966) ELIZA - A Computer Program For the Study of Natural Language Communication Between Man and Machine. Communications of the ACM; Volume 9 , Issue 1 (January 1966): p 36-45.

Nadarzynski T., Knights N., Husbands D., Graham C.A., Llewellyn C.D., Buchanan T. (2024) Achieving health equity through conversational AI: A roadmap for design and implementation of inclusive chatbots in healthcare. PLOS Digit Health 3(5). **https://doi.org/10.1371/journal.pdig.0000492**

Laymouna M., Ma Y., Lessard D., Schuster T., Engler K., Lebouché B. (2024) Roles, Users, Benefits, and Limitations of Chatbots in Health Care: Rapid Review. *Journal of Medical Internet Research*. **https://doi.org/10.2196/56930**

Heinz, M. V., Mackin, D. M., Trudeau, B. M., Bhattacharya, S., Wang, Y., Banta, H. A., Jewett, A., Salzhauer, A., Griffin, T., & Jacobson, N. C. (2025). Randomized trial of a generative AI chatbot for mental health treatment. NEJM AI, 2(4). **https://ai.nejm.org/doi/full/10.1056/AIoa2400802**



Meer weten



Wil je je nog verder verdiepen? Hieronder vind je een beperkte selectie van toegankelijke websites, artikels en andere bronnen van de projectpartners en organisaties die bij AI-CARES betrokken waren. Ze bieden extra achtergrond, praktijkvoorbeelden en verdieping bij de thema's uit deze praktijkgids.

Over AI-CARES

AI-CARES projectwebsite

Dit handboek bouwt voort op de inzichten en ervaringen uit het onderzoeksproject AI-CARES. Op de projectwebsite vind je meer informatie over de achtergrond, doelstellingen, proeftuinen en resultaten van het project.

https://www.arteveldehogeschool.be/ai-cares

Testomgeving AI-CARES

Zelf aan de slag? In de testomgeving kan je verschillende chatbotprototypes uit het project uitproberen.

https://aicares.ilabt.imec.be/nl/

De projectpartners

Mens, Samenleving en Digitalisering – Arteveldehogeschool

De onderzoeksgroep Mens, Samenleving en Digitalisering bundelt onderzoek, tools en methodieken rond digitale zorg, online en blended hulpverlening, artificiële intelligentie en digitale inclusie. Interessant voor wie AI wil bekijken binnen de bredere digitalisering van welzijn en sociaal werk.

https://msd.arteveldehogeschool.be/nl

Psychologie en technologie – Thomas More-Hogeschool

Deze onderzoekslijn onderzoekt hoe technologie zoals AI, apps, wearables en immersieve technologie verantwoord kan worden ingezet in de geestelijke gezondheidszorg. Je vindt er projecten, tools en praktijkgericht onderzoek over de ontwikkeling, evaluatie en implementatie van digitale toepassingen.

https://thomasmore.be/nl/zorg-en-welzijn-mens-en-welzijn/expertise/psychologie-en-technologie

IDLab – Universiteit Gent

IDLab is een onderzoeksgroep van Universiteit Gent en imec met expertise in onder meer datawetenschap en artificiële intelligentie. Binnen AI-CARES bracht IDLab de technische expertise in voor het ontwikkelen van de chatbotprototypes.

https://idlab.ugent.be/home

Verder lezen

Dossier 'AI in welzijn en zorg' - Onlinehulp Vlaanderen

Een uitgebreid dossier over wat conversationele AI betekent voor welzijn en zorg. Tim Vanhove (Arteveldehogeschool) en Marie Van der Cam (SAM vzw) bespreken kansen en risico's, professioneel AI-gebruik, organisatiebeleid en aandachtspunten bij het ontwikkelen of inzetten van eigen AI-toepassingen.
https://onlinehulp-vlaanderen.be/dossiers/ai-in-welzijn-en-zorg

Dossier AI in welzijn en zorg op Mediawijs-website

In deze reeks van zes toegankelijke artikels verkennen onderzoekers van Artevelde-hogeschool en Thomas More-hogeschool wat AI betekent voor welzijn en zorg.

AI in de hulpverlening: 'Het wordt overal gebruikt. Op de juiste en op de verkeerde manier' - Sociaal.Net

In dit interview bespreekt Tim Vanhove (Arteveldehogeschool) wat AI concreet kan betekenen voor sociale professionals en hulpzoekers. Het artikel gaat onder meer over de betrouwbaarheid van chatbots, professionele verantwoordelijkheid, privacy, organisatiebeleid, guardrails en het gebruik van betrouwbare kennisbronnen.
https://sociaal.net/achtergrond/ai-in-de-hulpverlening-het-wordt-overal-gebruikt-tim-vanhove/

'Praat jij met een AI-therapeut?' – Thomas More-hogeschool

Een toegankelijk artikel met Tom Van Daele en Dries Van Craen (Thomas More-Hogeschool) over hun experiment met een AI-versie van psychiater Dirk De Wachter. Het toont heel concreet hoe overtuigend een AI-therapeut kan lijken en bespreekt de risico's rond schijnempathie, privacy, foutieve antwoorden en het vervangen van menselijke hulp.
https://stories.thomasmore.be/praat-jij-met-een-ai-therapeut

Chatbots als mentale steun: wat kan AI voor je doen en waar moet je op letten? - Leef

Een korte en zeer toegankelijke bijdrage waarin Tom Van Daele (Thomas More-Hogeschool) uitlegt waarom mensen chatbots gebruiken voor mentale ondersteuning, wat zulke systemen mogelijk kunnen betekenen en waar hun grenzen liggen.
https://www.cm.be/leef/mentaal-welzijn/chatbots-als-mentale-steun

'Artificiële intelligentie en de klinisch psycholoog: zegen of vloek?' – Fien Buelens, Tom Van Daele en Sylvie Bernaerts

Deze narratieve review biedt een toegankelijke wetenschappelijk onderbouwde introductie tot AI in de geestelijke gezondheidszorg. De auteurs bespreken zowel kansen en risico's als concrete toepassingen voor klinisch psychologen, met aandacht voor onder meer chatbots, privacy, wetgeving en menselijke verantwoordelijkheid.
https://tkp-psychologie.eu/artificiele-intelligentie-en-de-klinisch-psycholoog-zegen-of-vloek

www.arteveldehogeschool.be/ai-cares