

Verantwoord gebruik van generatieve AI in de Zorg

Bettine Pluut | Strategisch adviseur | Informatieveiliggedrag in de zorg
Chris van Zwolle | Product owner kennisproducten | Z-CERT



Informatieveilig
gedrag **in de zorg**



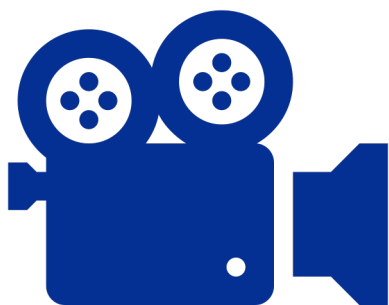
Platform voor de
InformatieSamenleving



6 februari 2025

Welkom!

We nemen het webinar op



Wil je niet in beeld?
Zet je camera uit!



Zet je microfoon uit

Programma

1. Waarom een webinar over AI-chatbots?
2. Chris van Zwolle (Z-CERT): De techniek achter AI-chatbots
3. Bettine Pluut (IVGZ): Hulpmiddelen voor AI-chatbots op de werkvloer
4. Vragenronde

Enkele vragen aan jullie

Generatieve AI

Een kijkje achter de schermen



COMPUTER EMERGENCY
RESPONSE TEAM
VOOR DE ZORG





De Nederlandse zorg digitaal veilig

Wij zijn Z-CERT, het expertisecentrum voor cybersecurity in de zorg. Onze missie is de zorg digitaal veiliger maken. Dit doen wij door zorginstellingen weerbaar te maken tegen cybercriminelen.

Z-CERT is vanuit de Cyberbeveiligingswet aangewezen als de sectorale CERT voor de zorg.



www.z-cert.nl

Wat zit er achter die telefoon?



Bron: Bright.nl, 4 februari 2025

AI is niet vanzelf veilig, dat moeten wij doen

Een AI praatplaat



Wet en regelgeving

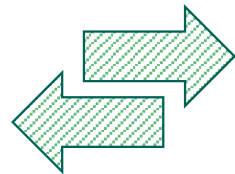
Organisatie



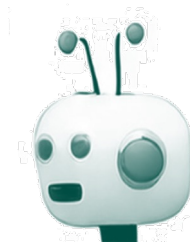
Keten van leveranciers



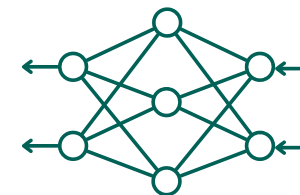
Gebruiker



Vraag & Antwoord



Programma's



Model (LLM)



Trainingsdata

Leren



Wat er gebeurt met jouw input

AVG

Wet en regelgeving

Opslag in meerdere clouds

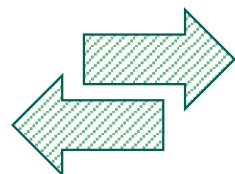
Organisatie



Keten van leveranciers



Antwoord!



Gebruiker

Vraag?

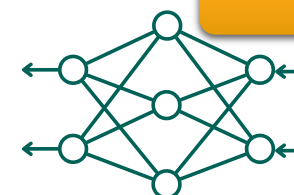
Vraag & Antwoord

Programma's

Verwerkt privacy data

Leren

Historie van vragen en antwoorden worden vastgelegd



Model (LLM)

Verwerkt privacy data



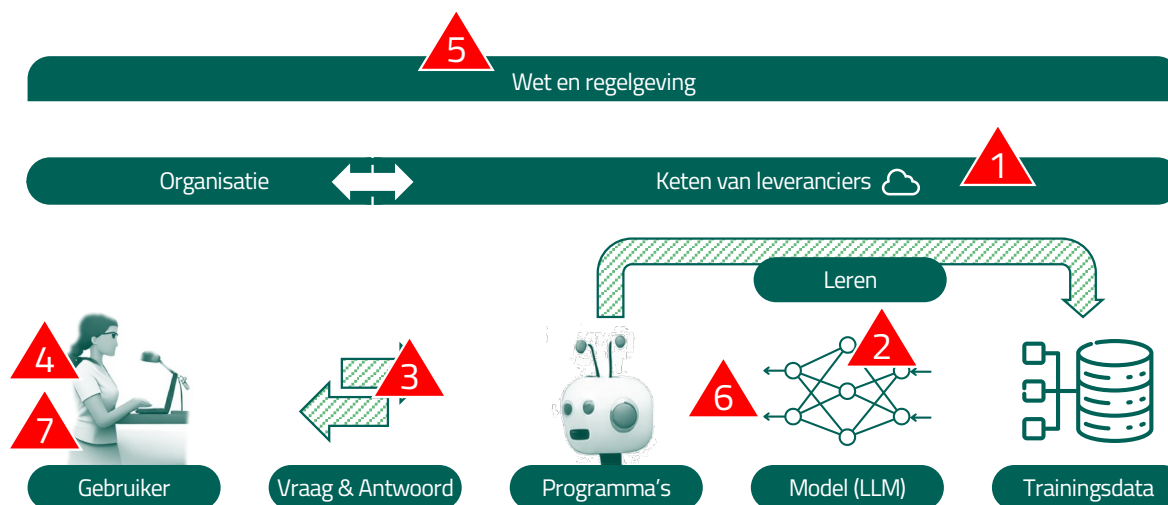
Trainingsdata

Leren



Enkele risico's bij gebruik van AI, slechts een selectie

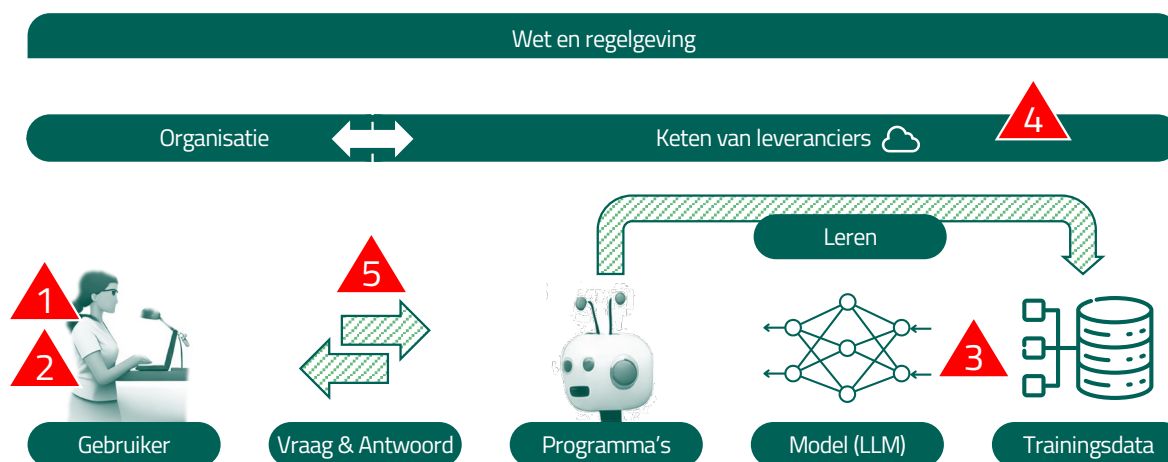
1. **Datalek in de keten.** Je hebt maar zeer beperkt grip op je keten, terwijl je werkt met zeer gevoelige informatie.
2. **Bias.** De AI wordt vaak als expert gezien, terwijl elke AI een bias heeft en de bronnen ook. Een neutrale samenvatting bestaat niet.
Garbage in = garbage out.
3. **Context verwarring.** Een AI koppelt latere vragen aan eerdere, waardoor de context van de vraag verandert. Dit gebeurt ook als de context onvolledig wordt meegegeven.
4. **Profilering van de gebruiker.** Uit combinaties van zoekopdrachten ontstaat een beeld. Eventueel samen met internetgedrag.
5. **Niet compliant.** Wetgeving ontwikkelt zich snel en de keten is heel internationaal. Hoe weet je dat jouw data compliant wordt verwerkt?
6. **Filtering.** Bij sommige AI's worden de antwoorden na de LLM gefilterd, denk aan DeepSeek. Vertrouw je naast de AI ook de hele context?
7. **Afhankelijkheid.** Hoeveel afhankelijk van deze ondoorzichtige systemen is verantwoord?





Enkele risico's bij een aanval, een selectie

1. **Prompt Code Injection.** Via de prompt kan een aanvaller code invoeren en die code laten uitvoeren door de achterliggende systemen. Dat kan leiden tot heel veel narigheid!
2. **Discovery.** Een aanvaller achterhaalt persoonsinfo door de 'juiste' vragen te stellen en legt zo relaties die je zelf wilt voorkomen.
3. **Data Poisoning.** Een aanvaller kan de trainingsdata vervuilen, waardoor de LLM andere conclusies trekt.
4. **Data theft.** Door het massale gebruik van AI ontstaan ook data goudmijnen.
5. **AI in the Middle.** Met wie praat je eigenlijk?



Meer informatie op:
atlas.mitre.org/matrices/ATLAS



Wie kan er wat aan doen?

Gebruikers

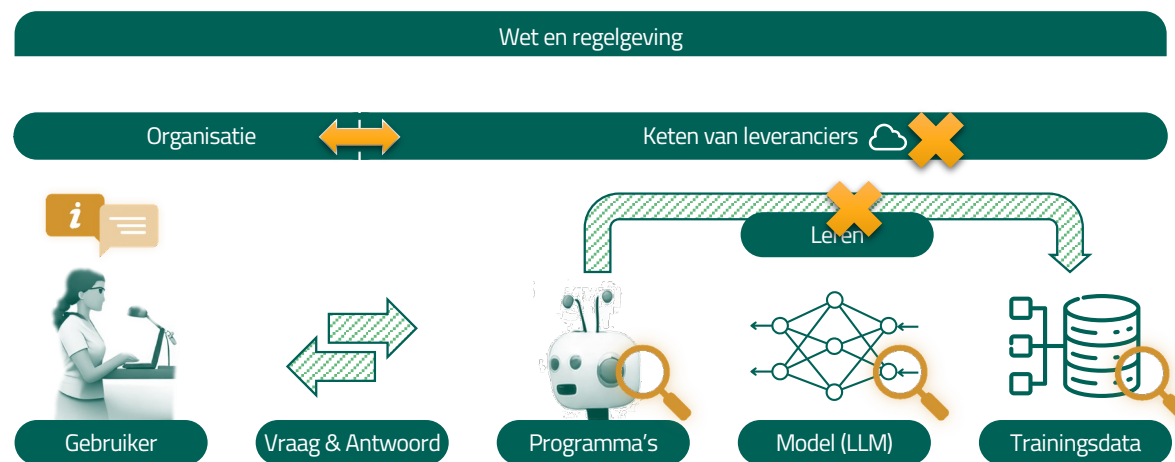
- Bepaalde AI programma's *niet* gebruiken
- Goedgekeurde AI programma's *goed* gebruiken

Organisaties

- Maak duidelijk wat wel en niet kan. Biedt een goed alternatief.
- Zoveel mogelijk aanbieden vanuit de eigen IT omgeving
- Pentesten van de AI omgeving

Leveranciers

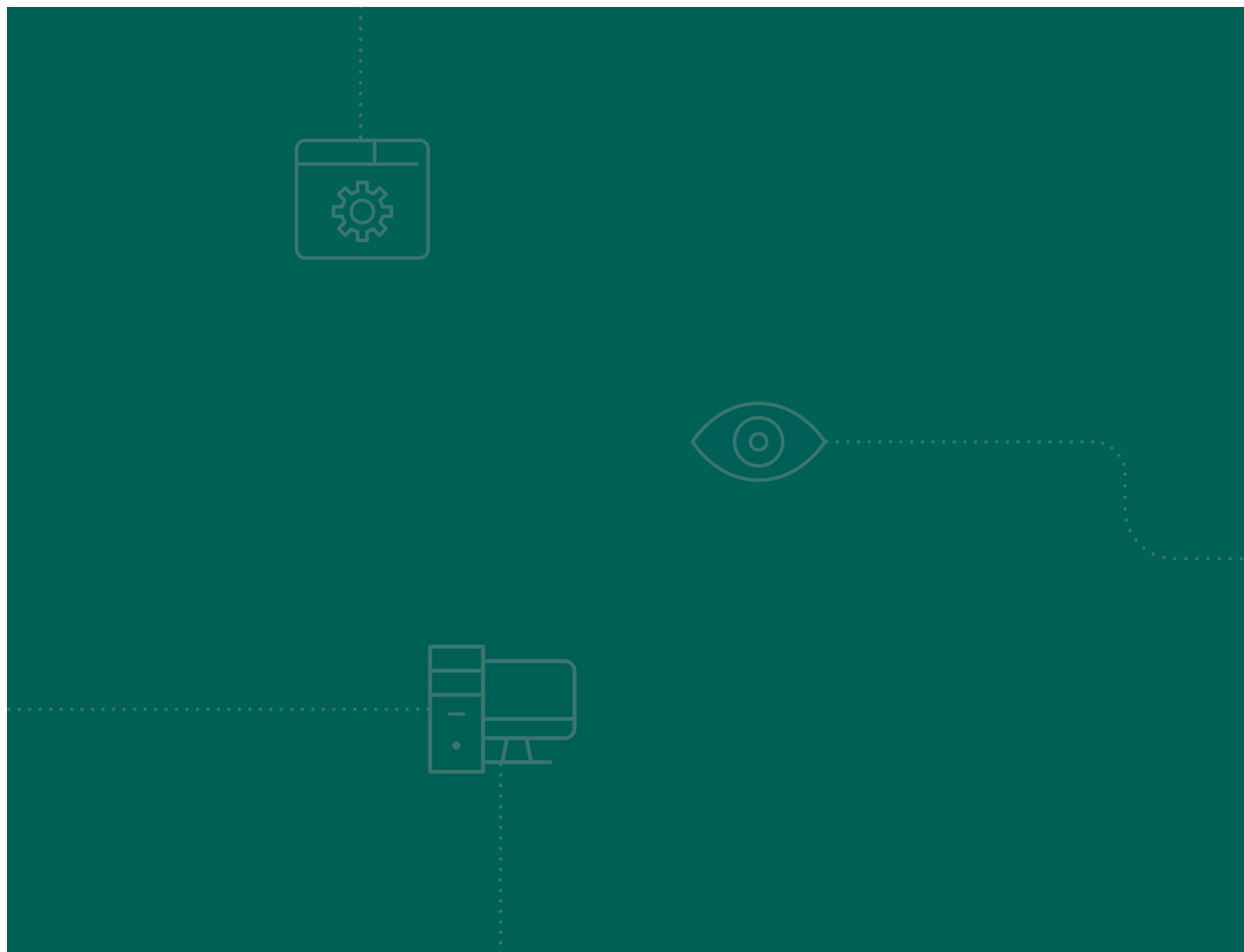
- Afspreken dat de leverancier geen data opslaat
- Afspreken dat jouw vragen niet worden gebruikt om het model te leren
- Procesafspraken rondom data of bias correcties



Let op:
Voor **AI security** geldt reguliere **IT security**!



Stichting Z-CERT
www.z-cert.nl



Hulpmiddelen AI-chatbots op de werkvloer



Informatieveilig
gedrag **in de zorg**



Platform voor de
InformatieSamenleving

Bettine Pluut - 6 februari 2025

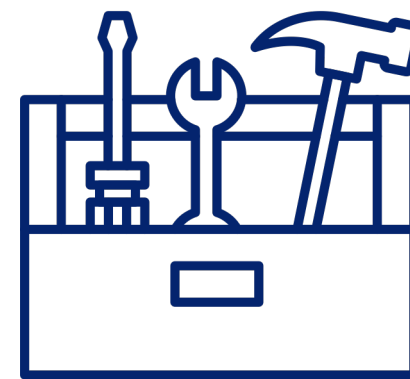
Programma Informatieveilig Gedrag in de Zorg



Informereren, inspireren en activeren



De gedragscomponent van veiligheid centraal



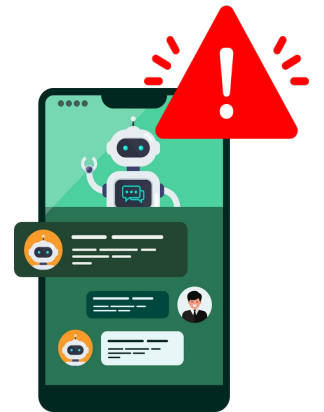
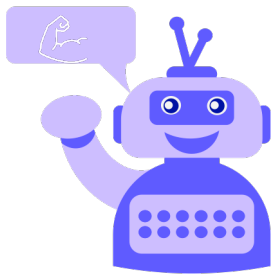
Gratis toegankelijke trainingen webinars en hulpmiddelen

“

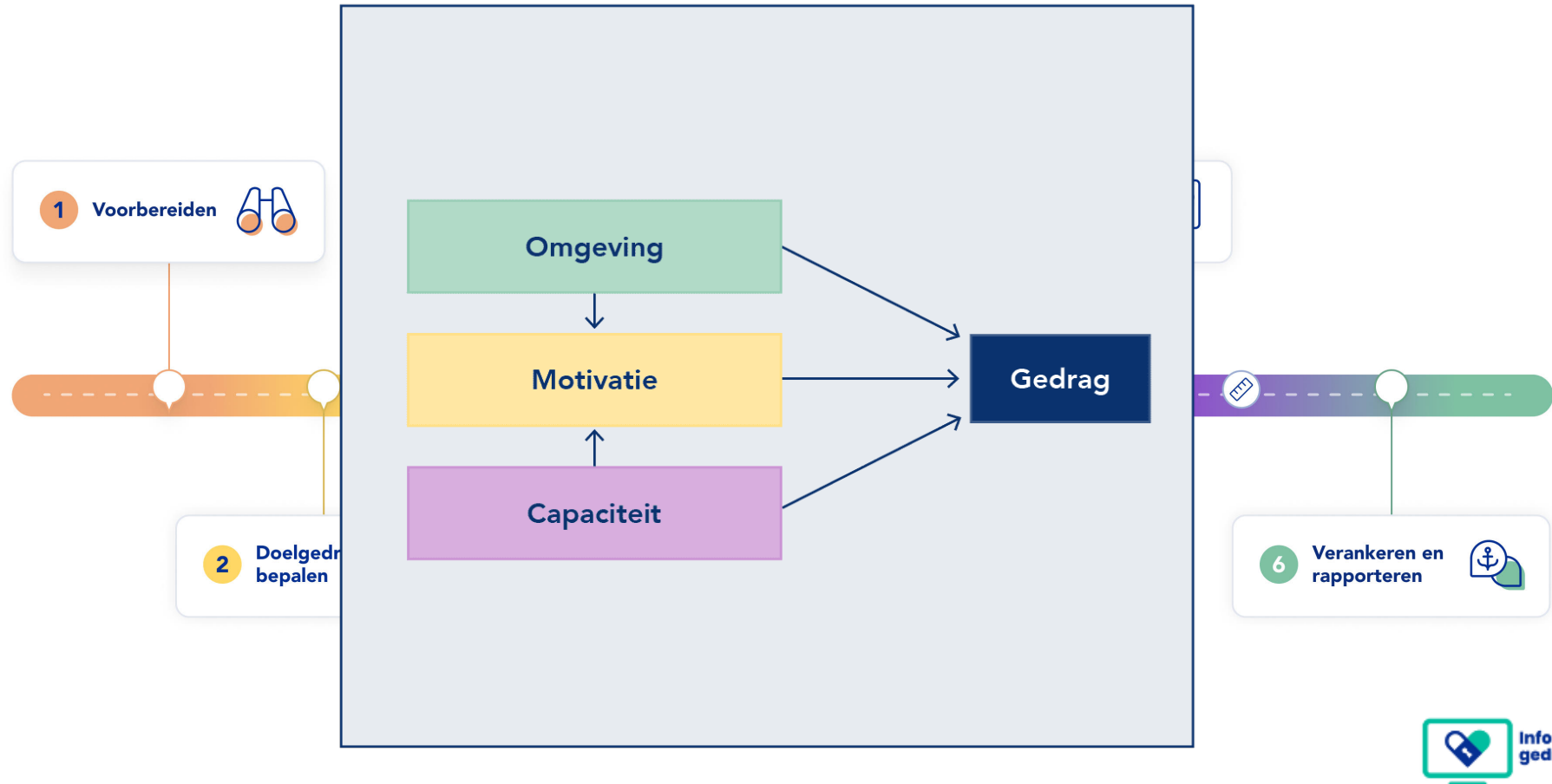
“Het voelt kwetsbaar”

”





De Wegwijzer



Van Wegwijzer naar hulpmiddel



Uitgangspunten:

- Sámen een bruikbaar resultaat in drie bijeenkomsten
- Collectieve denkkraft voor gemeenschappelijk probleem
- Besluitvorming met deep democracy

Praktisch

- 12 - 14 deelnemers
- Online bijeenkomsten
- 4 - 5 maanden
- Testen onder de doelgroep en experts

AI-chatbots zijn handig, gebruik ze verstandig



Informatieveiligheidsprofessionals



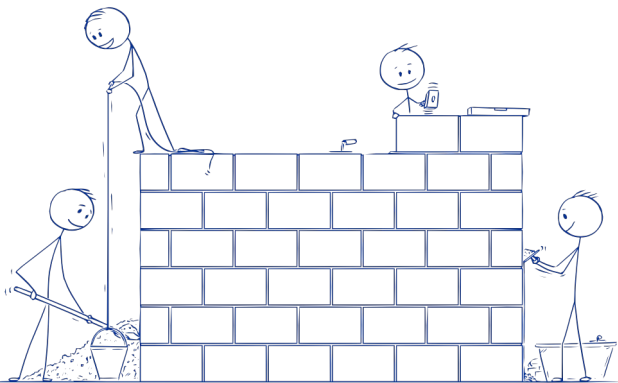
Zorgmedewerkers

Infographic voor de veiligheidsprofessional

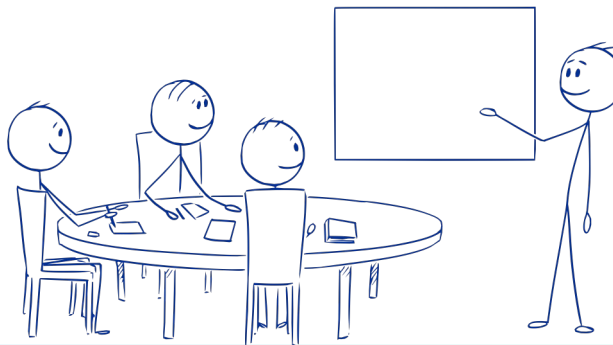
Hoe kun je informatieveilig gedrag stimuleren?



Bouw een muur



Bespreek in werkoverleg



Jaag gesprekken aan



Infographic voor zorgmedewerkers

Wat staat er in deze infographic?



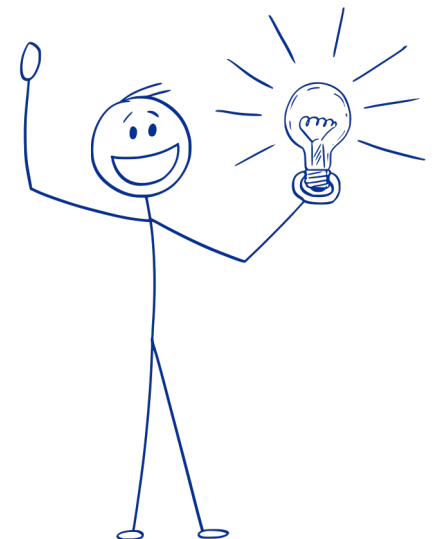
Tips voor gebruik

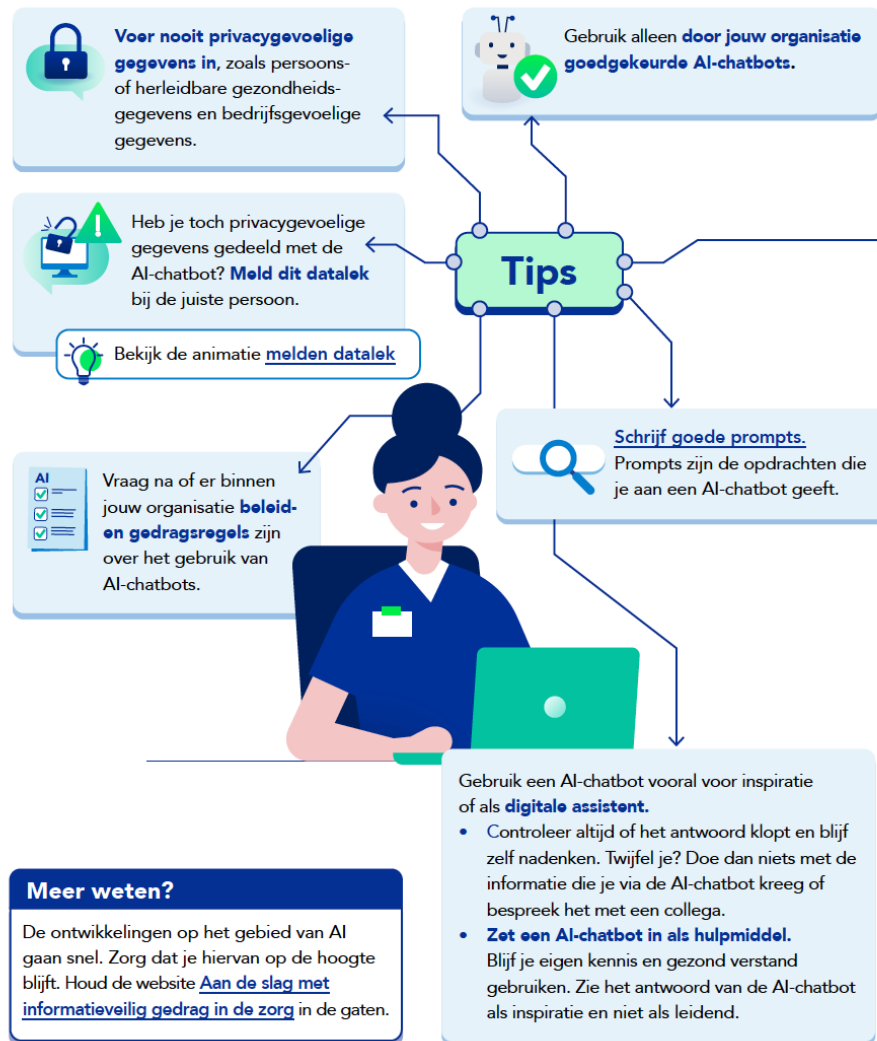


Datalek gemeld, leren is wat telt

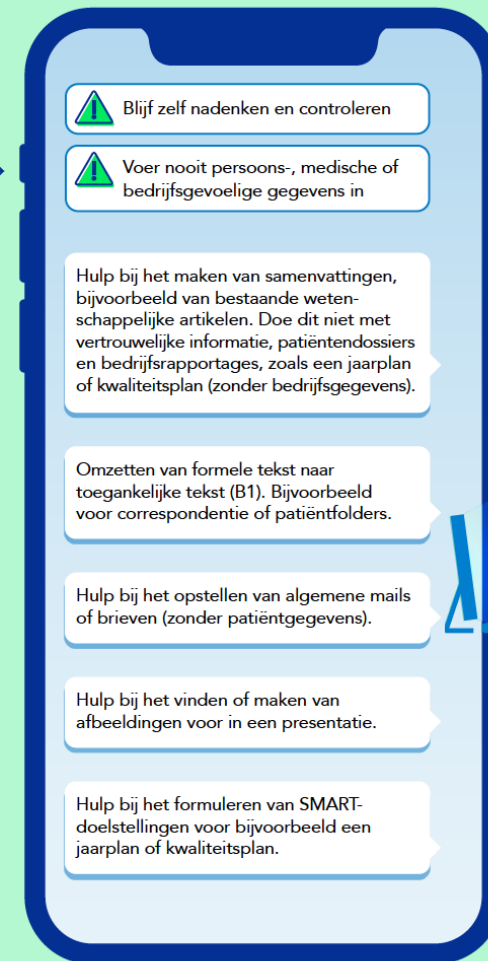


Waar kan ik het voor gebruiken?





Waar kan je AI-chatbots wél bij gebruiken?



Vragen

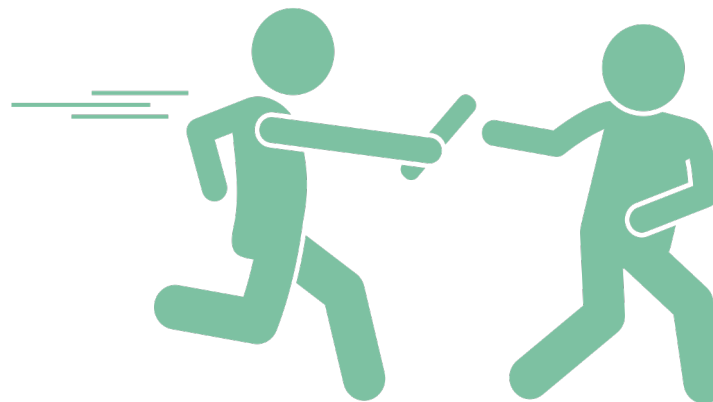


Levende case study



Delen van ervaringen

***Doe jij mee?
Meld je aan!***



Van beginnend tot in beweging

Meer weten of meedoen? Dat kan!

- Volg ons op **LinkedIn** en/of **meld je aan voor de nieuwsbrief**
 - www.informatieveiliggedragzorg.nl
- Meld je aan om mee te doen:
 - **Werkgroep:** Eerste bijeenkomst op maandag 10 maart (online)
 - **Cyberveilig in 1 dag:** 13 maart in Amersfoort
 - **Masterclass:** Eerste bijeenkomst op 31 maart
- Kijk op ons **Youtube** kanaal voor de uitleg over hoe je een **gedragsanalyse met de COM-B methode** doet
 - https://youtu.be/3hwVRH_PMpk?si=KivunWF9Ds1vr2pC



Bedankt voor jullie aandacht!



Informatieveilig
gedrag **in de zorg**



Platform voor de
InformatieSamenleving

