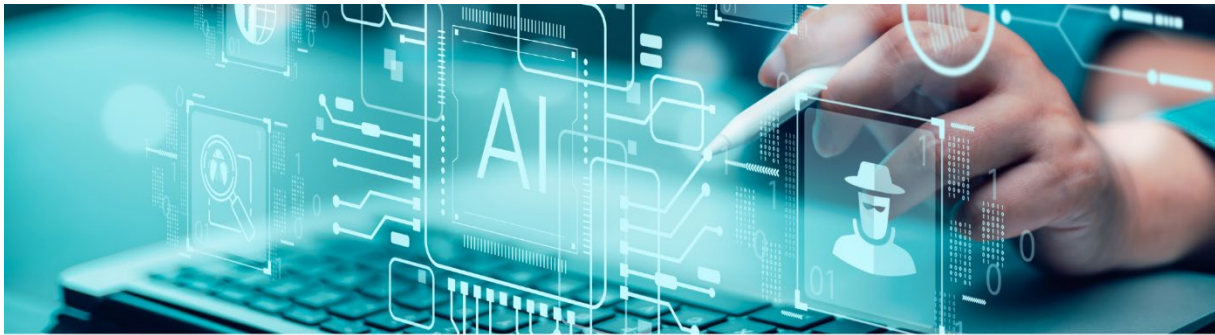




Break-out sessie: Generative AI & GDPR

Onder leiding van Julius Obdeijn (Legal Advisor Enforcement, AP) en Matthijs Baart (Senior Adviseur, AP)

Tijdens deze interactieve sessie verkenden deelnemers de relatie tussen generatieve AI en de AVG aan de hand van drie fictieve casussen. De sessie begon met een overzicht van de AI-keten – van dataverzameling tot toepassing – en behandelde juridische, ethische en praktische implicaties van AI-gebruik op persoonsgegevens. Thema's als transparantie, grondslagen voor verwerking, profilering en menselijke controle kwamen uitgebreid aan bod.

In de eerste casus werd besproken hoe het hergebruik van klantenservice-chatlogs voor AI-training juridische risico's met zich meebrengt, vooral vanwege het ontbreken van transparantie en de moeilijkheid van volledige anonimisering. De tweede casus richtte zich op AI in werving en selectie, waarbij zorgen werden geuit over geautomatiseerde besluitvorming zonder menselijke tussenkomst, gebrek aan transparantie en risico's op bias. De derde casus betrof het samenvatten van patiëntendossiers in de zorg. Hoewel dit potentieel waardevol is, bleek het juridisch en technisch complex vanwege de gevoeligheid van medische gegevens en het black box-karakter van AI-modellen.

De sessie maakte duidelijk dat verantwoord gebruik van generatieve AI binnen de kaders van de AVG vraagt om heldere afspraken, transparantie richting betrokkenen, en waarborgen zoals "human in the loop". Deelnemers benadrukten het belang van praktische guidance, vooral in sectoren als HR en zorg, waar de impact van AI groot is.

Casus 1: Finetuning met chatlogs van klantenservice

Achtergrond: Een bedrijf overweegt om chatlogs van de klantenservice te gebruiken voor het trainen van een generatief AI-model dat klantvragen zelfstandig kan beantwoorden. Deze logs bevatten echter persoonsgegevens en zijn nooit met dat doel verzameld. Klanten zijn hierover niet geïnformeerd.

Belangrijkste observaties:

1. Klanten worden niet geïnformeerd: gebrek aan transparantie.

2. Anonimisering is wenselijk, maar nooit volledig haalbaar.
3. De gegevens worden hergebruikt, dus in die zin kan je niet met een nieuwe grondslag het probleem oplossen, omdat het gaat om hergebruik van gegevens. Gesprekken opnemen voor trainingsdoeleinden kan, mits juridisch afgedekt.
4. Verwijderen van data bij generatieve AI is bijzonder moeilijk.
5. Belangrijke afweging: gerechtvaardigd belang versus privacy van betrokkenen.
6. Toestemming kan worden ingetrokken; dit maakt het juridisch complex

Casus 2: Generatieve AI in werving & selectie

Achtergrond: Een organisatie wil AI inzetten om het selectieproces van kandidaten te versnellen. Het model maakt op basis van CV's en motivatiebrieven een beoordeling van de kandidaat. Er is geen menselijke tussenkomst en kandidaten worden niet geïnformeerd over de inzet van AI. Verder, worden de persoonsgegevens niet gebruikt om het AI-model verder te trainen.

Belangrijkste observaties:

1. Persoonsgegevens (zoals op CV's) mogen niet zomaar verwerkt worden, toestemming is vereist.
2. Transparantie richting sollicitanten ontbreekt volledig.
3. Geautomatiseerde besluitvorming zonder menselijke tussenkomst is juridisch en ethisch problematisch.
4. Het AI-model is niet getraind op interne data, waardoor relevantie en biasrisico's groot zijn.
5. Onduidelijkheid over waar de data naartoe gaan.
6. Profileren op basis van persoonlijkheidskenmerken is AVG-gevoelig en niet ethisch.
7. Eerlijkheid en gelijke behandeling van kandidaten staan ter discussie.
8. Voorwaarden voor geautomatiseerde AI-besluitvorming: human in the loop én bezwaaroctie.

Casus 3: De inzet van generatieve AI voor het maken van samenvattingen van patiëntendossiers

Achtergrond: In de zorg wordt overwogen om generatieve AI in te zetten voor het samenvatten van patiëntendossiers, met als doel om artsen en zorgverleners te ondersteunen in hun werkzaamheden. De AI zou echter ook getraind worden op deze gegevens.

Belangrijkste observaties:

1. Voor het samenvatten van patiëntendossiers zouden deze volledig geanonimiseerd moeten zijn. Anonimisering van patiëntgegevens is zeker bij verdere training van de AI-chatbot cruciaal, maar zeer complex. Hoe krijg je die gegevens eruit als het misgaat?
2. AI-modellen zijn moeilijk te controleren of te corrigeren; black box-probleem.
3. Gebruik van AI voor samenvatten moet passen binnen het oorspronkelijke zorgdoel: zorg verlenen.
4. AI-modellen zijn moeilijk te controleren of te corrigeren; black box-probleem.
5. Patiënten zouden al vroeg in het proces betrokken moeten worden, niet pas achteraf.

6. Tip uit publiek: het betrekken van een derde partij met zelf afgesloten datasets, die je achteraf gewoon kan wissen, in eigen beheer.
7. Er is nood aan heldere afspraken over opslag, gebruik en verwijdering van data. Vooral in de zorg waar een datalek al snel grote gevolgen kan hebben.