mbo digitaal

NETWERK Informatiebeveiliging en Privacy

Verantwoord omgaan met AI en de FRIA

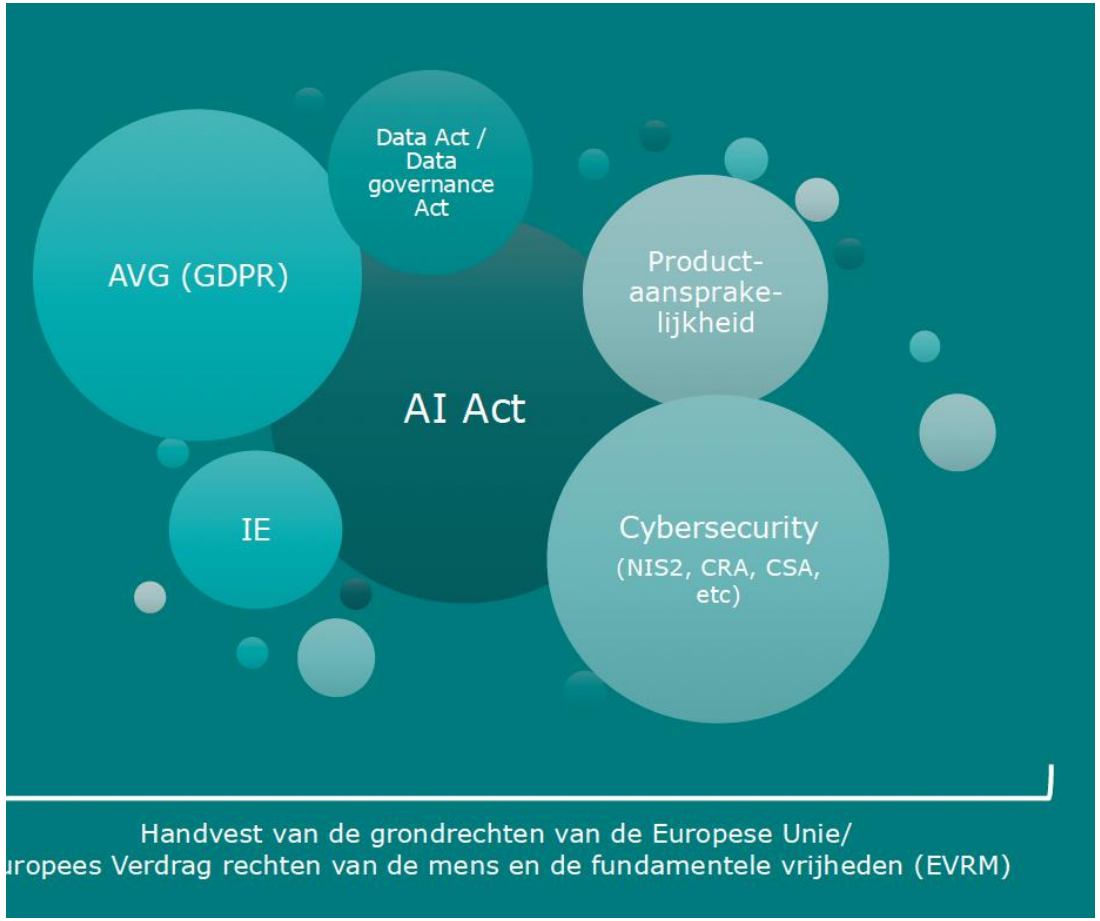
Niels Dutij

Voorstellen

- Niels Dutij
- Adviseur MBO Digitaal bij Programma Cyberveiligheid
- Functionaris voor gegevensbescherming bij meerdere MBO’s

mbo digitaal

Overzicht van de wetgeving



- Veel verschillende soorten wetgeving omtrent data en technologie
- AI-ACT bestaat naast de AVG

Wat is AI?

Wat is een AI-systeem?

“Een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na het inzetten ervan **aanpassingsvermogen kan vertonen**, en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren zoals **voorspellingen, inhoud, aanbevelingen of beslissingen** die van invloed kunnen zijn op fysieke of virtuele omgevingen.”

mbo digitaal

FRIA en DPIA

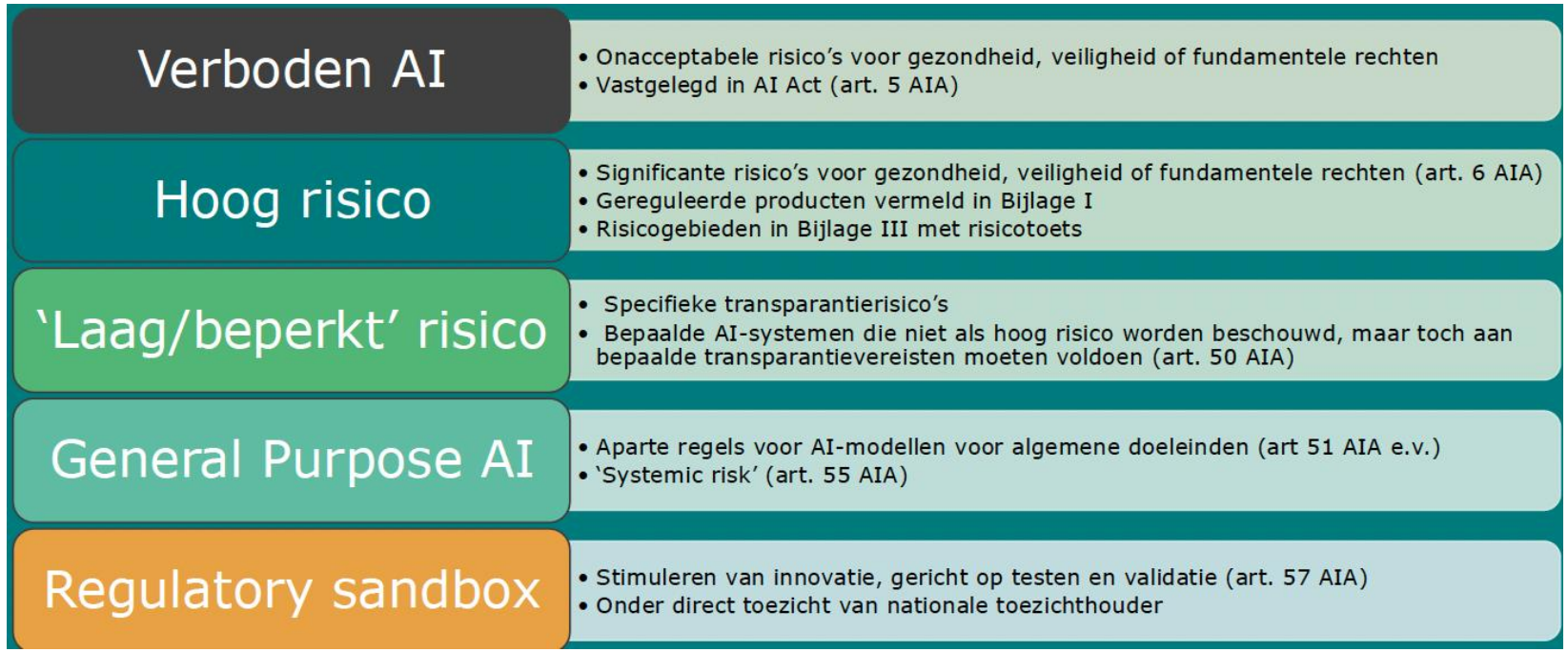
Fundamental Rights Impact Assessment (FRIA)

- Inschatting en mitigatie van risico's voor (grond)rechten en maatschappij bij hoog risico AI-systeem.
- Uitgevoerd door deployers in overheid, semi-overheid, kredietinstellingen en levens/gezondheid verzekering.
- Herhalen indien verouderd.

Data Protection Impact Assessment (DPIA)

- Inschatting en mitigatie van risico's voor privacy en (grond)rechten.
- Uitgevoerd door verwerkingsverantwoordelijke.
- Opnieuw bij veranderde risico's.
- In de AIA wordt verwezen naar artikel 35 AVG. Je mag de resultaten van een DPIA meenemen in een FRIA.

Verschillende soorten AI



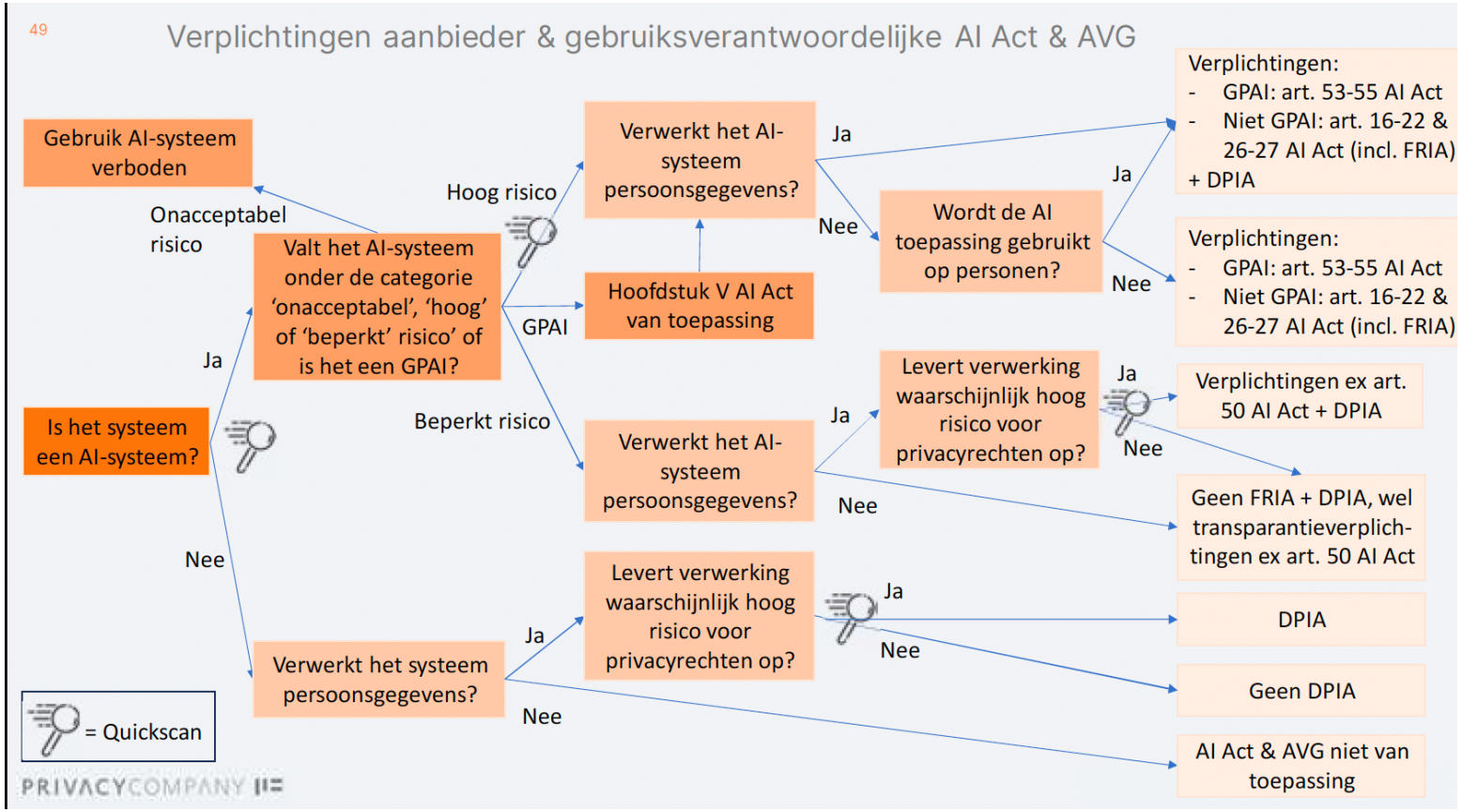
mbo digitaal

Verplichtingen hoge risico's

- *Hoog risico* voor gezondheid, veiligheid, grondrechten van personen of voor het milieu
- Classificatie: formeel proces (art. 6 t/m 15 + Bijlage I + III)
  - AI is veiligheidscomponent van gereguleerd product (Bijlage I): liften, speelgoed, medische apparaten, luchtvaart, auto's.
  - AI wordt gebruikt in hoogrisico *use cases* (Bijlage III): kritieke infrastructuur, werkgelegenheid, rechtshandhaving, migratie en onderwijs en beroepsopleiding.
  - AI wordt gebruikt voor *profiling*
  - Uitzonderingen indien AI 'louter ondersteunend' is
- Strenge eisen voor AI-systeem
- Verplichtingen voor zowel aanbieders als gebruiksverantwoordelijken, maar meeste voor aanbieders

mbo digitaal

Overzicht verplichtingen



mbo digitaal

Fundemental Human Rights Impact Assessment (FRIA)

1. een beschrijving van de processen van de exploitant waarin het risicovolle AI-systeem zal worden gebruikt in overeenstemming met het beoogde doel;
2. en beschrijving van de periode en frequentie waarin elk risicovol AI-systeem bedoeld is om te worden gebruikt
3. de categorieën natuurlijke personen en groepen die waarschijnlijk gevolgen zullen ondervinden van het gebruik ervan in de specifieke context;
4. de specifieke risico's op schade die waarschijnlijk gevolgen zullen hebben voor de categorieën personen of groepen personen die zijn geïdentificeerd overeenkomstig punt c), rekening houdend met de informatie die door de provider wordt verstrekt overeenkomstig artikel 13;
5. en beschrijving van de implementatie van maatregelen voor menselijk toezicht, volgens de gebruiksaanwijzing;
6. de maatregelen die moeten worden genomen in het geval dat deze risico's zich voordoen, met inbegrip van de regelingen voor interne governance en klachtenmechanismen.

Hoog risico in het onderwijs

- Toegang, toelating en toewijzing tot onderwijs
- Evalueren van leerresultaten en sturen van het leerproces
- Beoordelen van het passend onderwijsniveau
- Monitoren en detecteren van ongeoorloofd gedrag tijdens toetsen

Casus: Geautomatiseerde examinering

Een examenplatform biedt een geweldige mogelijkheid voor het afnemen van mondelinge examens. Studenten spreken de opdracht in, waarbij het AI-platform tot een score komt.

Dit proces is volledig geautomatiseerd. Docenten en examinatoren krijgen enkel de uitslag, maar hebben geen toegang tot het gemaakte examenwerk.

Er is wel door de medewerkers van het examenplatform de mogelijkheid voor een review van gemaakt werk als examen. Dit is echter reactief.

Casus DPIA

Wat zijn de pijnpunten onder de AVG bij dit platform?

- Verwerking van persoonsgegevens
- Geautomatiseerde besluitvorming, DPIA verplicht
- Wat komt er uit de DPIA als mogelijke problemen?

Risico 1: Recht op gelijke behandeling

Is er bias in het AI-model (accenten, taalgebruik, spraakstoornissen)?

Worden kwetsbare groepen (bv. dyslectische of anderstalige studenten) benadeeld?

Vereisten AI-ACT

- AI-systemen moeten worden getest op bias en discriminatie.
- De dataset moet representatief zijn om oneerlijke beoordelingen te voorkomen.

Risico 2: Rechtmatigheid verwerking gegevens

- Hoe wordt stemdata verwerkt en opgeslagen?
- Geautomatiseerde besluitvorming
- Kunnen studenten hun gegevens inzien en laten verwijderen?

Vereisten AVG en AI-ACT
Studenten moeten **inzage krijgen** in hun gegevens.
Er moet een duidelijke **grondslag** zijn voor het verwerken van biometrische data.
Dataminimalisatie: Data mag niet langer bewaard worden dan strikt noodzakelijk.

mbo digitaal

Risico 3: Recht op een eerlijk proces (due process)

- Kunnen studenten hun beoordeling inzien en aanvechten?
- Is er een menselijke controle voordat beslissingen definitief zijn?

Riscio 4. Transparantie

Transparantie en uitlegbaarheid

- Black-box karakter: Studenten en docenten krijgen geen inzicht in hoe de AI tot een score komt.
- Gebrek aan controle: Docenten kunnen de beoordeling niet valideren controleren.
- Gebrek aan uitleg voor studenten: Studenten hebben geen toegang tot hun ingesproken examen of de beoordeling, wat in strijd kan zijn met het recht op uitleg.

Vereisten onder de AI Act:

- De AI moet begrijpelijk en uitlegbaar zijn.
- Studenten en docenten moeten kunnen begrijpen hoe beslissingen tot stand komen.

mbo digitaal

Risico 5 Aansprakelijkheid en verantwoordelijkheid

Aansprakelijkheid en verantwoordelijkheid

- Wie is verantwoordelijk bij fouten (de school of het platform)?
- Wat zijn de gevolgen voor studenten als ze door een AI-fout onterecht zakken?

Welke maatregelen zien we hiervoor?

--> Is de inzet van dit examenplatform mogelijk en onder welke voorwaarden?

Bedankt voor jullie aandacht!