HOE KAN IK MINIMAAL GEGEVENS VERWERKEN?

Het principe van minimale gegevensverwerking houdt in dat persoonsgegevens **toereikend** moeten zijn, **ter zake dienend** zijn en **beperkt** zijn **tot wat noodzakelijk is** voor de doeleinden waarvoor zij verwerkt worden. Zij mogen enkel worden verwerkt indien het doel van de verwerking niet redelijkerwijze op een andere wijze kan worden bereikt.

Indien een gelijkaardig resultaat dus bereikt kan worden zonder (bepaalde) persoonsgegevens te gebruiken, mogen die persoonsgegevens daarvoor niet worden gebruikt.

Minimale gegevensverwerking moet ruim bekeken worden en geldt niet enkel bij het **verzamelen van persoonsgegevens**, maar ook bij de **interne toegang** en het **gebruik** van persoonsgegevens.

ACTIEPUNTEN

- Ken de **herkomst**, de eraan verbonden **rechten**, de **eigenschappen** en de **attributen** van de gegevens om te weten …
  - of ze correct verzameld werden en;
  - voor welke doeleinden ze gebruikt mogen worden.
- Bepaal vooraf welke **types persoonsgegevens** en welk **volume aan persoonsgegevens** strikt noodzakelijk zijn om het AI-systeem te kunnen trainen en/of gebruiken. Werk hiervoor multidisciplinair, motiveer en documenteer.
- **Evalueer** of technieken en/of systemen kunnen worden toegepast die toelaten om …
  - met minder persoonsgegevens te werken;
  - met versleutelde persoonsgegevens te werken of;
  - methoden te gebruiken zoals *Generative Adversarial Networks* (GAN's), federated learning of transfer learning, die minder gegevens nodig hebben om te leren.
- Zet een performant **pseudonimiserings - en anonimiseringsbeleid** op.
- Kuis **datasets** regelmatig op en verwerk gegevens niet langer dan noodzakelijk.
- **Documenteer en registreer** elke stap en iedere evaluatie van de AI-toepassing, conform de verantwoordingsplicht.

Kenniscentrum Data & Maatschappij

Kenniscentrum Data & Maatschappij (2021). Fiche: Hoe kan ik minimaal gegevens verwerken? Brussel: Kenniscentrum Data & Maatschappij.

1 HEB IK WERKELIJK (ECHTE) PERSOONSGEGEVENS NODIG?

Je mag persoonsgegevens enkel verwerken indien het doel van de verwerking niet redelijkerwijze op een andere manier kan worden verwezenlijkt.

Ga dus na of je het AI-systeem ook kan trainen met **andere gegevens**, bijv. met synthetische persoonsgegevens (die persoonsgegevens nabootsen) of geanonimiseerde gegevens.

2 KAN DE VERWERKING GEBEUREN MET MINDER PERSOONSGEGEVENS?

Ga na welke gegevens **echt nodig** zijn om het vooropgestelde doel te bereiken. Hou hierbij rekening met twee elementen:

- Over hoeveel personen worden gegevens verwerkt;
- Om hoeveel soorten persoonsgegevens gaat het.

Ga na of het gebruik van de beoogde persoonsgegevens **proportioneel** is ten aanzien van het beoogde doel en ten aanzien van de risico's die de betreffende personen daarbij lopen.

Analyseer en bepaal op voorhand hoeveel en welke persoonsgegevens vereist zijn voor het verwerkingsdoel. Documenteer dit.

3 KAN IK HET VOLUME AAN (CENTRAAL) VEREISTE PERSOONSGEGEVENS IN DE TRAININGSFASE VERMINDEREN?

Je hebt verschillende **mogelijkheden** om het volume aan (centraal) vereiste persoonsgegevens **in te trainingsfase** te verminderen:

- Leg minimale gegevensbescherming op als productvereiste bij aankoop of opdracht tot ontwikkeling van een AI-systeem waarmee zelf nog getraind zal worden;
- Evalueer welke aanwezige parameters in een dataset noodzakelijk zijn voor het trainingsproces en verwijder de andere parameters vooraleer de dataset te gebruiken;
- Gefedereerd leren laat toe om op verschillende databases op lokale toestellen te trainen, zonder dat de persoonsgegevens het lokale toestel verlaten;
- Door gebruik te maken van een GAN of *Generative Adverserial Network* wordt het vereiste volume aan (persoons-)gegevens voor training van een AI-systeem verminderd;
- Dankzij overdrachtsleren ('*transfer learning*') leer je van nieuwe (persoons-)gegevens, maar vertrek je van een bestaand model van een andere toepassing.

4 KAN IK HET VOLUME AAN (CENTRAAL) VEREISTE PERSOONSGEGEVENS IN DE GEBRUIKSFASE VERMINDEREN?

Je hebt ook nog een aantal **mogelijkheden** om het volume aan (centraal) vereiste persoonsgegevens **in de gebruiksfase** te verminderen zoals:

- minimale gegevensverwerking als productvereiste;
- minimale gegevensbescherming als productvereiste bij aankoop of opdracht tot ontwikkeling van een AI-systeem;
- persoonsgegevens-beschermende verzoeken (*privacy preserving queries* of P2Q).

5 BEWAAR GEGEVENS NIET LANGER DAN NODIG.

Je moet **persoonsgegevens verwijderen of anonimiseren van zodra** zij **niet meer nodig** zijn voor de doeleinden waarvoor zij verzameld werden.

- Bepaal bewaartermijnen voor elk type van gegevens. Indien dit niet mogelijk is, bepaal dan de parameters waarmee de termijn kan worden bepaald.
- Overweeg regelmatig of de verwerkte persoonsgegevens nog wel nodig zijn, en zo niet, verwijder en/of anonimiseer ze.
- Voor archiverings-, onderzoeks- of statistische doeleinden kunnen persoonsgegevens wel langer bewaard worden.
- Houd er rekening mee dat personen het recht hebben om hun gegevens in bepaalde omstandigheden te laten verwijderen. Schat in welke impact dit kan hebben op de werking, ontwikkeling en uitrol van de AI-toepassing.
- Bewaringstermijnen hebben een zeer technisch karakter, werk daarom samen met andere afdelingen binnen de organisatie.

6 KAN IK PERSOONSGEGEVENS AL VERNIETIGEN OF ANONIMISEREN?

Wanneer het niet langer noodzakelijk is om persoonsgegevens aan de betrokken personen te kunnen koppelen, heb je **twee mogelijkheden**:

- **Vernietiging:** de gegevens worden vernietigd op alle plaatsen van bewaring (ook gegevens in de back-up).
- **Anonimisering:** de AVG beschouwt anonimisering als een vorm van vernietiging doordat de persoonsgegevens niet langer persoonsgegevens zijn en de AVG dus hierop niet langer van toepassing is.

Er bestaan verschillende **methodes** die toelaten om persoonsgegevens te anonimiseren. Elk hebben ze hun voor- en hun nadelen, reden waarom vaak een combinatie van verschillende methodes toegepast wordt. Hieronder zijn enkele voorbeelden van dergelijke **methodes**:

- *Attribute suppression* of het verwijderen van alle meldingen van een bepaald type (identificerende) eigenschappen;
- *Record suppression* of het verwijderen van gegevens met betrekking tot bepaalde opvallende betrokkenen;
- *Character masking* of het verwijderen van bepaalde tekens;
- Generalisering;
- *Swapping, shuffling* of het 'dooreenschudden' van gegevens.

7 KAN IK DE LEESBAARHEID EN DE BRUIKBAARHEID VAN DATASETS VOOR DERDEN VERMINDEREN?

Je kan verschillende technieken toepassen om de **leesbaarheid** en de **bruikbaarheid** van persoonsgegevens door derden te verminderen. Zo is het moeilijker om de betrokken personen en/of bepaalde persoonsgegevens te identificeren.

Enkele van deze gegevensbeschermings-bevorderende **technieken**, ook wel PET's (*Privacy Enhancing Techniques*) genoemd, zijn:

- Differentiële privacy (*Differential privacy*) of verstoring (*Perturbation*);
- Pseudonomisering;
- Onleesbaar maken van persoonsgegevens;
- Versleuteling en homomorfe versleuteling.

Zorg voor een **afdwingbaar rol- en toegangsbeleid** in de organisatie, waarbij personen en applicaties enkel toegang hebben tot de rauwe persoonsgegevens als ze dat echt nodig hebben en na bijkomende identificatie of machtiging.

Bekijk het volledige rapport **'Artificiële Intelligentie en gegevensbescherming: een verkennende gids'** (*minimale gegevensverwerking: pg. 37-49*).