

Samen Slimmer 'de toekomst van data en jouw rol hierin'

Spreker: Layla Hassan

AGENDA

- ❑ Impact van digitalisering op datahoeveelheden binnen de overheid
- ❑ Opkomst van AI en de creatie van nieuwe databronnen
- ❑ Discussies rondom synthetische data en archivering
- ❑ Ethische vraagstukken: bias, transparantie en privacy
- ❑ Zijn er hulpmiddelen?
- ❑ Grote verschillen tussen de Data Act en de Archiefwet
- ❑ De Data Act (dataverordening)
- ❑ Casus: De uitdagingen rondom data-uitwissling
- ❑ Casusvragen: Data-uitwisseling en privacy (Data Act)
- ❑ Nieuwe verantwoordelijkheden door Data Act en AI Act
- ❑ Voorspelde toekomstige rollen van de informatie- en archiefbeheerder

IMPACT VAN DIGITALISERING OP DATAHOEVEELHEDEN BINNEN DE OVERHEID

- De digitalisering binnen de overheid heeft geleid tot een exponentiële groei van datahoeveelheden. Deze toename stelt overheidsorganisaties voor uitdagingen op het gebied van opslag, beheer en beveiliging van gegevens. Een oplossing die wordt overwogen, is de overstap naar clouddiensten om dataopslag te optimaliseren.
- De opkomst van kunstmatige intelligentie (AI) heeft geleid tot de creatie van nieuwe databronnen, wat aanzienlijke ethische overwegingen met zich meebrengt op het gebied van bias, transparantie en privacy. Daarnaast heeft de Europese Unie gereageerd met de introductie van de AI Act om hoogrisico AI-systemen te reguleren.

IMPACT VAN DIGITALISERING OP DATAHOEEVEELHEDEN BINNEN DE OVERHEID

De Algemene Rekenkamer publiceerde in januari 2025 het rapport "Het Rijk in de cloud", waarin wordt geconcludeerd dat de rijksoverheid zonder voldoende risicoafwegingen is overgestapt op het gebruik van clouddiensten. Van de 126 belangrijkste public cloud-diensten is bij 84 diensten (67%) geen risicoafweging gemaakt. Dit gebrek aan strategische afwegingen brengt risico's met zich mee voor de continuïteit van de dienstverlening en de bescherming van gegevens.

Conclusie: de digitalisering binnen de overheid leidt tot grotere datahoeveelheden, wat vraagt om effectieve beheerstrategieën. Hoewel clouddiensten aantrekkelijke oplossingen bieden voor dataopslag, moeten overheidsorganisaties zorgvuldig de balans bewaken tussen efficiëntie en controle om de privacy en rechten van burgers te waarborgen.

Opkomst van AI en de creatie van nieuwe databronnen

AI-technologieën, met name machine learning en deep learning, zijn sterk afhankelijk van grote hoeveelheden data voor training en verbetering. Deze data kunnen afkomstig zijn uit diverse bronnen, zoals sociale media, sensoren en transactiesystemen. Een recente ontwikkeling is het gebruik van synthetische data, waarbij AI wordt ingezet om datasets te genereren die de echte wereld nabootsen. Dit biedt voordelen op het gebied van privacy en het verminderen van bias in trainingsdata. Onderzoekers hebben frameworks ontwikkeld om eerlijkeheidsbeperkingen op te nemen in het zelflerende proces, wat resulteert in representatieve en eerlijke synthetische data.

Voorbeelden

Fairway: A Way to Build Fair ML Software

Dit framework combineert pre-processing en in-processing benaderingen om ethische bias in trainingsdata en getrainde modellen te verminderen. Het richt zich op het identificeren en mitigeren van bias zonder de voorspellende prestaties van het model significant te schaden.

ACROCPoLis: A Descriptive Framework for Making Sense of Fairness

Dit framework biedt een gedeelde vocabulaire waarin de factoren die relevant zijn voor eerlijkeheidsbeoordelingen expliciet worden gemaakt, evenals hun onderlinge relaties.

Discussies rondom synthetische data en archivering

1. Authenticiteit & herleidbaarheid

1. Aangezien synthetische data niet uit de werkelijkheid afkomstig zijn, roept dit vragen op over de betrouwbaarheid en herleidbaarheid van beslissingen die op deze data zijn gebaseerd.

2. Transparantie & verantwoordelijkheid

1. Moet er een documentatieplicht zijn over hoe en waarom synthetische data zijn gegenereerd?
2. Als beleid is gebaseerd op synthetische data, wie is dan verantwoordelijk als er fouten in zitten?

3. Opslagduur & bewaarplicht

1. Moeten synthetische datasets net zo lang bewaard blijven als echte datasets? Of kan men ze verwijderen na gebruik?
2. Hoe onderscheid je relevante synthetische data van irrelevante, aangezien ze vaak in grote hoeveelheden worden gegenereerd?

Ethische vraagstukken: bias, transparantie en privacy

- **Bias:** AI-systemen kunnen onbedoeld bestaande vooroordelen in data overnemen, wat leidt tot discriminerende uitkomsten. Dit kan bijvoorbeeld resulteren in benadeling van bepaalde demografische groepen bij sollicitatieprocedures of kredietbeoordelingen. Het is essentieel om diverse en representatieve datasets te gebruiken en AI-systemen regelmatig te auditen om bias te identificeren en te verminderen.
- **Transparantie:** Veel AI-modellen, vooral deep learning-systemen, functioneren als 'black boxes', waardoor het moeilijk is om hun beslissingen te begrijpen. Gebrek aan transparantie kan het vertrouwen van gebruikers ondermijnen en maakt het uitdagend om biases te detecteren. Het implementeren van uitlegbare AI-methoden is cruciaal om de werking van deze systemen inzichtelijk te maken.
- **Privacy:** AI-systemen verwerken vaak gevoelige informatie, wat risico's met zich meebrengt voor de privacy van individuen. Het is belangrijk om data te anonimiseren en technieken zoals differentiële privacy toe te passen om de privacy te waarborgen. Daarnaast moeten organisaties transparant zijn over welke data ze verzamelen en hoe deze worden gebruikt.

Zijn er hulpmiddelen?

Onderzoek naar AI-gebruik binnen de rijksoverheid

De Algemene Rekenkamer heeft in oktober 2024 een rapport gepubliceerd over het gebruik van AI binnen de rijksoverheid. Hieruit bleek dat er 433 AI-systemen in gebruik zijn, waarvan 30 als hoog risico worden aangemerkt. Het rapport benadrukt de noodzaak van risicomanagement en controlemechanismen rondom AI, om zo transparantie en verantwoord gebruik te waarborgen.

Handreiking voor omgaan met desinformatie

Het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties heeft een handreiking opgesteld voor medeoverheden over het omgaan met desinformatie. Dit document helpt lokale overheden bij het herkennen van AI-gegenereerde desinformatie en biedt richtlijnen voor communicatieadviseurs en lokale politie om hier adequaat op te reageren.

Zijn er hulpmiddelen?

Ontwikkeling van ethische kaders voor digitalisering

Op provinciaal niveau zijn gidsen ontwikkeld die dienen als roadmap voor ethische digitalisering. Deze gidsen helpen overheidsorganisaties bij het implementeren van AI op een manier die rekening houdt met ethische overwegingen, privacy en transparantie.

Transparantie in AI-gebruik door gemeenten

Gemeenten werken aan het vergroten van transparantie door het registreren van algoritmen in nationale registers. Bijna 75% van de algoritmen in het nationale algoritmeregister zijn afkomstig van gemeentelijke organisaties. Dit initiatief bevordert openheid over het gebruik van AI en stelt burgers in staat om inzicht te krijgen in de toepassing ervan.

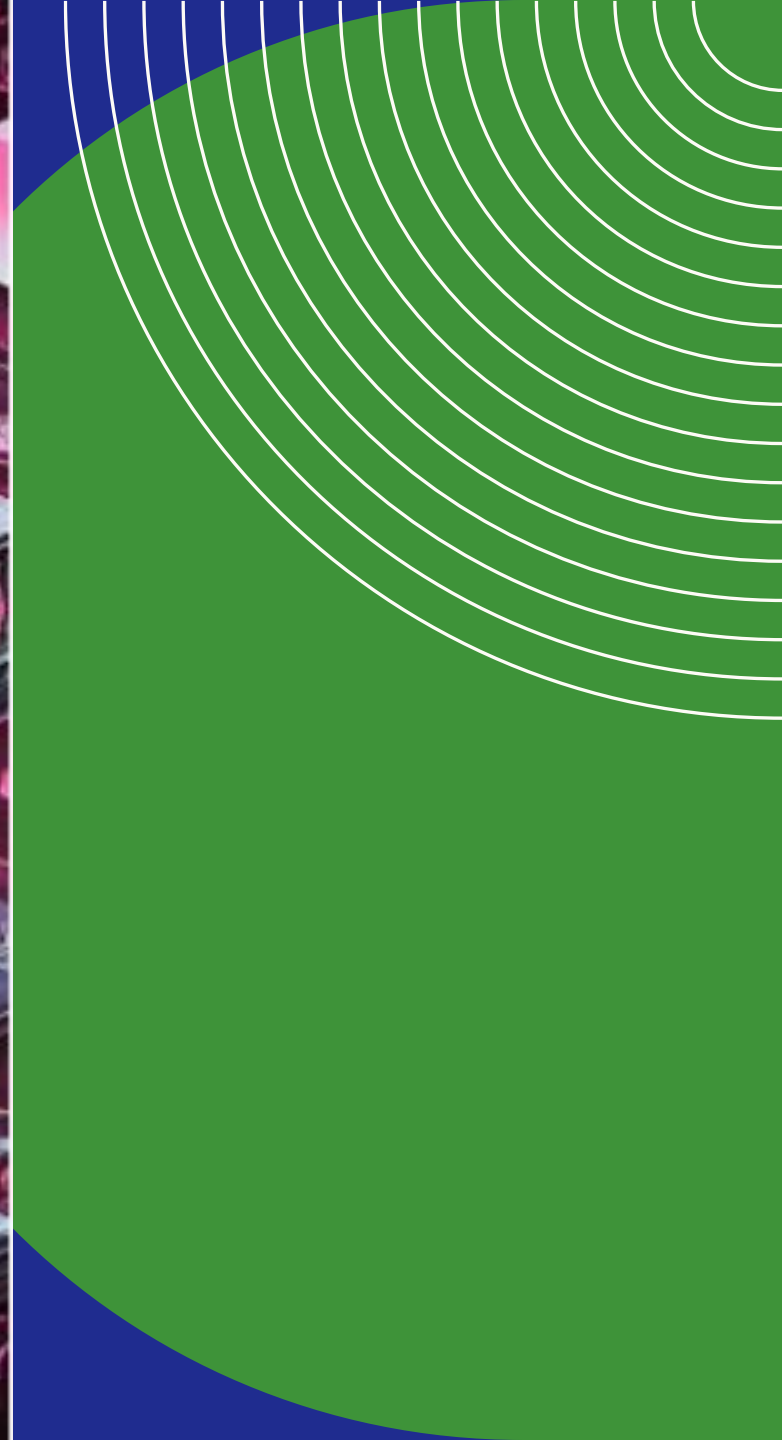
GROTE VERSCHILLEN TUSSEN DE DATA ACT EN DE ARCHIEFWET

Kenmerk	Data Act (EU)	Archiefwet (NL)
Doel	Stimuleren van data-uitwisseling en hergebruik binnen de EU	Beheer en bewaring van overheidsinformatie
Toepassing	Commerciële en publieke data in de digitale economie	Alleen overheidsarchieven en informatiebeheer
Privacy	Richt zich op data-eigendom en toegang tot niet-persoonlijke data	Houdt rekening met de bescherming van persoonsgegevens en de AVG
AI en Automatisering	Gericht op datadeling voor AI-ontwikkeling	Gericht op behoud en toegankelijkheid van informatie
Ethische Vraagstukken	Wie mag data gebruiken en hoe?	Hoe bewaren en vernietigen we data verantwoord?
Regulering	Europese wetgeving, van toepassing op bedrijven en overheden	Nederlandse wetgeving, specifiek voor overheidsinstellingen

De Data Act (dataverordening)

De Europese Data Act richt zich op het bevorderen van data-uitwisseling en het toegankelijk maken van data voor innovatie. Ethische overwegingen hierbij zijn:

- **Data-eigendom en -toegang:** Het delen van data tussen organisaties roept vragen op over eigendomsrechten en de bescherming van gevoelige informatie.
- **Gelijkwaardigheid:** Er is een risico dat grotere organisaties meer profiteren van data-uitwisseling dan kleinere entiteiten, wat ethische bezwaren oproept.



Casus: De uitdagingen rondom data-uitwissling

De gemeente **X**, een innovatieve middelgrote Nederlandse gemeente, heeft de ambitie om volledig digitaal te werken. Dit betekent dat vrijwel alle documentatie, archieven en burgerservices worden gedigitaliseerd. De gemeente gebruikt **kunstmatige intelligentie (AI) voor besluitvorming** en heeft een groeiende hoeveelheid gegevens die moet worden beheerd, gedeeld en beveiligd.

Uitdagingen waarmee gemeente X wordt geconfronteerd:

1. **Exponentiële groei van digitale informatie** – Hoe moet de gemeente alle data beheren en classificeren volgens de **Archiefwet**?
2. **Gebruik van AI bij dienstverlening** – Gemeente X gebruikt AI om **vergunningaanvragen te verwerken en sollicitaties te beoordelen**, maar dit roept vragen op over bias en transparantie (**AI Act**).
3. **Data-uitwisseling tussen overheidsinstanties en bedrijven** – Door de **Data Act** wordt gemeente X verplicht om data toegankelijk te maken voor innovatie, maar hoe bewaakt ze privacy en eigendomsrechten?

Casusvragen: Data-uitwisseling en privacy (Data Act)

Gemeente X krijgt een verzoek van een bedrijf dat slimme mobiliteitsoplossingen ontwikkelt. Het bedrijf wil **geanonimiseerde verkeersgegevens ontvangen** om AI-modellen te trainen voor verkeersoptimalisatie. De **Data Act verplicht overheidsorganisaties om data beschikbaar te stellen**, maar de gemeente maakt zich zorgen over privacy en mogelijk commercieel misbruik van de gegevens.

Vragen:

- Hoe bepaalt gemeente X welke data gedeeld mag worden zonder de privacy van burgers in gevaar te brengen?
- Hoe kan gemeente X voldoen aan de **Data Act** zonder de ethische waarden van privacy en publieke belangen te schenden?
- Welke rol kunnen archief- en informatiebeheerders spelen in het toezicht houden op **data-eigendom en transparantie**?

Nieuwe verantwoordelijkheden door Data Act en AI Act

De rol van de informatie- en archiefbeheerder zal in de toekomst sterk veranderen door de impact van digitalisering, kunstmatige intelligentie (AI), en nieuwe wet- en regelgeving zoals de **Archiefwet, de Data Act en de AI Act**. De focus zal verschuiven van **klassiek archiveren en documentbeheer** naar **strategisch data- en informatiebeheer** waarbij ethiek, juridische naleving en technologie een grote rol spelen.

- Door de **Data Act** en de **AI Act** zal de nadruk komen te liggen op **datagovernance, ethische besluitvorming en naleving van regelgeving**.
- Het beheer van **grote hoeveelheden dynamische en gestructureerde data** zal belangrijker worden dan statische documenten.

Voorspelde toekomstige rollen van de informatie- en archiefbeheerder

Door deze ontwikkelingen kunnen we verwachten dat de traditionele functie zich opsplitst in nieuwe specialistische rollen, zoals:

- 1. AI-gestuurde informatiebeheerder** – Werkt met AI-tools om informatiebeheer efficiënter te maken en bias in AI-systemen te signaleren.
- 2. Data governance specialist** – Richt zich op ethische vraagstukken en zorgt dat data wordt beheerd volgens de wetgeving.
- 3. Transparantie- en compliance manager** – Zorgt ervoor dat AI-beslissingen uitlegbaar en controleerbaar zijn.
- 4. Informatie-ethicus** – Houdt zich bezig met morele en ethische vraagstukken bij gegevensverwerking en archivering.

BEDANKT!

