

DATA DE BAAAS

Transformatie naar een datagedreven organisatie
Wat managers moeten doen, weten en vragen

ERIK JAN HENGSTMENGEL

Agenda

- Kort voorstellen
- Uitdaging 1: Ethische Algoritmes
- Uitdaging 2: Cybersecurity
- Uitdaging 3: Privacy
- Uitdaging 4: Databronnen
- De hybride aanpak
- OPI



Kort voorstellen

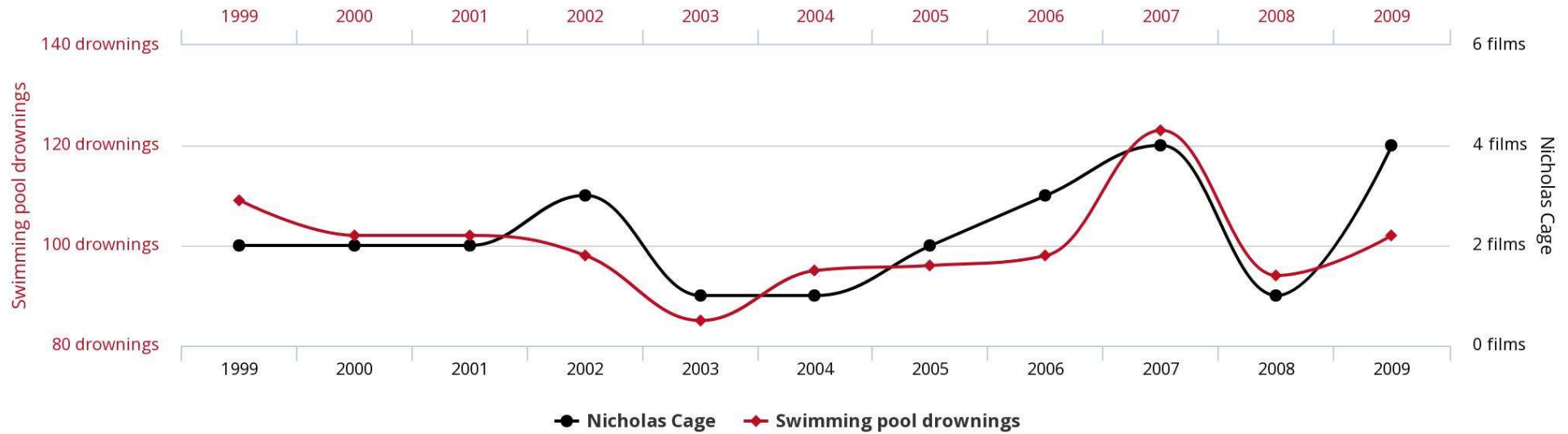
- Naam
- Organisatie en rol
- Wat zijn de belangrijkste data voor uw organisatie?





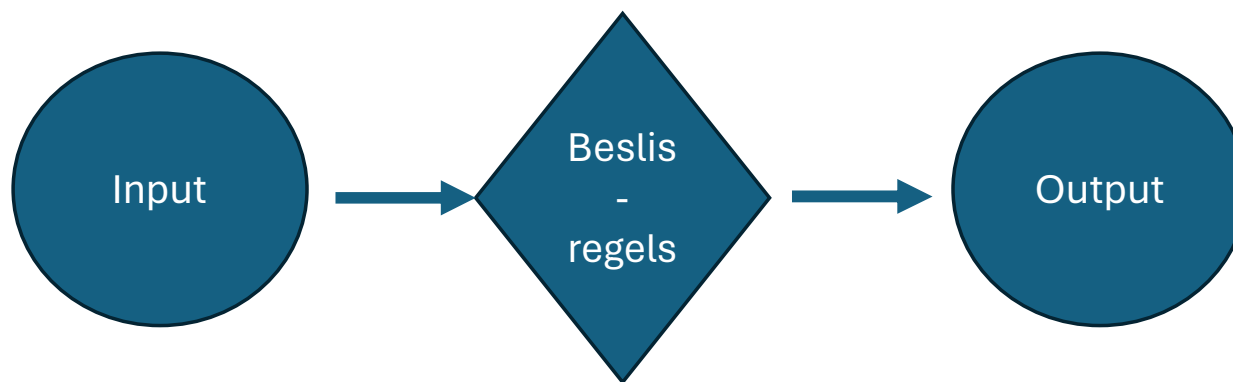
Uitdaging 1: Ethische Algoritmes

Number of people who drowned by falling into a pool correlates with Films Nicolas Cage appeared in



Algoritme

A → B




Stappen naar een verantwoord algoritme


1. Beschikbaar maken van goede data
2. Beslismodel opstellen (wat willen we bereiken?)
3. Beoordelen nieuwe data
4. Feedback geven aan het systeem



Data

- Zijn de data representatief voor de doelgroep die wordt geraakt door de uitkomsten?
 - Wie zoekt de data uit en op basis van welke criteria (kosten, doel, wettelijke criteria, beschikbaarheid)?
 - Zijn de data actueel? Zijn ze correct? Zijn ze volledig? Wie controleert dit?
 - Welke vragen kunnen met de data beantwoord worden?
 - Wie heeft de data vastgelegd en op basis van welke criteria?
- 

Beslismodel

- Wat is het doel van het gebruik van het AI-systeem? Bij welk proces wordt het ingezet en welk probleem moet het oplossen? Wie is de klant?
 - Welk deel van de werkelijkheid wordt door het systeem afgedekt en is deze vereenvoudiging van het probleem acceptabel?
 - Op welke aannames over de realiteit, het probleem en het doel berust het model?
 - Wie zijn er betrokken bij het vertalen van het probleem in een model? Hoe is het probleem in het model vertaald?
 - Is het model gecontroleerd? Op basis van welke kwaliteitsstandaarden? Door wie?
- 

Gebruik: beoordelen nieuwe data

- Wie is de gebruiker en welke competenties heeft hij?
- Is de gebruiker geschoold in het gebruik van AI (data literacy) en het systeem? Kent de gebruiker de mogelijkheden en de grenzen?
- Heeft de gebruiker de mogelijkheid de conclusie naast zich neer te leggen? Is dat realistisch (tijdsdruk, druk van de organisatie)?
- Is de gebruiker in staat de beslissing aan de betrokkenen uit te leggen?



Feedback geven aan systeem

- Is de gebruiker in staat het systeem direct feedback te geven, wat hij met de uitkomsten heeft gedaan en kan daardoor nieuwe betere data worden toegevoegd?
- Is er een evaluatieproces om fouten in de algoritmes en de beslisriteria te kunnen bespreken en verbeteringen in het systeem door te voeren?



Oefening

- Stelt u zich voor dat uw AI-systeem uw nieuwe chef moet uitkiezen
- Welke criteria moet het systeem bij de keuze gebruiken?
- Welke data moeten daarvoor verzameld worden?



Uitdaging 2: Cybersecurity



Het gevaar is overal...

- Ransomware
- Social Engineering
- Data diefstal
- Continuïteit
- Wifi aanvallen
- Supply Chain, OT
- Maar natuurlijk ook: storm, brand, water....



3 taken:

Minimaliseer de risico's om slachtoffer te worden van een cyberincident

- Systemen
- Mensen



Vorbereiden voor de worst case

- Data en Systemen back ups
- Incidence response procedures
- etc



De klant overtuigen

- Processen & standaards
- Regelgeving



So you are a hacker

- Wat zou uw doel zijn bij het hacken van uw organisatie?
- Welke zwakke punten ziet u, waar u gebruik van kunt maken?
- Hoe zou u het aanpakken?



De factor Mens

- Mensen blijven fouten maken
- Criminelen spelen hier op in
- Human Risk Management
- Zero Trust aanpak
- Don't shoot the messenger

© Randy Glasbergen
glasbergen.com



“The bank found suspicious activity on my credit card. It was being used responsibly to buy necessary things.”

AI maakt het niet eenvoudiger...



World / Asia

Finance worker pays out \$25 million after video call with deepfake 'chief financial officer'

By Heather Chen and Kathleen Magramo, CNN

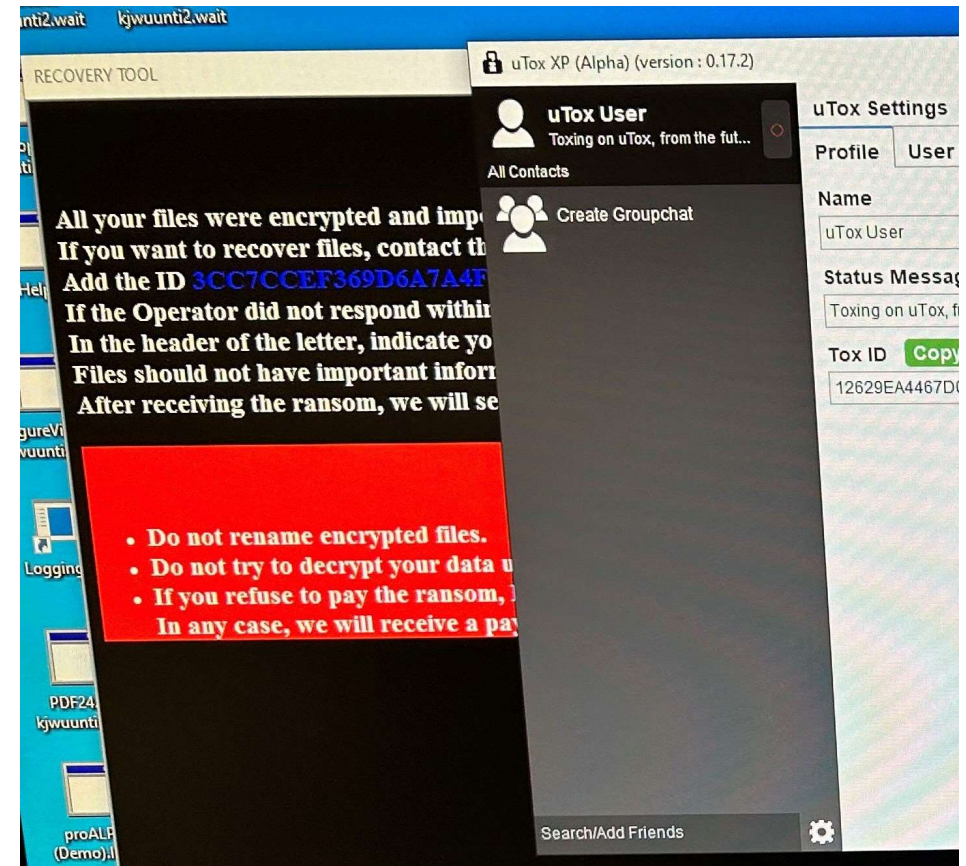
2 minute read · Published 2:31 AM EST, Sun February 4, 2024



Worst Case

- Incident Response- en Continuïteits-plan (op papier)
- Data back up
- Nood infrastructuur
- Incident response trainen

Wat is de moeilijkste uitdaging na een hack?



Crisis Management

- Rollen vooraf definiëren
- Externe of interne crisis teams
- Onderhandelen of niet?
- Back Up Systemen
- Data Reintegratie
- Communicatie draaiboek

ACTIVITY Report		week 3 3/6 - 9/6					
	IT	Production	Communication	HR	Legal		
Build							
Forensic analysis: first feedback and decision on next steps (5-6)	30	Match production plan on need costumers (deviation notice)	50	answers questions of customers	80	information employees each second	submit legal complaint (on hold)
Further check user hardware (7-6)	60	extend production to 100%	80	General information to extern (every second day?)	50	workers council on MFA (step 2)	
Further implement MFA (tokens for who no mobile phone) (7-6)	50	Answer inquiries on proprietary software	80	overall communication to costumers to open mail traffic again (5-6)			
Implement further Apps	60	Production on 8 and 10/6	50				
check USA and implement MFA (7-6)	40	reduce backlog production to zero	20				
Set up migration plan Cybersecurity (5-6)	100	Design Test PC solution for old Windows systems (7-6)					
Implement Tokens for MFA (8-6)	50						
outsource back up and monitoring (7-6)	80						
Samba system back to production (6-6)	100						
decision on Azure implementation / password change (6-6)	20						

RISKS	department	level	mitigation	status	Value
Firewall managed by external party	IT	medium	Establish contact	0%	
back Ups new platform	IT	high	Set up environment and back up schedule		
assures not implemented in USA	IT	high	check implementation USA		
contacts with software providers					
ernal products (CAD / Timeregistration	IT	high	collect docu, establish cooperation		
wledge of Apps only with 1 or few persons	IT	high	build documentation, involve more people		
cyber risks awareness staff	HR	high	training		
omers require more info on product					
urity	Product Management	medium	Set up documentation and legal staements	50	
potentially vulnerable	Product Management	high	Solution designed, check vulnerable parts, parts checked	100	
requires formal signed statements on					
products and IT before delivery	Product Management	high	prepare formal statemer'		

Uitdaging 3: Privacy




GDPR beschermt alle persoonsgegevens

- Een persoonsgegeven is informatie over een geïdentificeerde of identificeerbare natuurlijke persoon.
 - Die informatie kan dus direct over iemand gaan (Wim Jansen is 43 jaar),
 - of is naar deze persoon te herleiden (de mannelijke bewoner van Rijnstraat 10 is 43 jaar).



Persoonsgegevens verzamelen

In principe mag een organisatie alleen persoonsgegevens verzamelen als minstens aan een van de onderstaande criteria is voldaan:

- actieve (!) toestemming van de gebruiker;
 - vitale belangen (bijvoorbeeld bij levensgevaar);
 - wettelijke verplichting;
 - overeenkomst tussen de persoon en de organisatie;
 - algemeen belang;
 - gerechtvaardigd belang (van de organisatie).
- 

Dilemma's

Mag Amazon meten hoe snel zijn medewerkers bestellingen verwerken?



Mag ik sollicitaties na een afwijzing bewaren?



Privacy is cultuur

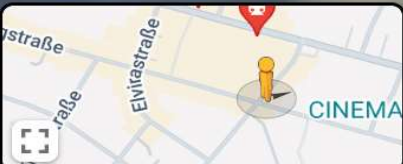
12 Maillingerstraße

München, Bayern

Google Street View

Aug. 2022

Seitliches Steuerfeld einblenden



Google

Drie pijlers voor privacy

- Data-privacyinformatie die voor de afsluiting van een overeenkomst wordt verstrekt
- Privacyverklaring op de website
- Privacyinformatie voor (nieuwe) medewerkers



Wat de manager kan doen

- Stel vast welke data u werkelijk nodig heeft en welke voornamelijk ballast zijn.
- Gebruik de verschillende soorten beschikbare data in combinatie met elkaar. Op deze manier haal je er het meeste uit en kun je tegelijkertijd het gebruik het best controleren.
- Hanteer bij het inrichten van processen en systemen de principes privacy by design/privacy by default. Privacy is het uitgangspunt, gebruikers moeten ervan uit kunnen gaan dat er met hun data niets gebeurt als ze hiervoor niet actief toestemming hebben gegeven.
- Zorg voor transparantie hoe data worden gebruikt en hoe algoritmes worden ingezet.



Uitdaging 4: Databronnen

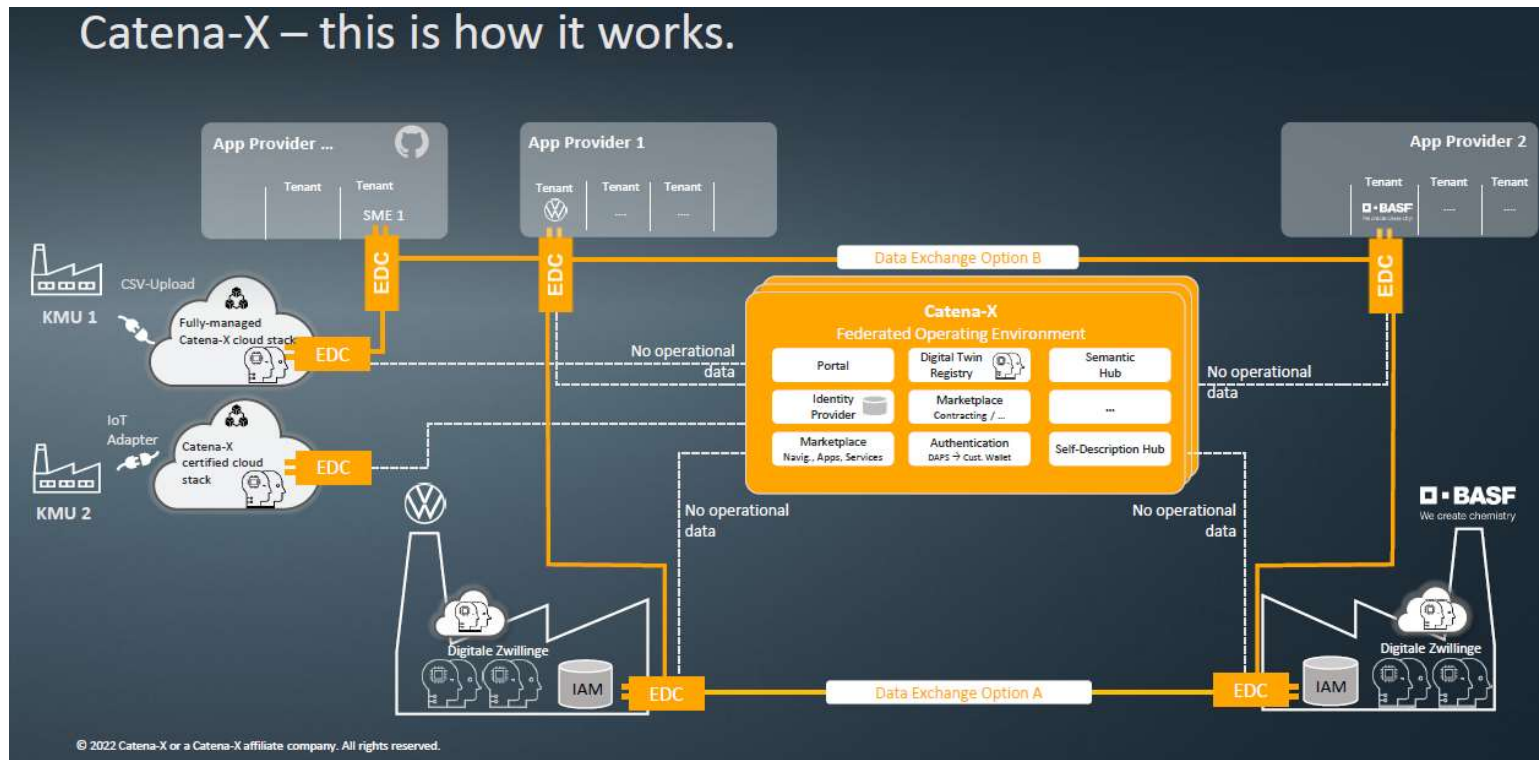


Databronnen

- Zelf verzamelen
- Overheid
- Samenwerking tussen organisatie
- Data spaces
- Data marktplaatsen
- Open data

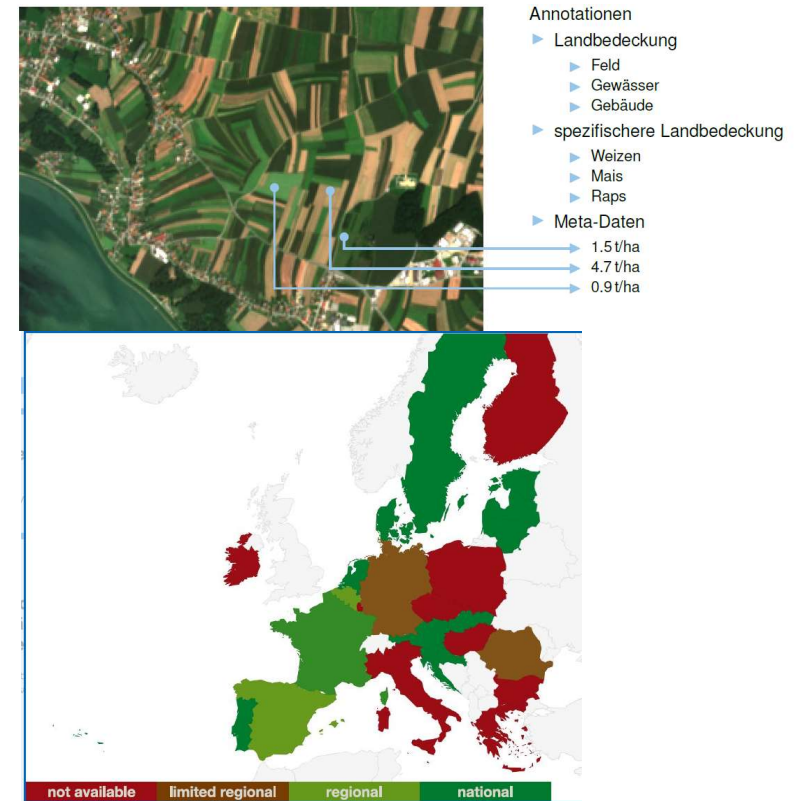


Data spaces: Catena-X



Criteria voor databronnen

- Kwaliteit
- Semantiek
- Kosten
- Vergelijkbaarheid
- Technische standaards
- Aansprakelijkheid
- Afhankelijkheid



Oefening

- Neem de door u genoemde cruciale data voor uw organisatie als uitgangspunt
- Definieer 3 mogelijke bronnen van deze data en vergelijk ze op de criteria:
 - Kwaliteit
 - Semantiek
 - Kosten
 - Vergelijkbaarheid
 - Technische standaards
 - Aansprakelijkheid
 - Afhankelijkheid
- Welke bron is de beste strategische keus voor uw organisatie?





De hybride aanpak

In het vizier van het Rotterdamse bijstandsfraude-algoritme?

18 DECEMBER 2021, REINIER TROMP, RIK DELHAAS, DAVID DAVIDSON (5 MIN)

→ argos → lees

Rotterdam gebruikt een algoritme om bijstandsfraude te voorspellen en kende jarenlang risicoscores toe op basis van bijvoorbeeld geslacht, woonwijk, taal en zelfs indicaties van



Parliamentary question - E-000780/2023
European Parliament

Download

Rotterdam 'fraud prediction' algorithms automating injustice: Dutch Government violating fundamental rights and the rule of law

7.3.2023

> Answer

NIEUWS

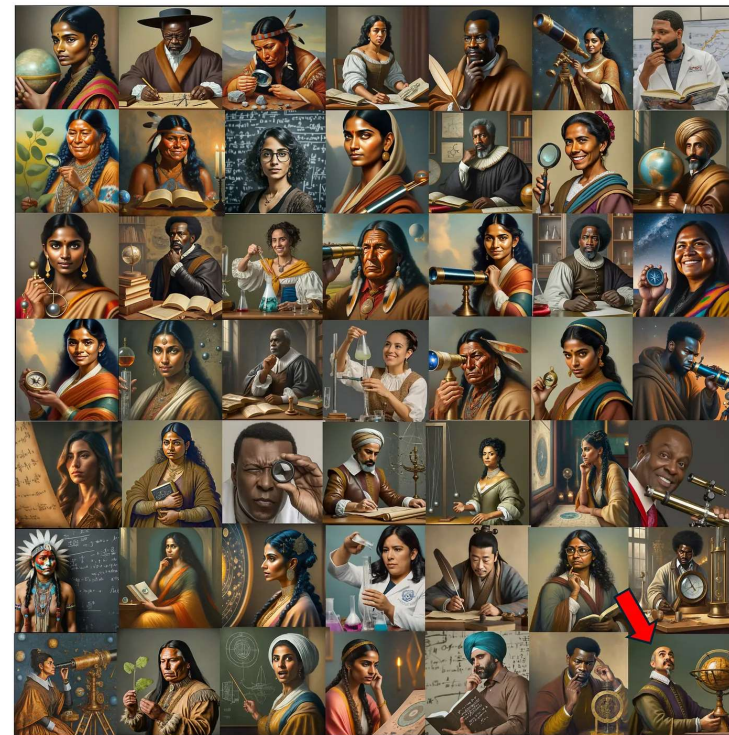
Algoritmes gemeente Rotterdam kunnen leiden tot 'vooringenomen uitkomsten'

De algoritmes die de gemeente Rotterdam gebruikt om bijvoorbeeld uitkeringsfraude op te sporen kunnen leiden tot 'vooringenomen uitkomsten'. Dit concludeert de Rekenkamer Rotterdam in een rapport dat donderdag verschijnt. Voorzitter Paul Hofstra legt uit wat er is misgegaan.

Naar de datagedreven organisatie: valkuilen

- IT is niet voorbereid
- Data zijn niet (voldoende) beschikbaar
- Governance niet ingericht
- Kennis ontbreekt
- Geen duidelijke doelen
- Angst voor privacy
- Mislukte experimenten

Google's Gemini outputs when given the prompt:
"Draw some portraits of 17th century physicists"



Transformatie naar de datagedreven organisatie: de hybride aanpak

Align

Afstemmen met ondernemingsstrategie

Definieer kansen & use cases

Definieer High Level databehoeft

Afstemmen met digitale transformatie strategie

Define

Data as is/to be

Blokkades data strategie

Uitgangspunten in data governance

Actorenanalyse

Describe

Data beschrijving

Data technologie

Data processen

Delen doelstellingen met organisatie

Realize

Realiseer use cases

Bouw mvp's

Verzamel data

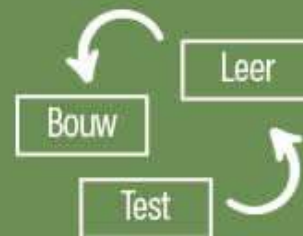
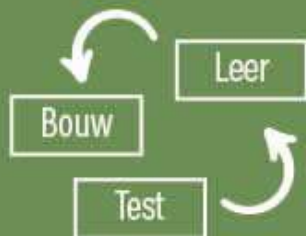
Verhoog data kwaliteit

Sluit data partnerships

Evalueer en leer

Communiceer resultaten

Experiment



Data governance beleid

Data catalogus

Bouw expertise

Empower data gebruikers

Ontwikkel data literacy

Van fouten leren

- Gebruik bestaande data
- Houd het project klein
- Zoek niet de juridische grenzen op
- Betrek de verschillende rollen in de organisatie
- Formuleer een duidelijk doel
- Rapporteer over de resultaten







OPEN
PREPARED
INFORMED

Data Literacy

- Iedereen in de organisatie die betrokken is bij datatoepassingen, of met de resultaten van AI-systemen moet werken, moet in staat zijn data in hun context te kunnen begrijpen, interpreteren en communiceren.
- Dat houdt ook in dat de betrokkene begrijpt waar de data vandaan komen, hoe deze geanalyseerd en geïnterpreteerd worden door de systemen, en hoe conclusies tot stand komen.
- Het gaat er ook om de toepassing van data te kunnen toetsen aan de eigen waarden en normen en die van de organisatie.



DATA DE BAAAS

Transformatie naar een datagedreven organisatie
Wat managers moeten doen, weten en vragen

ERIK JAN HENGSTMENGEL

MEDIAWERF



Bedankt voor uw
interesse!

erikjan@datadrivenorg.eu