



E-mailarchivering & AI: tussen wetgeving en werkelijkheid

CIO-dag 24 juni 2026



Rijksprogramma
Duurzaam
Digitale
Informatiehuishouding

Het kader

De pilot

Het spel

Inzichten



Archiefwet

AVG

WOO



E-mailarchivering

› Van **idee** naar **impact**



Casus

- › De pilot richt zich op het toepassen en testen van automatische e-mailclassificatie binnen een afgebakende set e-mailboxen.
- › RDDI i.s.m. het Ministerie van Sociale Zaken & Werkgelegenheid.
- › E-mailboxen van vertrokken bewindspersonen (BWP) en oud-bestuursraadleden (BR-leden).
- › E- mailboxen uit de periode 2014 – 2025.
- › E-mails categoriseren in 5 categorieën (functioneel, niet – functioneel, privé, personeelsvertrouwelijk en partijpolitiek).



Rijksprogramma
Duurzaam
Digitale
Informatiehuishouding



De data

Getal	Duiding
20	e-mailboxen
834.531	e-mails in totaal
752.074	unieke e-mails in totaal
13.709	e-mails handmatig geclassificeerd
670	e-mails gecorrigeerd (4,9% van het geheel)



Praktijkervaringen en lessen uit de AI-pilot

- › Van alle mails is 54% niet-functioneel.
- › Functionele en niet-functionele mails zijn vrij nauwkeurig te onderscheiden (85%).
- › Twee modellen (logistic regression en PEFT-taalmodel) die evenredig presteren.





Praktijkervaringen en lessen uit de AI-pilot

- › Winst is te behalen door uitbreiding en verbetering van de trainingsdata voor de kleinere categorieën voor hogere accuraat.
- › Bredere organisatorische vraag blootgelegd die nog niet is beantwoord: Welke drempelwaarde voor privacygevoelige e-mails in een te archiveren dataset is acceptabel, en hoe weegt de organisatie de verplichtingen uit de Archiefwet af tegen de vereisten van de AVG?





Het Spel



Rijksprogramma
Duurzaam
Digitale
Informatiehuishouding



80%

De norm



Rijksprogramma
Duurzaam
Digitale
Informatiehuishouding



Het Spel

- › Ieder team start bij hun eigen tafel met de stapel van 50 e-mails voor zich.
- › Het doel is om deze mails zo snel mogelijk te categoriseren en in het juiste bakje te plaatsen.
- › E-mails in de categorie: *niet-functioneel, privé, partijpolitiek of personeelsvertrouwelijk* worden in de daarvoor bestemde team-doos op tafel geplaatst.
- › *Functionele e-mails* bewaar je op je eigen tafel.
- › Het team dat alle mails heeft verwerkt steekt de handen in de lucht en roept STOP!
- › Het team met de snelste tijd én het laagste foutpercentage wint het spel.
 - TIP: Twijfel je? Overleg het even met je team.



Een **functionele e-mail** is een e-mail die onderdeel uitmaakt van het uitvoeren van een publieke taak of van het bestuurlijk handelen van een overheidsorganisatie.

Een **niet-functionele e-mail** maakt geen onderdeel uit van het uitvoeren van een publieke taak of van het bestuurlijk handelen van een overheidsorganisatie. Het is dus alles wat niet direct bijdraagt aan die taak, maar wél via het werkaccount kan lopen.

Een **privé-e-mail** is een bericht dat geen verband houdt met de uitvoering van de publieke taak of het bestuurlijk handelen van een overheidsorganisatie.

Een **personeelsvertrouwelijke e-mail** is een bericht dat (bijzondere) persoonsgegevens bevat over de arbeidsrelatie tussen werkgever en medewerker, en dat primair thuishoort in een afgeschermd personeelsdossier.

Partijpolitieke e-mails zijn communicatie met partijgenoten over onderwerpen die de partij aangaan (bijv. interne partijaangelegenheden en inhoudelijke partijpolitieke standpunten).

COUNTDOWN

10.00
MIN.

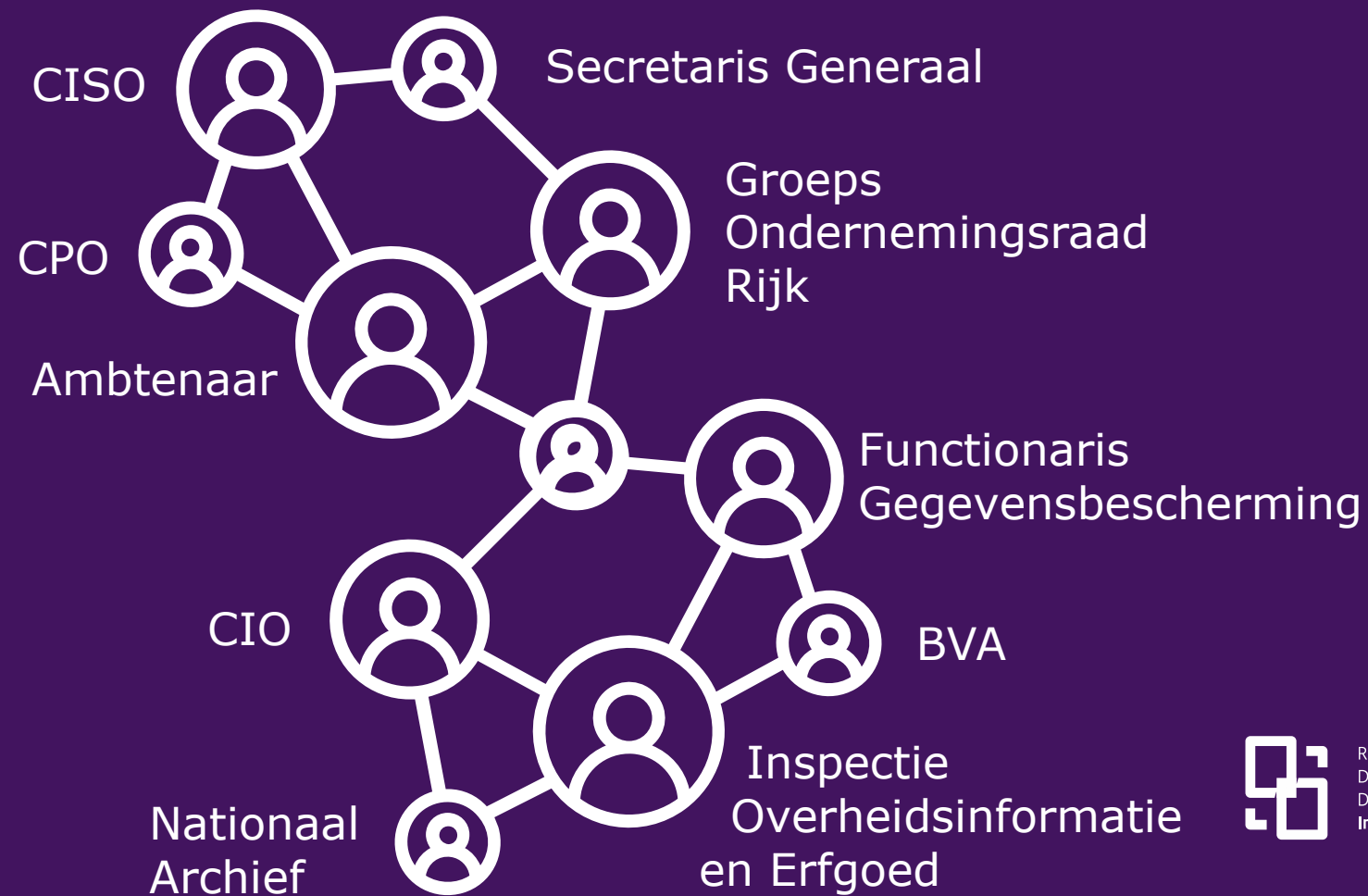
with SIMPLE BEEPS



Inzichten

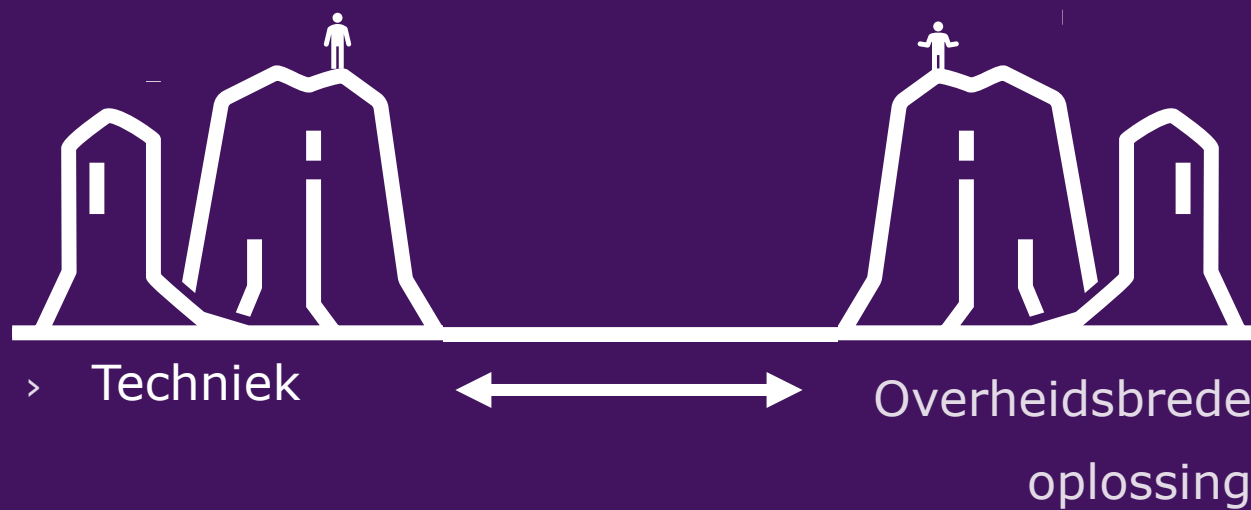


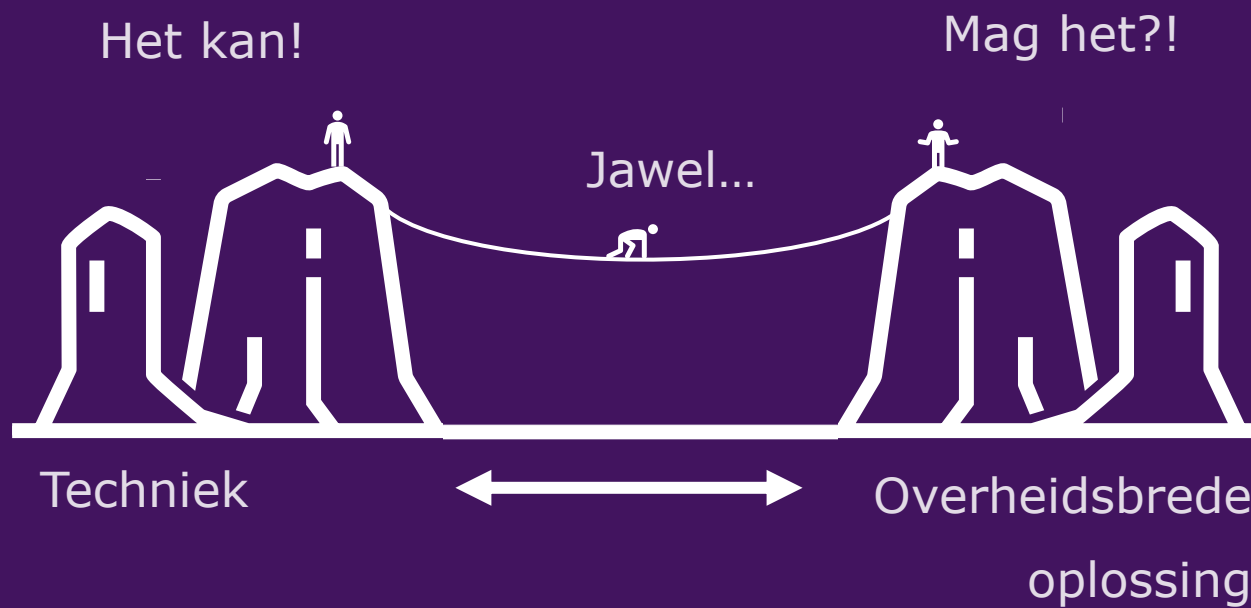
Rijksprogramma
Duurzaam
Digitale
Informatiehuishouding





Geleerde lessen





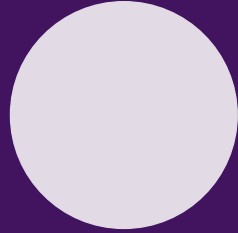
Rijksprogramma
Duurzaam
Digitale
Informatiehuishouding



De vergelijking

De AI heeft in 6 minuut en 25 seconde, 50 mails verwerkt met 82% accuratesse.
Deden jullie het beter dan de AI?

	Team Geel	Team Oranje	Team Rood	Team Paars	Team Blauw	Team Groen
Tijd	5 minuten en 30 seconde	7 minuten en 45 seconde	6 minuten	5 minuten en 30 seconde	8 minuten	3 minuten en 30 seconde
Fout geclassificeerd	16 fout van de 50	14 fout van de 50	12 fout van de 50	17 fout van de 50	13 fout van de 50	8 fout van de 50
Goed geclassificeerd	34 goed van de 50	36 goed van de 50	38 goed van de 50	33 goed van de 50	37 goed van de 50	42 goed van de 50
Accuratie	68%	72%	76%	66%	74%	84%



~ 1-2 miljard e-mails per jaar

~ 10 miljard e-mails aan legacy



Rijksprogramma
Duurzaam
Digitale
Informatiehuishouding



Afsluitend

Meer weten?
Open de QR voor de uitkomsten van de pilot

