

Foto: Fritzens Fritz / <https://betterimagesofai.org>
/ <https://creativecommons.org/licenses/by/4.0/>

Rapport TNO 2025 R12653

Overheidsbrede Monitor Generatieve AI

Inzicht in Generatieve AI inzet bij
Nederlandse Overheidsorganisaties

Auteurs:

Marissa Hoekstra
Tessa Bruijne
Marianne Schoenmakers
Pippa Jones

December 2025

TNOvector
Centre for Societal Innovation and Strategy

Inhoudsopgave

1. Inleiding	3
2. Onderzoeksmethode	5
3. Inzet van GenAI bij Nederlandse Overheidsorganisaties	7
4. Succesfactoren en knelpunten bij de inzet van GenAI bij Nederlandse Overheidsorganisaties	15
5. Praktijkvoorbeelden	21
6. Conclusies en aanbevelingen	30
Acknowledgements	35
Referenties	36

1. Inleiding

Inleiding

Achtergrond

De opkomst van generatieve AI (GenAI) heeft ertoe geleid dat AI technologieën breed beschikbaar zijn voor individuen en organisaties. Ook bij Nederlandse overheidsorganisaties is er interesse in GenAI. In 2024 is de Overheidsbrede visie Generatieve AI gepubliceerd¹, waarin werd aangekondigd dat er jaarlijks een monitor generatieve AI uit wordt gebracht om de inzet van GenAI door overheden te volgen. Aan het begin van 2025 zijn de vernieuwde overheidsbrede standpunt voor de inzet van GenAI² en de handreiking voor verantwoorde inzet van generatieve AI³ gepubliceerd, waarin de spelregels op innoveren en experimenteren met GenAI wordt gepresenteerd. In Juni 2025 is tevens de Nederlandse Digitaliseringsstrategie⁴ gepubliceerd.

Onderzoeksdoel

In het kader van deze ontwikkelingen, voert TNO in samenwerking met het

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) onderzoek uit naar de inzet van GenAI toepassingen bij overheidsorganisaties in Nederland. Het doel van dit onderzoek is driedelig:

- Inzicht krijgen in de **inzet van GenAI door overheidsorganisaties**;
- Inzicht krijgen in de **succesfactoren en knelpunten bij de inzet van GenAI toepassingen** die Nederlandse overheidsorganisaties ervaren;
- Inzicht krijgen in de **overwegingen** die worden gemaakt bij de inzet van GenAI toepassingen door Nederlandse overheidsorganisaties.

De bevindingen van dit onderzoek worden in dit rapport gepresenteerd. De resultaten die worden besproken in dit rapport zijn gebaseerd op publiek beschikbare informatie en inzichten die gedeeld zijn gedurende het onderzoeksproces. **De mate van representativiteit** voor de inzet van GenAI bij Nederlandse overheidsorganisaties kan daarom niet worden vastgesteld.

Desondanks geeft dit onderzoek inzicht in voorbeelden van hoe GenAI bij Nederlandse overheidsorganisaties wordt toegepast en in relevante succesfactoren en knelpunten bij de inzet van GenAI.

Onderzoeksvragen

Het onderzoek naar de inzet van GenAI bij Nederlandse overheidsorganisaties bestaat uit drie delen. Het eerste deel richt zich op het gebruik van GenAI toepassingen. Het tweede deel gaat dieper in op de succesfactoren en knelpunten bij de ontwikkeling en inzet van GenAI. In het derde deel worden vier voorbeelden uit de praktijk toegelicht. Voor elk voorbeeld wordt een beschrijving gegeven van de toepassing, hoe deze tot stand is gekomen en welke overwegingen belangrijk zijn bij de ontwikkeling en inzet. Ook wordt uitgediept welke succesfactoren en knelpunten een rol speelden bij de ontwikkeling en inzet van de GenAI uit het voorbeeld.

Voor elk deel van dit onderzoek zijn onderzoeksvragen geformuleerd:

1. Welke GenAI toepassingen worden door de overheid ingezet en wat zijn daarvan de kenmerken?
2. Wat zijn succesfactoren en knelpunten bij de inzet van GenAI toepassingen die Nederlandse overheidsorganisaties ervaren?
3. Wat zijn overwegingen die worden gemaakt bij de ontwikkeling en inzet van GenAI toepassingen voor Nederlandse overheidsinstellingen?

Definitie van generatieve AI

In dit onderzoek hanteren we de definitie van GenAI en taalmodellen van het Rathenau Instituut. Volgens deze definitie zijn GenAI “AI systemen die geautomatiseerd content kunnen maken, op verzoek van een gebruiker.”⁵

Er zijn verschillende vormen van content, bijvoorbeeld: tekst, stem, foto's, video's, software of audiovisuele content waarin aspecten worden gecombineerd.^{6, 7} GenAI kan nieuwe content genereren door een synthese te maken van de ingevoerde

1 [Overheidsbrede visie Generatieve AI](#)

2 [Overheidsbreed standpunt voor de inzet van generatieve AI](#)

3 [Overheidsbrede handreiking voor verantwoorde inzet van generatieve AI](#)

4 [De Nederlandse Digitaliseringsstrategie](#)

5 Ibid., p. 5

6 Ibid.

7 Garon, 2023

informatie, daarom wordt deze content ook wel ‘synthetische content, output, of media’ genoemd.^{8,9}

Een GenAI-systeem gebruikt vaak, maar is niet beperkt tot, een groot taalmodel, ook wel een Large Language Model (LLM)¹⁰. Dit zijn software modellen die op basis van berekeningen voorspellen welk woord het meest waarschijnlijke volgende woord is in een zin of om ontbrekende woorden in zinnen aan te vullen.¹¹ Er wordt al lange tijd gewerkt aan het automatiseren van het verwerken van menselijke taal (denk aan stem assistenten zoals Siri of Alexa of vertaalhulpmiddelen zoals Google Translate).¹² Recente ontwikkelingen maken het echter mogelijk voor modellen om meer data te analyseren, waardoor de voorspellingen beter kloppen en accurater lijken.^{13, 14, 15}

8 Garon, 2023

9 TNO, 2024, p. 14

10 Bender et al., 2021

11 Rathenau Instituut, 2024

12 Oomen, 2025

13 Rathenau Instituut, 2023

14 Bommasani et al., 2022

15 Bender et al., 2021

2. Onderzoeksmethode

Onderzoeksmethode deel 1: Inzet van GenAI

Het doel van deel 1 van het onderzoek is om inzicht te krijgen in de inzet van GenAI door Nederlandse overheidsorganisaties. Met behulp van desk research en survey onderzoek hebben we voorbeelden van toepassingen van GenAI verzameld en gecategoriseerd.

Deel 1 van het onderzoek is uitgevoerd aan de hand van de volgende stappen:

1. Bepalen van de scope en reikwijdte van het onderzoek.

Het Register van Overheidsorganisaties was leidend¹⁶. Type overheidsorganisaties die binnen de scope van dit onderzoek vallen zijn onder andere: agentschappen; adviescolleges; gemeenten; inspecties; ministeries; politie; provincies; rechtspraak; regionale samenwerkingsorganen; uitvoeringsinstellingen; waterschappen en zelfstandige bestuursorganen (zbo). Zorg en onderwijsorganisaties zijn niet meegenomen in dit onderzoek. Net zoals bij de Quickscan AI in de Publieke Dienstverlening III (2024)¹⁷ is het individueel gebruik van GenAI van

ambtenaren buiten scope van dit onderzoek. Dit onderzoek focust zich op toepassingen die top-down worden aangeboden door overheidsorganisaties

2. Desk research en survey onderzoek

vond plaats tussen 18 februari 2025 en 16 juni 2025. Geraadpleegde bronnen zijn:

- Google search naar publiek beschikbare informatie en search op overheidsorganisaties en semi-overheid databases zoals overheid.nl. Zoektermen: (“GenAI” OR “generatieve AI” OR “generative AI” OR “LLM” OR “GPT” OR “taalmodel” OR “chatbot”) AND (“overheid” OR “overheidsorganisatie” OR “[naam overheidsorganisatie]”)
- Algoritmeregister van de Nederlandse overheid¹⁸. Het algoritmeregister is (nog) niet verplicht, dus een volledig overzicht van alle GenAI toepassingen is niet gegarandeerd¹⁹.
- Uitvraag met een online survey via het netwerk van BZK, VNG en Platform AI en Overheid.

3. **Analyse en categorisering van de toepassingen** aan de hand van de volgende kenmerken, zoals: type organisatie, status toepassing, type proces, type taak, beoogde eindgebruiker, risico inschatting, impact toets, ontwikkelvorm.

4. **Contextualiseren van inzichten en verwerken** in rapportage.

Onderzoeksmethode deel 2: Succesfactoren en knelpunten

Het doel van deel 2 van het onderzoek is om inzicht te krijgen in knelpunten en succesfactoren die overheidsorganisaties ervaren bij de inzet van GenAI. Deze inzichten hebben we opgehaald door middel van **interviews** met sleutelrollen die betrokken zijn bij de inkoop, ontwikkeling of gebruik van GenAI toepassingen bij verschillende overheidsorganisaties.

Tussen 26 mei 2025 en 25 juni 2025 hebben we 17 interviews gehouden met 20 deelnemers (drie interviews met twee deelnemers van één organisatie). Hoewel dit onderdeel van het onderzoek hiermee niet representatief is voor de hele doelgroep, hebben we een breed scala aan

perspectieven en ervaringen verzameld die waardevolle inzichten bieden.

De deelnemers zijn betrokken bij GenAI toepassingen in overheidsorganisaties, waarvan enkelen ook technische kennis hebben van (Gen)AI. Er is een gevarieerde set aan casussen bij verschillende overheidsorganisaties geselecteerd. Zie hiervoor **tabel 1**. Selectie vond plaats op basis van verwijzing via het netwerk of na deelname aan de survey uit deel 1 van dit onderzoek.

De interviews volgden een semigestructureerde opzet, waarvoor vooraf een set vragen over onderwerpen is opgesteld, zoals aanleiding toepassing, ontwikkelvorm, ontwikkelproces, data, technologie & infrastructuur, AI geletterdheid, innovatiebewustzijn, leiderschap, organisatie en processen, maatschappelijk doel en lessen voor andere organisaties. De interviews zijn met toestemming opgenomen en getranscribeerd ten behoeve van de analyse en rapportage.

¹⁶ [Register van overheidsorganisaties](#)

¹⁷ TNO, 2024

¹⁸ [Algoritmeregister](#)

¹⁹ [Over het algoritmeregister](#)

Tabel 1: aantal interviews per type overheidsorganisatie

Type organisatie	Aantal interviews
Gemeente	6
Provincie	3
Netwerkorganisatie	3
Agentschap	2
Regionaal samenwerkingsorgaan	1
Ministerie	1
Private partij	1

Analyse van de interviews vond plaats op basis van thematische analyse.²⁰ Deze methode laat toe om vooraf geselecteerde thema's op te halen en ook nieuwe thema's te identificeren in de data.²¹ De gepresenteerde voorbeelden uit de praktijk aan het eind van dit rapport zijn gebaseerd op zowel de interviews als op publiek beschikbare informatie.

Onderzoeksmethode deel 3: Praktijkvoorbeelden

Het doel van het derde deel van dit onderzoek is om inzicht te krijgen in overwegingen die worden gemaakt bij het ontwikkelen en inzetten van GenAI door Nederlandse overheidsinstellingen. Dit doen we aan de hand van **vier praktijkvoorbeelden**.

In elk voorbeeld worden resultaten uit deel 1 en deel 2 van dit onderzoek gecombineerd. De kenmerken van de toepassing, zoals doel, ontwikkelvorm en technische kenmerken komen uit deel 1. Aanvullingen op deze kenmerken en de inzichten over succesfactoren en knelpunten komen uit de interviews deel 2. Daarnaast is voor het uitdiepen van de praktijkvoorbeelden extra aandacht besteed aan de overwegingen of belangrijke keuzes die in de interviews zijn besproken door de respondenten.

De keuze voor welke praktijkvoorbeelden worden toegelicht is gemaakt op basis van de kenmerken van de GenAI toepassing, en onderlinge variaties, zodat verschillende contexten aan bod komen. De kenmerken zijn onder andere: doel van de toepassing, eindgebruiker van de toepassing, ontwikkelvorm en betrokken partijen bij ontwikkeling en inzet van de toepassing en technische kenmerken van de toepassing.

Dit heeft geleid tot de volgende selectie:

1. Chatbot GEM
2. Leeshulp Tolkie
3. Archiefassistent Nationaal Archief
4. Transcriptie tool WerkSaam Westfriesland

²⁰ Braun & Clarke, 2006

²¹ Braun et al., 2019

3. Inzet van GenAI bij Nederlandse Overheidsorganisaties

81 toepassingen van GenAI bij Nederlandse Overheidsorganisaties

Op basis van desk research zijn 81 voorbeelden van GenAI toepassingen bij Nederlandse overheidsorganisaties gevonden (zie bijlage: 'Lijst met voorbeelden GenAI in de overheid'). De gevonden voorbeelden zijn gecategoriseerd en geanalyseerd aan de hand van de volgende categorieën:

- Type overheidsorganisatie
- Status toepassing
- Type proces
- Type taak
- Eindgebruiker toepassing
- Risico inschatting
- Impact assessment
- Technische kenmerken van GenAI

Type overheidsorganisatie

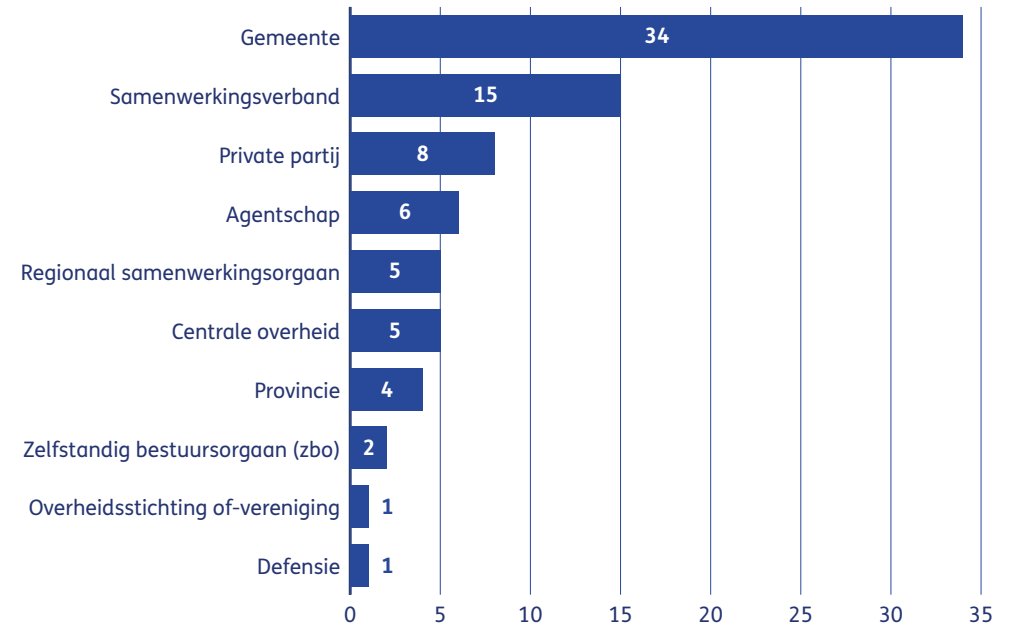
Allereerst is er gekeken naar het type overheidsorganisaties die GenAI inzetten:

- 34 toepassingen (42%) bij **gemeenten**. Dit betreft voornamelijk chatbots die assisteren in het vergaren en verwerken van informatie voor ambtenaren en die assisteren in het contact met de burger. De meeste toepassingen zijn dan ook gevonden bij gemeenten. Een mogelijke verklaring hiervoor is tweeledig. Ten eerste, er zijn in totaal 342 gemeenten in Nederland. Hierdoor hebben ze een

groot aandeel in het totaal aantal overheidsorganisaties in Nederland. Ten tweede, de meeste registraties in het algoritmeregister zijn afkomstig van gemeenten.

- 15 toepassingen in **samenwerkingsverbanden**, waarbij verschillende type overheidsorganisaties en eventueel private partijen gezamenlijk een toepassing ontwikkelen. Denk bijvoorbeeld aan het samenwerkingsverband van Chatbot GEM. Een voorbeeld is geclassificeerd als samenwerkingsverband (TNO, 2021; TNO, 2024) indien:
 - Meer dan drie verschillende partijen betrokken zijn bij de ontwikkeling van het voorbeeld; of
 - Er sprake is van samenwerking tussen verschillende type overheidsorganisaties (bijv. provincie en gemeenten, inspectie en centrale overheid).
 - 8 toepassingen worden aangeboden door **private partijen**, maar het is onduidelijk wat de reikwijdte is van het gebruik van deze toepassingen.
- 5 toepassingen bij de **centrale overheid**, zoals ministeries. Dit zijn voornamelijk chatbots of kennisverwerkingshulpmiddelen voor ambtenaren.

Figuur 1: Type overheidsorganisaties



- 6 toepassingen bij **agentschappen**, ook wel uitvoeringsorganisaties zoals KNMI en RVO.
- 5 toepassingen bij **regionaal samenwerkingsorganen** en 4 toepassingen bij **provincies**.
- 1 toepassing bij **Defensie**. Dit betreft DefGPT, een intern alternatief voor de

bekende commerciële AI-diensten voor medewerkers van Defensie.

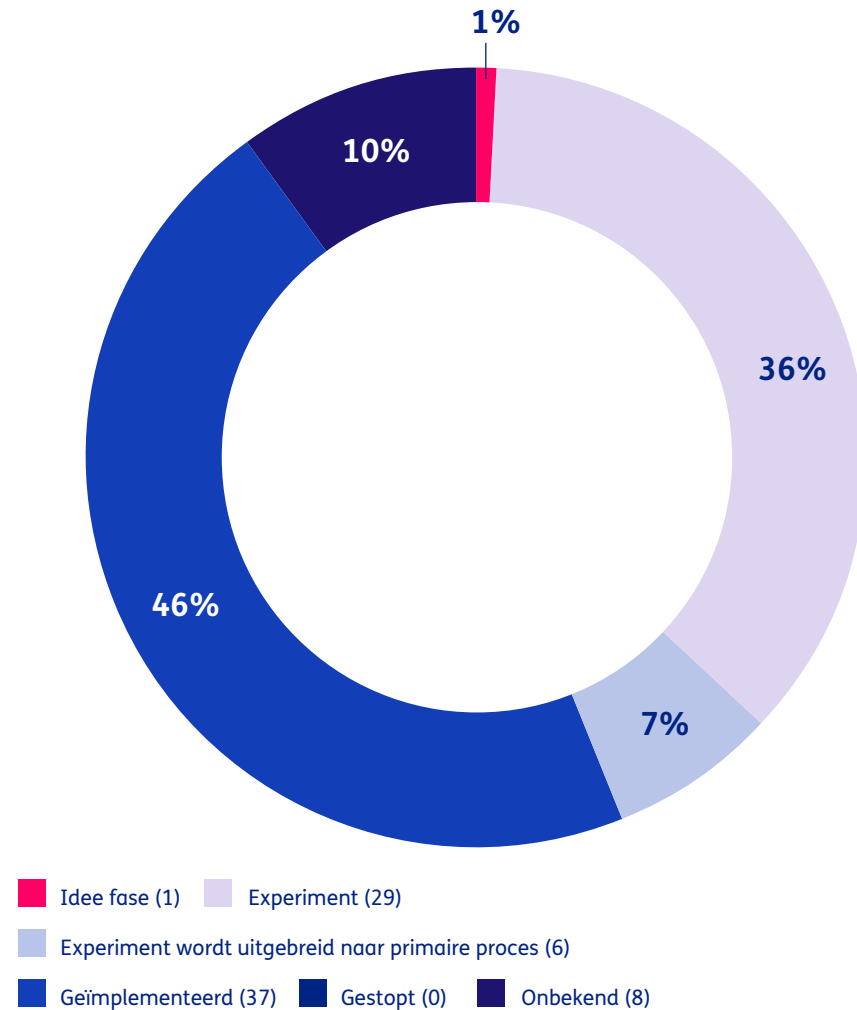
- 1 toepassing bij een **ZBO** en bij **overheidsstichting ICTU**.

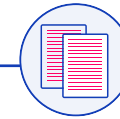
Status toepassing

Toepassingen bevinden zich in verschillende fasen van ontwikkeling.

- **Idee fase.** Een toepassing bevindt zich in de idee fase als het experiment of pilot nog moet starten. Er is 1 voorbeeld gevonden dat zich in de **idee fase (1%)** bevindt en nog moet starten. Waarschijnlijk zijn er nog veel meer toepassingen die zich in de idee fase bevinden. Informatie over deze toepassingen worden doorgaans pas openbaar toegankelijk zodra ze de experimentele fase bereiken.
- **Experiment.** Een toepassing bevindt zich in de experiment fase indien het nog in ontwikkeling is, een proof of concept is of pilot is. 29 toepassingen (36%) zijn momenteel in de **experiment fase**. Dit zijn meestal pilots waarbij de werking en toegevoegde waarde wordt getest.
- **Uitgebreid naar het primaire proces.** Een toepassing wordt uitgebreid naar het primaire proces als het bezig is de stap te maken van pilot naar implementatie. 6 experimenten worden momenteel **uitgebreid naar het primaire proces (7%)** en worden voorbereid om geïmplementeerd te worden genomen.
- **Geïmplementeerd.** Een toepassing wordt als geïmplementeerd beschouwd als het in gebruik is genomen door de organisatie, voorbij de pilot fase is en onderdeel uit maakt van de dagelijkse praktijk. Op dit moment zijn 37 toepassingen (46%) geïmplementeerd.
- We hebben geen toepassingen gevonden die zijn stopgezet of momenteel buiten gebruik zijn.
- Van 8 toepassingen is de huidige status onbekend (10%).

Figuur 2: Status toepassingen





Democratisch proces

GenAI wordt ingezet ter bevordering van het democratisch proces. Bijvoorbeeld ter ondersteuning van een gemeenteraad.

Maatwerk & dienstverlening

Personalisatie van dienstverlening voor burgers en bedrijven. Deze toepassing is extern gericht, bijvoorbeeld het gebruik van Chatbots.

Generieke processen & bedrijfsvoering

Inzet van GenAI in interne processen ter bevordering van bedrijfsvoering. Bijvoorbeeld een interne Chatbot die ambtenaren gebruiken voor meerdere doeleinden, bijvoorbeeld vertalen, transcriberen, opzoeken van informatie of schrijven van e-mails.

Kennisvergaring & beleidsvorming

GenAI dat wordt ingezet voor het opdoen van nieuwe informatie waar (beleids)keuzes op worden gebaseerd. Dit betreffen specifieke toepassingen die ondersteuning bieden bij een specifiek beleidsdoel.

Kennisverwerking, archivering & anonimiseren

GenAI wordt ingezet voor het verwerken en archiveren van informatie, en voor anonimiseren van persoonsgegevens in documenten.

Type proces: categorisering

GenAI toepassingen worden in verschillende type overheidsprocessen ingezet. In dit onderzoek definiëren we type proces als processen die zijn ingericht op organisatie-niveau om de doelstelling van een organisatie te behalen. Een aantal van deze processen is extern gericht, zoals het democratisch proces en maatwerk en dienstverlening. Ook zijn er interne processen die zijn gericht op beleidsvorming en bedrijfsvoering. De hier

gepresenteerde categorisering is verfijnd en gebaseerd op de categorie ‘type toepassing’ uit de derde versie die is opgesteld door TNO voor de ‘Quickscan AI in de publieke Dienstverlening III.’²² Ten opzichte van de indeling uit dit onderzoek²³ vallen een aantal aspecten op. Ten eerste, de geïdentificeerde GenAI toepassingen worden met name ingezet in overheidsprocessen waarbij een hoge mate van gegevens en tekstverwerking plaatsvindt. Er zijn er geen toepassingen in de fysieke

omgeving geïdentificeerd. De aard van machine learning toepassingen zijn dan ook diverser dan GenAI toepassingen. Ten tweede valt op dat de geïdentificeerde toepassingen geen nieuwe toepassingen zijn, maar het vooral verbeterde toepassingen van tekstherkenning betreffen. Wel valt één ander aspect op: GenAI toepassingen kunnen gemakkelijker voor meerdere doeleinden beschikbaar worden gesteld. Er is dan ook één nieuwe categorie toegevoegd: generieke processen &

bedrijfsvoering. Deze categorie is gericht op de inzet van GenAI in interne processen ter bevordering van bedrijfsvoering waarbij ambtenaren de toepassing voor meerdere type taken kunnen gebruiken. De andere twee interne categorieën zijn kennisvergaring & beleidsvorming en kennisverwerking, archivering & anonimiseren. Bij deze twee categorieën worden GenAI toepassingen ingezet voor een specifiek proces met een specifieke taak.

22 TNO, 2024

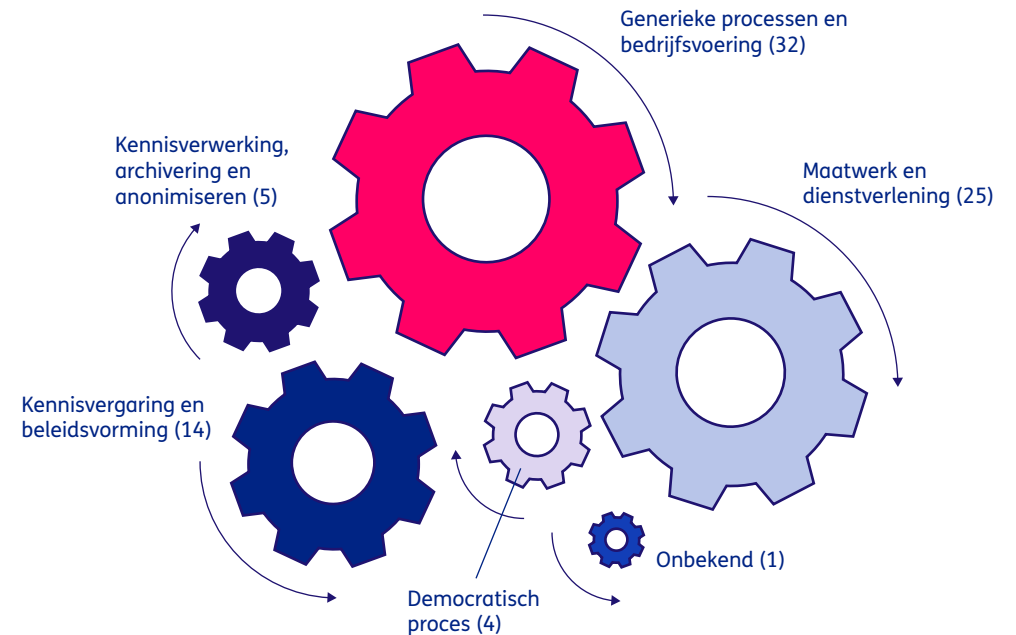
23 Ibid.

De geïdentificeerde GenAI toepassingen worden met name ingezet in overheidsprocessen waarbij een hoge mate van gegevens en tekstverwerking plaatsvindt.

- 32 toepassingen (40%) worden ingezet voor **generieke processen en bedrijfsvoering**, waarbij vaak gebruik wordt gemaakt van een interne chatbot, die alleen beschikbaar is voor medewerkers, om deze processen te ondersteunen.
- 25 toepassingen (31%) zijn gericht op **maatwerk en dienstverlening**, deze processen zijn extern gericht, zoals personalisatie van dienstverlening en interactie met burgers en bedrijven. Een voorbeeld van inzet van GenAI in het proces van maatwerk en dienstverlening is een chatbot op een gemeentelijke website.

- 14 toepassingen (17%) zijn gericht op **kennisvergaring en beleidsvorming**, waarmee medewerkers inzichten verzamelen om beter onderbouwde beslissingen te nemen.
- 5 toepassingen (6%), zijn ontwikkeld voor **kennisverwerking, archivering en het anonimiseren** van gegevens.
- Tot slot worden 4 toepassingen (5%) gebruikt ter ondersteuning van het **democratisch proces**, bijvoorbeeld bij ondersteuning bij verslaglegging bij vergaderingen van de gemeenteraad of bij het ondersteunen van ambtenaren bij het beantwoorden van Kamervragen.

Figuur 3: Type proces



Tabel 2: Type taak

Type taak	Aantal
Generieke taken	24
Opzoeken en analyseren van informatie	18
Contact met burger	11
Tekstoptimalisatie	6
Vertalen	5
Transcriberen	4
Controleren van gegevens	3
Doorzoekbaar maken van informatie	3
Samenvatten van informatie	2
Onbekend	2
Anonimiseren	1
Contact met bedrijven	1
Sentimentanalyse	1

Type taak

Type taak definiëren we in dit onderzoek als een activiteit dat wordt uitgevoerd om een type proces goed uit te kunnen voeren. Een taak kan daarom ook worden gezien als praktische acties en werkzaamheden die medewerkers uitvoeren. In dit onderzoek hebben we in kaart gebracht bij welke

taken GenAI toepassingen ondersteuning bieden.

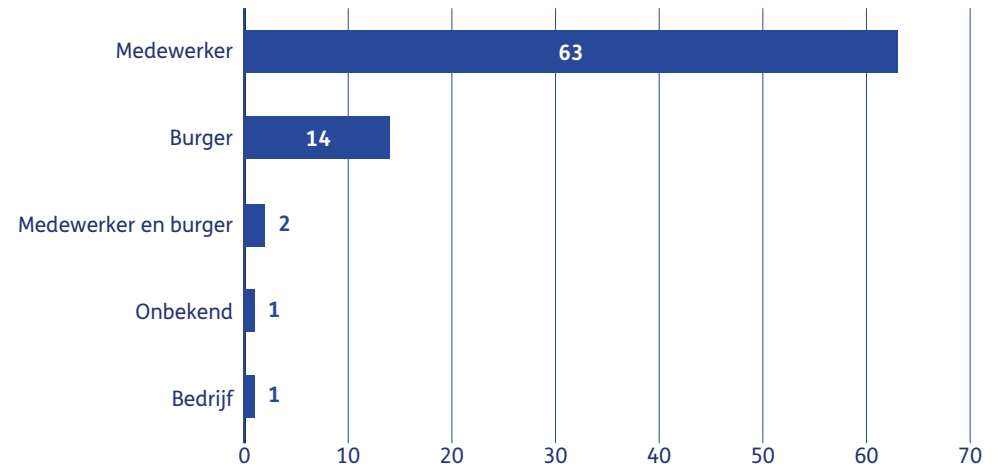
- 24 (30%) van de toepassingen zijn chatbots die gericht zijn op het ondersteunen van medewerkers bij **meerdere generieke taken en kunnen dus worden ingezet voor meerdere doeleinden**. Voorbeelden van toepassingen zijn Co-Pilot, DefGPT van het ministerie van Defensie, Chatbot Robin van JenV, VlamChat van SSC-ICT, PZH Assist van de Provincie Zuid-Holland, GovChat/LAICA van de Provincie Limburg en ChatAmsterdam van de gemeente Amsterdam.
- De overige 55 toepassingen (68%) worden ingezet bij specifiekere taken en vaak voor één specifiek doel. Zo ondersteunen bijvoorbeeld 18 (22%) van de toepassingen bij het **opzoeken en analyseren van informatie**, wat medewerkers helpt bij het ondersteunen van besluitvorming en kennisdeling.
- 11 (14%) toepassingen richt zich op **contact met burgers**, meestal via een gemeentelijke chatbot die 24/7 beschikbaar is.
- Andere voorbeelden van toepassingen met een specifieke taak zijn:

- Optimaliseren van tekst voor contentcreatie bij RVO
- Transcriberen van audio tijdens clientgesprekken bij WerkSaam Westfriesland
- Versimpelen van teksten op de gemeentelijke website voor laaggeletterden bij de gemeente Wijchen en de gemeente Druten
- Uitvoeren van sentimentanalyse over gevoelige gesprekken bij de gemeente Amsterdam

Eindgebruiker toepassing

- Bij 63 toepassingen (78%) zijn **medewerkers** van overheidsorganisaties de eindgebruiker.
- Bij 14 toepassingen (17%) zijn **burgers** de beoogde eindgebruiker. Dit betreffen met name chatbots op gemeentelijke websites, maar ook toepassingen die laaggeletterden helpen door teksten op websites versimpelen.
- Bij 2 toepassingen (3%) zijn zowel **medewerkers als burgers** de eindgebruiker de beoogde eindgebruiker.
- Bij 1 toepassing zijn **bedrijven** de beoogde eindgebruiker.
- Van 1 toepassing is **onbekend** wie de beoogde eindgebruiker is.

Figuur 4: Eindgebruiker toepassing



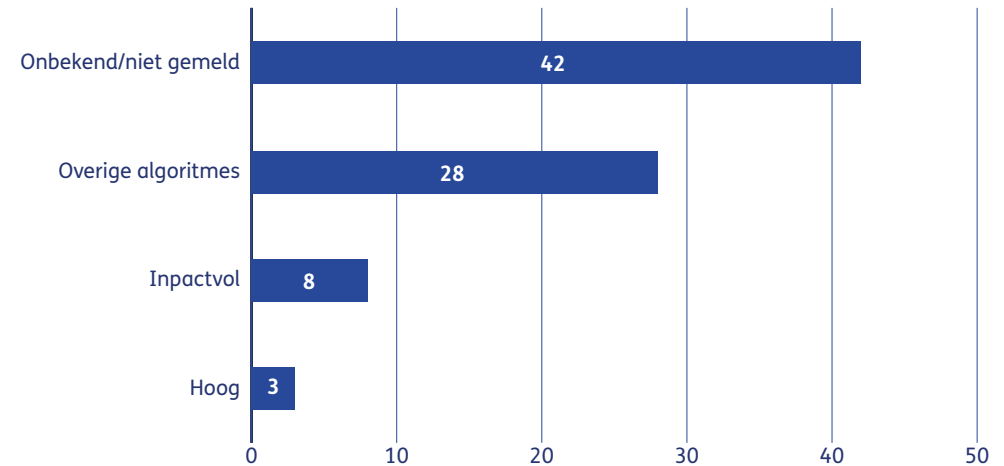
Risico inschatting

- Het gebruik van (Gen)AI brengt ook risico's met zich mee. Om de mogelijke impact van algoritmes te bepalen en om te bepalen welke algoritmes gepubliceerd moeten worden heeft het algoritmeregister drie publicatie-categorieën gedefinieerd in hun publicatiestandaard²⁴:
 - **Hoog-risico AI systeem** (categorie A): deze categorie wordt gedefinieerd in de AI-verordening van de EU.
 - **Impactvolle algoritmes** (categorie B): het algoritme heeft directe rechtsgevolgen voor een betrokkene en/of beïnvloedt hoe de overheid een betrokkene of groep classificeert.
 - **Overige algoritmes** (categorie C): indien categorie A of B niet van toepassing is kunnen er andere overwegingen meespelen om algoritmes te publiceren, bijvoorbeeld omdat een algoritme onderwerp is van het maatschappelijk debat of een indirecte impact heeft op de omgeving.
- Dit onderzoek heeft deze categorieën overgenomen en daarbij ook de categorie onbekend/niet gemeld

toegevoegd. De resultaten die zijn weergegeven in de grafiek zijn gebaseerd op informatie uit het algoritmeregister en geeft dus inzicht in hoe overheidsorganisaties zelf een risico inschatting maken van een GenAI toepassing. Hierbij is het ook goed om te benoemen dat er momenteel geen verplichting is om algoritmes te registreren in het algoritmeregister.

- De resultaten die zijn weergegeven in figuur 5 zijn gebaseerd op informatie uit het algoritmeregister en geeft dus inzicht in hoe overheidsorganisaties zelf een risico inschatting maken van een GenAI toepassing.
- Van 42 van de geïdentificeerde GenAI voorbeelden (52%) is het **onbekend** in welke categorie de toepassing valt, bijvoorbeeld omdat het voorbeeld niet in het algoritmeregister staat, als er geen publieke informatie over gedeeld is of als het niet duidelijk is welke definities voor de risico categorieën zijn gebruikt om de inschatting te maken.
- 28 toepassingen (34%) zijn ingedeeld door overheidsorganisaties in de categorie **overige algoritmes**.

Figuur 5: Risico inschatting van toepassing



- 8 toepassingen (10%) zijn ingedeeld door overheidsorganisaties in de categorie **impactvol**.
- Slechts 3 toepassingen (4%) zijn door overheidsorganisaties ingedeeld in de categorie **hoog risico AI systeem**.
- Ook observeren we dat een dezelfde soort toepassing bij verschillende organisaties in een andere risico categorie wordt ingedeeld. Ter illustratie zijn hierbij twee voorbeelden:
 - **Chatbots op de website van gemeenten**. De toepassing Chatbot Guus is door de gemeente Goes

- ingedeeld in de categorie overige algoritmes, terwijl Chatbot MAI door de gemeente Montferland is ingedeeld in de categorie impactvolle algoritmes.
- **CoPilot**. De gemeente Zeist en het Commissariaat van de Media heeft het gebruik van CoPilot ingedeeld als Hoog-risico AI-systeem, terwijl de Veiligheidsregio Zuid-Holland Zuid, ICTU, Provincie Groningen, Veiligheidsregio Zaanstreek-Waterland en Veiligheidsregio Amsterdam-Amstelland het gebruik van CoPilot hebben ingedeeld in de categorie overige algoritmes.

²⁴ [Handreiking Algoritmeregister](#)

Impact toetsen

Naast een risico inschatting, is het in sommige gevallen verplicht om een impact toets uit te voeren. Er bestaan verschillende vormen van impact toetsen. De toetsen die door veel organisaties worden genoemd zijn:

- Een Data Protection Impact Assessment (DPIA)²⁵ is een instrument waarmee mogelijke privacy risico's van gegevensverwerkingen in kaart gebracht kunnen worden en maatregelen worden gedefinieerd om de risico's te voorkomen. Een DPIA is een vereiste onder de Algemene Verordening Gegevensbescherming (AVG) wanneer er (grote hoeveelheden) persoonsgegevens worden verwerkt of bij nieuwe technologieën. Er bestaat ook een verkorte variant zoals de DPIA light, en een toets of een DPIA nodig is, de pre-DPIA.²⁶
- Het Impact Assessment Mensenrechten en Algoritmes (IAMA)²⁷ helpt bij gespreksvoering over grondrechten, data en ethiek bij het ontwikkelen en inzetten van AI.

- Andere vormen van impact toetsen of ethische toetsen die ook bekend zijn in de praktijk, zijn:
 - Aanpak Begeleidingsethiek is een methode die organisaties helpt bij het maken van verantwoorde keuzes bij de ontwikkeling en inzet van algoritmes en AI, waarbij ethiek als hulpmiddel in plaats van remmende factor wordt gezien. Het brengt verschillende betrokkenen samen en laat ze samen de casus uitdiepen, bespreken welke actoren, waarden en effecten betrokken zijn en het werkt uiteindelijk toe naar het formuleren van handelingsopties.
 - Het CODIO-instrument²⁸ is bedoeld als instrument om de afweging en borging van waarden te verzorgen. Dit instrument heeft ook meerwaarde voor de inzet van GenAI zodat publieke waarden ook meegenomen kunnen worden.
 - Sommige organisaties hebben een eigen kader of methode ontwikkeld om een impact toets of eigen toets uit te voeren. Ook wordt er gebruikt gemaakt van ethische commissies

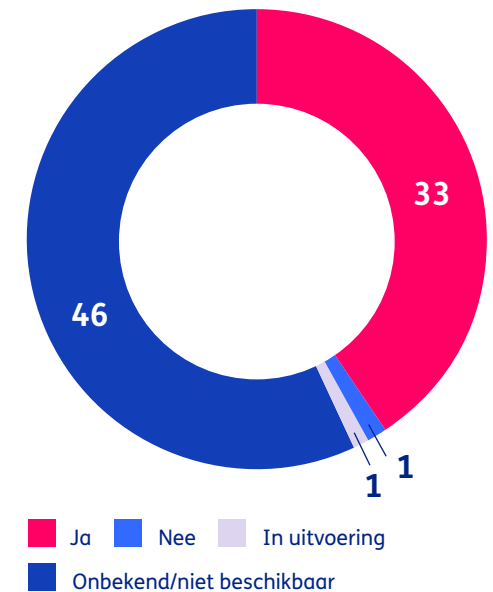
om een (vorm van) een impact toets uit te voeren en ethische vraagstukken mee te nemen bij de ontwikkeling of inzet van GenAI.

Om vast te stellen of er een impact toets is uitgevoerd, zijn de volgende bronnen gebruikt:

- Algoritmeregister: Het algoritmeregister heeft een veld waarin overheidsorganisaties kunnen aangeven of er een impact toets is uitgevoerd.
- Publiek beschikbare informatie: in sommige gevallen is bekend welke impact toets is uitgevoerd op basis van rapporten of andere publicaties.
- Interviews: in de interviews is expliciet gevraagd naar uitgevoerde of geplande impact toetsen. Deze inzichten zijn gebruikt als aanvulling.
- 33 toepassingen (41%) hebben aangegeven een vorm van een impact toetsen te hebben **uitgevoerd**. Dit kan een of meerdere impact toetsen zijn die op de vorige pagina staan beschreven.
- Bij 46 toepassingen (56%) is het

onbekend of er een impact toets is uitgevoerd. Dit betekent echter niet automatisch dat er géén toets is gedaan of nog zal worden uitgevoerd.

- 1 toepassing (1%) geeft aan geen vorm van een impact toets te hebben uitgevoerd omdat de toepassing als laag risico is geclassificeerd.



²⁵ [Data protection impact assessment \(DPIA\)](#)

²⁶ [Impact Assessment Mensenrechten en Algoritmes](#)

²⁷ [Aanpak Begeleidingsethiek](#)

²⁸ [Het CODIO-instrument](#)

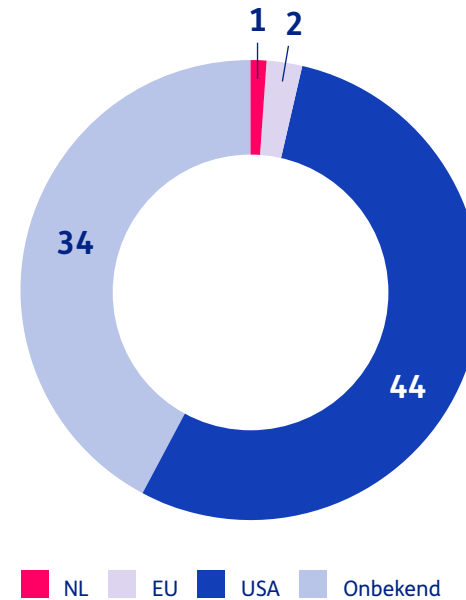
Technische kenmerken van GenAI

In dit onderzoek hebben we een poging gedaan om ook informatie te verzamelen over kenmerken die inzicht geven in hoe de 81 GenAI toepassingen op technisch vlak worden vormgegeven. We hebben informatie opgevraagd over: de ontwikkelvorm, type model, land van oorsprong van het model, type aanbieder, land van oorsprong aanbieder. In de praktijk blijkt dat informatie over technische specificaties en de onderliggende digitale infrastructuur niet altijd publiekelijk beschikbaar of niet volledig is.

We observeren het volgende over type modellen, omgang met data en ontwikkelvormen:

- **Oorsprong foundation modellen.** Van 34 toepassingen (42%) is het onbekend wat het land van oorsprong van het foundation model is. Bij 44 toepassingen (54%) zijn de gebruikte foundation modellen ontwikkeld door bedrijven die zijn gevestigd in de Verenigde Staten. Ook wanneer de leveranciers van GenAI toepassingen Nederlandse bedrijven zijn, zijn de foundation modellen vaak ontwikkeld door OpenAI, Meta, Microsoft of Google. Europese modellen die ook voorkomen zijn van DeepL of Mistral.

- **Ontwikkelvorm.** Overheidsorganisaties kunnen ervoor kiezen om (Gen)AI-toepassingen in-house te ontwikkelen, in een samenwerkingsverband te ontwikkelen of in te kopen. De meerderheid van toepassingen zijn ingekocht, maar er zijn ook initiatieven die in samenwerkingsverband worden ontwikkeld of waar overheidsorganisaties de toepassing bouwen met bestaande (open source) modellen.
- **Van klassieke NLP technieken naar GenAI.** Een aantal toepassingen maakte oorspronkelijk gebruik van klassieke NLP technieken maar zijn nu sinds de introductie van GenAI overgestapt op GenAI. Een voorbeeld hiervan zijn Chatbot Guus van de gemeente Goes en Chatbot GEM.
- **Data en infrastructuur.** Op het gebied van data wordt veelal expliciet gerapporteerd dat eigen **data niet (weer) wordt gevoed in systemen** voor verdere training van taalmodellen, of **hosting lokaal wordt ingericht**.



4. Succesfactoren en knelpunten bij de inzet van GenAI bij Nederlandse Overheidsorganisaties

Succesfactoren en knelpunten bij GenAI inzet

Deel 2 van dit onderzoek richt zich op de **succesfactoren en knelpunten** die worden ervaren bij de inzet van GenAI bij overheidsorganisaties. Deze inzichten zijn opgehaald uit de interviews die we in het kader van dit onderzoek hebben afgenomen. Zoals beschreven in de onderzoeksmethode, heeft TNO 17 interviews uitgevoerd met sleutelrollen om een beter beeld te krijgen van de ervaringen in de Nederlandse publieke sector met GenAI toepassingen. De deelnemers zijn veelal innovatiemanagers die betrokken zijn geweest bij het ontwikkelen en inzetten van één of meerdere GenAI toepassingen.

Naast de succesfactoren en knelpunten, hebben de respondenten ook inzicht gegeven in de beweegredenen om te beginnen aan de ontwikkeling en inzet van GenAI.

De verzameling aan succesfactoren en knelpunten kan helpen bij het identificeren van kansen voor verdere GenAI inzet en het mitigeren van knelpunten. In de volgende pagina's schetsen we welke succesfactoren en knelpunten zijn besproken in de interviews. Of iets een succesfactor en knelpunt is, is

context-afhankelijk. Dat wil zeggen, de impact en mate van invloed hangt af van de situatie en hoe verschillende aspecten samenkomen in een GenAI toepassing. Daarnaast kan een knelpunt in een voorbeeld juist ook zorgen voor innovatieve oplossingen als er effectief of creatief mee wordt omgegaan.

Aanleiding voor GenAI inzet

In de interviews noemen respondenten een aantal redenen waarom organisaties ervoor kiezen om specifiek met GenAI als technologie aan de slag te gaan:

- **GenAI spreekt tot de verbeelding:** Meerdere respondenten deelden de ervaring dat het gemakkelijk is voor zowel experts als niet-experts om met GenAI aan de slag te gaan. GenAI spreekt volgens de respondenten tot de verbeelding en geeft inspiratie voor nieuwe ideeën en toepassingen. Respondent 4 zegt hierover: *“Mensen kunnen snel zelf aan de slag met ideeën. In de praktijk leer je, dan loop je tegen dingen aan die niet al te grote problemen verzorgen en niet met teveel risico's komen maar waarbij je toch al heel veel leert zodat je steeds dieper kunt gaan.”*

- **Evolutie van taaltechnologieën en schaalbaarheid:** Door recente ontwikkelingen in taaltechnologieën is de kwaliteit van de output in de Nederlandse taal van dermate hoog niveau dat GenAI een levensvatbare oplossing is geworden voor overheidsorganisaties.
- **Efficiëntie van een scriptbot ten opzichte van een GenAI bot:** Volgens respondent 6 zijn er, in tegenstelling tot klassieke NLP technieken, geen grote databases en scripts nodig met verschillende definities van woorden voor GenAI toepassingen. In vergelijking met klassieke NLP toepassingen kost de ontwikkeling van een toepassing hierdoor ook veel minder tijd. Daarnaast is de techniek een stuk simpeler wat betreft beheer. Andere respondenten noemen wel dat het belang van compliance bij GenAI echter wel groter is geworden en dat daar nu meer bij komt kijken.
- **Verwachting over de inzet van GenAI bij het oplossen van maatschappelijke opgaven.** Respondenten geven aan dat er veel aandacht is voor GenAI en de mogelijke meerwaarde die de technologie kan spelen bij huidige en

toekomstige uitdagingen, zoals krapte op de arbeidsmarkt door vergrijzing. Tegelijkertijd zijn overheidsorganisaties ook bewust dat er risico's zijn voor digitale autonomie bij GenAI.

- **Druk om GenAI te gebruiken.** Door de grote aandacht die er op dit moment is voor generatieve AI, de hoge verwachtingen die er zijn wordt door organisaties ook druk ervaren om GenAI in te zetten. *“Je wordt helemaal gek van die hype om generatieve AI heen. Alle aanbieders van oplossingen en trainingen. Daar zit een soort FOMO-gevoel bij. [...] Dat je bijna het gevoel hebt van, ik moet nu iets doen. Terwijl AI bestaat natuurlijk eigenlijk al 70 jaar. En machine learning deden wij de overheid al hartstikke lang. Maar nu voelt het ineens alsof er is en dat je er gelijk iets mee moet.”* (Respondent 1)
- **Voorkomen van schaduwgebruik van GenAI binnen de organisatie.** Overheidsorganisaties zien dat medewerkers GenAI toepassingen gebruiken in hun werk, maar beseffen ook dat dit risico's met zich meebrengt. Daarom bieden ze graag een veilig en verantwoord alternatief aan. Respondent 11 zegt hierover: *“We zien*

gigantisch veel gebruik naar dit soort tooling [GenAI]. We kunnen het afsluiten, maar het water vindt wel een weg en dan raken we het helemaal uit het zicht. Dus laten we een alternatief bieden.”

Succesfactoren bij de inzet van GenAI

In **tabel 3** staan de geïdentificeerde succesfactoren uit de interviews weer-geven. Deze succesfactoren worden in dit onderdeel nader toegelicht.

De succesfactoren die in de interviews zijn besproken, zijn niet uniek voor GenAI maar gelden ook voor andere (digitale) innovaties. Deze overlap is logisch, omdat GenAI ook een vorm van innovatie is. De snelle ontwikkeling en verspreiding van deze technologie en de impact op mens en maatschappij maakt dat aandacht voor en reflectie op deze succesfactoren extra belangrijk is.^{29, 30}

Samenwerking binnen het innovatie ecosysteem in Nederland

Uit de interviews is gebleken dat er een overwegend positieve houding is ten aanzien van samenwerking op innovatie met GenAI. De respondenten zijn zelf actief in

samenwerkingsverbanden of geven aan dat samenwerking nog sterker gestimuleerd en beter ingericht kan worden. De genoemde voordelen van samenwerking zijn het delen van middelen, lasten en ideeën. Hoewel er ook fragmentatie te zien is op toepassingsgebied, zoals de verschillende initiatieven rondom chatbots laten zien, wordt er vanuit meerdere overheidsinitiatieven, zoals Platform AI en Overheid, gewerkt aan het samenbrengen van partijen en van elkaar te leren. Respondent 1 zegt: *“Alleen maar ambtenaren gaan het niet redden. Alleen maar ondernemers gaan het niet redden. Alleen maar universiteiten gaan het niet redden. Je moet echt gezamenlijk optrekken, anders ga je er gewoon niet komen.”* Een voorwaarde om dit succesvol te doen is openheid naar anderen en de bereidheid om niet alleen aan de slag te gaan.

Maatschappelijke opgave centraal zetten

Een aantal van de geïnterviewde initiatieven zijn ontstaan vanuit de wens om een oplossing of bijdrage te leveren aan een maatschappelijke opgave, bijvoorbeeld de verbetering van dienstverlening richting de burger (voorbeeld Werksaam West-Friesland en GEM) en het helpen van

Tabel 3: Overzicht van succesfactoren bij de inzet van GenAI

Succesfactoren bij de inzet van GenAI
Samenwerking binnen het innovatie ecosysteem in Nederland
Maatschappelijke opgave staat centraal
Duidelijke probleemdefinitie en doel
innovatiebewustzijn en ruimte voor experimenteren
Leiderschap als positief boegbeeld
Bewust omgaan met mogelijk schaduwgebruik
Betrekken van de gebruiker tijdens ontwikkeling
Betrekken van domein experts tijdens ontwikkeling
Voldoende middelen (tijd, geld, expertise)
Formalisatie van processen en verantwoordelijkheden
Stapsgewijs aandacht voor privacy, security en ethische aspecten

laaggeletterden (voorbeeld Leeshulp). Hierdoor wordt de inzet van GenAI een middel, in plaats van een doel. Ook zorgt dit voor een duidelijke scope en nauwe aansluiting bij de praktijk.

Duidelijke probleemdefinitie en doel

Toepassingen die worden ontwikkeld vanuit een goed gedefinieerde probleemstelling uit de praktijk en waarvan de

meerwaarde voor een organisatie duidelijk is, worden als waardevoller beschouwd door de respondenten dan toepassingen die worden ontwikkeld vanuit een technische drijfveer. Volgens respondenten komt dit omdat er breder wordt nagedacht over de toepassing, dat er een scherpe kosten-baten analyse wordt uitgevoerd en dat er een plan komt over hoe een toepassing moet landen in een organisatie. Dat maakt

²⁹ Cinar et al., 2019

³⁰ Jöhnk et al., 2021

het ook gemakkelijker om goedkeuring te krijgen om te starten met een innovatie-traject.

Innovatiebewustzijn en ruimte voor experimenteren

Uit de interviews blijkt dat er draagvlak binnen de organisatie moet bestaan om met innovaties en met GenAI aan de slag te gaan. Een visie of strategie op AI kan een belangrijk element zijn om de visie en strategie kenbaar te maken aan iedereen in de organisatie en ruimte te creëren voor innovatietrajecten.

Uit de verhalen van respondenten blijkt dat het opmerken en aangrijpen van kansen voor innovatie wordt gestimuleerd door een combinatie van durf, proberen en leren door te doen. Een positieve houding ten aanzien van fouten maken of initiatieven moeten stopzetten is daarin ook waardevol. Respondent 5: *“Wij hebben het altijd ingeschoten met het idee van: wij willen leren over werken met AI. En een bot is de meest simpele manier om dat te doen. En we willen het als vliegwiel gebruiken voor andere AI toepassingen.”* Door meer kennis van en ervaring met GenAI op te bouwen, ontstaat er meer inzicht over wat er mogelijk is en wat de impact is in de praktijk.

Leiderschap als positief boegbeeld

Een positieve houding van leiderschap ten aanzien van innovatie en GenAI creëert ook draagvlak voor GenAI ontwikkeling en gebruik. Een goede relatie met leiderschap en beslissers in organisaties werkt ook faciliterend in processen om middelen te bemachtigen of toegang te krijgen tot nodige middelen en bronnen.

Bewust omgaan met mogelijk schaduwgebruik

Respondenten noemden echter ook voorbeelden waar het (schaduw) gebruik van medewerkers signaleerde dat een organisatie regels moest inrichten of tools moest aanbieden aan hun medewerkers. Respondent 5 zegt hierover: *“We hebben ook gewoon regels opgesteld. Als wij je AI tools leveren, hoef je niet na te denken, mag je alles doen wat je wilt. Om het gewoon helder te houden. Ook om grijs gebruik van AI tegen te gaan. Maar ga je zelf iets zoeken, dan is het ook jouw verantwoordelijkheid. En dan zijn er hier een hele lading aan spelregels.”*

Een andere organisatie is opzoek gegaan naar een alternatief voor ChatGPT om schaduwgebruik door medewerkers tegen te gaan: *“Dus we zagen dat het gebruik van ChatGPT onder onze collega’s echt gigantisch toenam. Daar kwam ook wel de zorg uit van: wat gebeurde met de data? Is het*

wel veilig? Nou ja goed de zorgen die iedereen wel kent. En toen hebben we besloten om een veilige alternatief te tegenover te stellen in plaats van de boel helemaal dicht te gooien en te verbieden.” Deze ervaring van respondent 9 toont aan dat richtlijnen en beleid het gedrag van medewerkers niet volledig kunnen sturen, maar dat er ook ruimte in beleid kan worden gevonden of gemaakt om mee te bewegen met innovatie en verantwoord gebruik mogelijk te maken.

Betrekken van de gebruiker tijdens ontwikkeling

Tijdens een aantal interviews kwam naar voren dat het belangrijk is om de gebruiker vanaf het begin te betrekken bij de ontwikkeling. Dit bevordert draagvlak voor de toepassing onder gebruikers. Ook helpt nauw contact met de gebruikersgroep om vroegtijdig inzicht te krijgen in de behoeften en wensen van gebruikers. Dit biedt ook de kans om te testen of de toepassing de juiste functionaliteiten heeft, gebruiksvriendelijk is en op de beoogde wijze wordt gebruikt.

Betrekken van domein experts tijdens de ontwikkeling

Ook kwam tijdens een aantal interviews naar voren dat het belangrijk is om domeinexperts te betrekken met kennis over het domein waarin de toepassing

wordt ingezet. Inzichten in en ervaring met specifieke processen uit de dagdagelijkse praktijk helpen bij het vormgeven van een toepassing en versterkt de bruikbaarheid en meerwaarde voor gebruikers. Dit geldt voor zowel het ontwerpen van de toepassing als de werking en output verfijnen, bijvoorbeeld door het opstellen van prompts.

Voldoende middelen (tijd, geld en expertise)

Een veel genoemde factor, ook bij knelpunten, is de beschikbaarheid van tijd, geld, expertise en andere middelen om te kunnen (starten met) innoveren. Uit de ervaring van de respondenten blijkt echter dat beschikbare tijd een centrale rol heeft, want om succesvol te innoveren, moet er ruimte zijn om met innovatie aan de slag te gaan. Respondent 3 zegt: *“Ik denk de belangrijkste les is: creëer echt ruimte voor innovatie. Op het moment dat je het erbij moet doen, dan is het risico heel groot dat het niet gebeurt [...] En maak er ook geld voor vrij.”*

Formalisatie van processen en verantwoordelijkheden

Wanneer een kleine groep start met samenwerking bestaat er veel flexibiliteit en kunnen beslissingen gemakkelijk gemaakt worden. Zodra meer partners aansluiten moet er meer ingericht en

geformaliseerd worden. Respondent 12 zegt: *“Het begint met twee, drie organisaties, informeel. En dan is gewoon alles en iedereen bij elkaar. Op een gegeven moment wordt dat een groep van twintig man. Super gezellig, maar dat werkt eigenlijk niet meer. En dus maken we een nieuwe stap door een stuurgroep op te zetten.”*

Stapsgewijs aandacht voor privacy, security en ethische aspecten

Meerdere respondenten beschreven hoe zij op een gelaagde manier privacy, security en ethische vraagstukken meenemen. Vanaf het begin van het ontwikkelproces wordt er aandacht aan deze aspecten besteed en als belangrijk ervaren. Maar om door te kunnen met een toepassing, kan niet alles in de vroege fase van ontwikkeling meegenomen worden. Privacy en security aspecten gaan vaak gelden wanneer de toepassing in productie genomen gaat worden. Het helpt om van te voren een duidelijk beeld te hebben van welke aspecten en stappen nodig zijn in het begin van het ontwikkelproces, en welke stappen pas in een latere fase van het ontwikkelproces kunnen worden ondernomen.

Knelpunten bij de inzet van GenAI

In **tabel 4** staan de geïdentificeerde knelpunten uit de interviews weergegeven. Deze knelpunten worden in dit onderdeel nader toegelicht.

Net als bij succesfactoren, zijn een deel van de knelpunten niet specifiek voor GenAI maar toepasselijk voor (digitale) innovatie in het algemeen. Bij knelpunten geven respondenten wel vaker aan wat er specifiek geldt voor een GenAI toepassing dan voor succesfactoren.^{31,32}

Innoveren met GenAI heeft vaak onzekere uitkomsten

De respondenten zijn vanuit hun rol bezig met (stimuleren van) innovatie met GenAI binnen hun overheidsorganisaties. Daarbij kan het zijn dat organisaties het spannend kunnen vinden om met GenAI aan de slag te gaan, en daar terughoudendheid in hebben. *“Als je kijkt naar wat innovatie werkelijk is, dat is niet het gebruiken van nieuwe technologieën. Dat is het ontdekken en leren van onzekere uitkomsten. Dat [deze lerende houding] is veel minder aanwezig.”* (respondent 9). Bovendien, dat geldt voor innovatie in brede zin, experimenten kunnen ook mislukken. Het hoort bij innoveren om iets te proberen, maar er

Tabel 4: Overzicht van knelpunten bij de inzet van GenAI

Knelpunten bij de inzet van GenAI
Innoveren met GenAI heeft vaak onzekere uitkomsten
Ontwikkelingen op het gebied van GenAI zijn moeilijk bij te houden
Risico's op fouten en misbruik van GenAI
Datamanagement en datakwaliteit is niet altijd op orde
Beperkte AI geletterdheid in organisaties
Beperkte centralisering op nationaal niveau
Onduidelijkheid rondom wet- en regelgeving en impact toetsen
Proportionele aandacht voor ethische overwegingen
Benodigde technische infrastructuur
De afwezigheid van budget, bronnen of andere middelen
Digitale afhankelijkheid
Onvoldoende inzicht in impact op duurzaamheid

in sommige gevallen toch mee te stoppen, omdat het niet werkt of dat er teveel risico's blijken te zijn bij een specifieke toepassing. Hoewel AI al langer bestaat en ook op verschillende manieren al door overheden wordt toegepast, bijvoorbeeld in de vorm van machine learning, is GenAI relatief nieuw. *“Ik denk eigenlijk dat de uitdagingen exact hetzelfde zijn, alleen dat*

de kennis over generatieve AI nog minder is. Ja, dat is zo'n nieuw veld, dat er zo weinig aanwezig is dat je diezelfde uitdagingen nog steeds heel moeilijk kan tackelen.” (respondent 1) . Onzekere uitkomsten zijn voor sommige organisaties een drempel om experimenteren met GenAI te omarmen.

31 Cinar et al., 2019

32 Jöhnk et al., 2021

Ontwikkelingen op het gebied van GenAI zijn moeilijk bij te houden

In een korte tijd hebben veel ontwikkelingen op het gebied van GenAI plaatsgevonden. Veel respondenten vinden het lastig om bij te houden wat de laatste technische ontwikkelingen zijn. Het vergt specialistische technische kennis en is en veranderlijk. Alle modellen hebben voor- en nadelen. En als je dan voor een ander model kiest voor je toepassing die je al in productie hebt genomen, moet je rekening houden met mogelijke andere output, ander soort fouten, omdat elk model (of versie van hetzelfde model) daarin anders kan zijn. *“Alle spullen die we nu hebben draaien, productioneel gezien, dat is gewoon oud, zeg maar. Dat is de vorige generatie AI. We zijn natuurlijk bezig met de nieuwe, moderne of volgende golf, maar we weten eigenlijk al, ja, over twee jaar lachen we daar om,”* aldus respondent 7.

Risico op fouten en misbruik van GenAI

Het gebruik van GenAI brengt ook risico's mee op misbruik en fouten. GenAI toepassingen kunnen bewust en onbewust beïnvloed of gemanipuleerd worden. Ook zijn de respondenten bewust van foute uitkomsten van GenAI, zoals hallucinaties. Als AI agents ingezet gaan worden, gaan die handelingen doen, acties uitvoeren of zelfs beslissingen maken. Dat zal een foutmarge hebben, die alsnog beter kan

zijn dan een medewerker. Toch blijft dan de vraag van respondent 5: *“Maar is dat goed genoeg voor iets dat automatisch gaat?”* Als iets 100% van de tijd moet kloppen, is volgens deze respondent een oplossing waarin altijd variatie bestaat (zoals met GenAI) eigenlijk niet geschikt. Toch kunnen er ook oplossingen ingebouwd worden, afhankelijk van de toepassing. Respondent 7 geeft aan dat een escalatie route is ingebouwd in hun chatbot, zodat er bij twee foute antwoorden contact met een medewerker wordt aangeboden aan de burger.

Data management en datakwaliteit is niet altijd op orde

De kwaliteit van de output van GenAI is afhankelijk van de kwaliteit van de data die erin gaat. Het is enorm belangrijk dat overheidsorganisaties hun data goed op orde hebben.

Beperkte AI geletterdheid in organisaties

Er is weinig kennis van en ervaring met (Gen)AI, waardoor innovatie met en gebruik van GenAI beperkt kan worden. In bijvoorbeeld beleidsdomeinen is vaak nog onvoldoende bewustzijn over de rol die GenAI kan spelen in het oplossen van maatschappelijke opgaven.

Beperkte centralisering op nationaal niveau

Er is behoefte aan sturing vanuit de centrale overheid door de nieuwigheid van GenAI en onzekerheden rondom wat er kan en mag met GenAI. Duidelijkheid in beleid, gezamenlijke standaarden en normen, en sturing rondom samenwerking en financiering zal naar verwachting bijdragen aan effectiever gebruik van GenAI en efficiëntere inzet van middelen en capaciteit.

Onduidelijkheid rondom wet- en regelgeving en impact toetsen

Er wordt hard gewerkt aan interpretatie van wetgeving zoals de EU AI Verordening en de samenhang met bijvoorbeeld de Algemene Verordening Gegevensbescherming. Ook impact toetsen zoals een DPIA en IAMA roepen vragen op en worden niet eenduidig ingezet. Dit geeft momenteel nog vertraging in het gebruik van GenAI maar brengt ook kansen met zich mee voor kennisontwikkeling en -deling waar andere partijen uit het Nederlandse ecosysteem van kunnen leren.

Proportionele aandacht voor ethische overwegingen

In de meeste organisaties zijn ethische afwegingen niet een vaststaand en expliciet onderdeel van het innovatieproces. Veel respondenten gaven aan dat

het niet werkbaar is om in een vroeg stadium van innovatie dieper in te gaan op ethische en compliance aspecten, omdat dit een sterke focus op risico's met zich meebrengt waardoor het moeilijk is om een toepassing van de grond te krijgen. Naar mate een toepassing verder ontwikkelt richting opschaling wordt er ook steeds meer gesproken over (het mitigeren van) risico's.

Daarnaast gaven een aantal respondenten aan dat er wel een ethisch kader beschikbaar was in de organisatie, maar hoe dit in de praktijk concreet een plek moet hebben in het innovatieproces is niet altijd duidelijk. *“We hebben een ethisch kader. Maar het mooie van een ethisch kader is dat iedereen zegt: we hebben een, maar niemand weet wat erin staat. Ik heb hem ooit doorgelezen. Je moet mij niet vragen wat erin staat.”*- Respondent 5

Benodigde technische infrastructuur

Ook rondom de technische infrastructuur bestaan uitdagingen. Er wordt nu vaak gewerkt met kleine initiatieven die lokaal op laptops draaien, maar de behoeftestelling is groter. Om GenAI toepassingen te kunnen laten werken is er voldoende rekenkracht nodig. Daarnaast spelen er ook vraagstukken rondom fysieke ruimte, stroom- en waterfaciliteiten.

In de interviews kwam de wens voor centralisatie rondom digitale vraagstukken vaak ter sprake. In de NDS³³ wordt besproken dat het streven is om naar één digitale overheid te werken. Er werden echter ook kanttekeningen bij gegeven in interviews. Bijvoorbeeld voor bestaande ondersteunende organisaties, zoals data of expertise centers, geldt vaak een beperkte opdracht of beperkt mandaat. Zij mogen zich alleen specifiek op het Rijk richten, en niet op andere type overheden. Bij het opschalen van technische infrastructuur kan dit voor belemmeringen zorgen.

De afwezigheid van budget, bronnen of andere middelen

Een veel voorkomend knelpunt is de beperkte beschikbaarheid van financiële middelen, tijd, data, capaciteit en rekenkracht. Respondent 3 geeft aan: *“Ik heb eigenlijk drie ontwikkellaptops nodig voor de prijs van €5000, maar dat is een investering en ik heb zelf als teamleider geen budget. Dan is het heel lastig om geld te regelen. Dus het is altijd een creatieve zoektocht naar goede spullen.”*

Daarnaast zijn bronnen voor GenAI ontwikkeling en inbedding een belangrijk knelpunt voor organisaties. Onderliggend aan effectieve GenAI is een goede

datahuishouding en een hoge kwantiteit en kwaliteit van data nodig. Respondent 7 geeft aan: *“Als we de chatbot op onze website zetten, moeten de webpagina’s allemaal up-to-date zijn, anders gaat de bot gekke dingen zeggen. Dus we moeten eerst aan die kant al een enorme kwaliteitsslag waarschijnlijk doen.”*

Maar tekorten aan middelen en beperkingen rondom infrastructuur spelen ook op de lange termijn. Respondent 14 geeft aan: *“Ik denk dat de grondstof die aan de basis ligt is natuurlijk gewoon de computing, de rekenkracht. Je moet voorstellen dat we met ons datacenter niet superveel ruimte meer hadden. Dus we hebben echt wel echt ons best moeten doen om te zorgen dat we voor komend jaar genoeg ruimte hebben. Maar dan [...] moet er ook waterkoeling komen om te zorgen dat we genoeg chips op een klein stukje kunnen stoppen, en nieuwe specialistische stroomaansluitingen.”*

Digitale afhankelijkheid

Een ander onderwerp dat de respondenten vaak aanhalen is digitale afhankelijkheid. Afwegingen rond het gebruik van open source modellen spelen een belangrijke rol in afwegingen rond technologie, maar respondenten geven ook aan dat

organisaties worstelen ook met het vinden van een passende oplossing. Veel organisaties kiezen op dit moment voor bekende leveranciers waar ze al contracten mee hebben, bijvoorbeeld omdat het binnen bestaande kantoorautomatisering past. Uit de interviews blijkt dat een aantal organisaties ook kijken naar (Europese) open source alternatieven, maar dat die op dit moment kwalitatief gezien nog onvoldoende zijn. Sommige respondenten geven aan dat ze proberen rekening te houden met de impact op digitale afhankelijkheid door middel van lokale hosting van data en/of model, en de mogelijkheid te behouden om nieuwe GenAI modellen in te zetten wanneer deze beschikbaar worden. Volledige oplossingen zijn echter nog niet voorhanden.

Respondent 10 geeft aan: *“In mijn beleving bij de Nederlandse overheid. Dat we de Microsoft fuik ingelopen zijn. Nu merk je dat we opnieuw op een soort kruispunt staan. Waar we verschillende opties hebben. En je ziet dat heel veel bestuurders het wel makkelijk vinden. Om een pot geld of een zak geld aan Microsoft te geven. En dan alle verantwoordelijkheden naar hun te schuiven.”*

Onvoldoende inzicht in impact op duurzaamheid

Het onderwerp duurzaamheid houdt respondenten ook bezig. Ze stellen zelf of krijgen vragen over water en energieverbruik van bijvoorbeeld datacentra. Dit hangt ook samen met het knelpunt over beschikbare infrastructuur. Volgens respondent 15, hoeft dit echter niet een automatische rem zijn om met GenAI te innoveren: *“Op het moment dat je GenAI een vraag stelt, even in chat variant, dat is ongeveer 2 tot 2,5 keer een reguliere Google vraag. Het tegenargument is dat de informatie die je terugkrijgt dermate efficiënt is dat je niet meer hoeft door te zoeken.”* Volgens deze respondent gaat het niet om een keuze is van wel of niet doen, maar dat het verder vraagt om na te denken over hoe de infrastructuur en andere zaken worden ingericht.

³³ De Nederlandse Digitaliseringsstrategie

5. Praktijkvoorbeelden

Chatbot Virtuele Gemeente Assistent GEM

Doel toepassing: GEM is een virtuele assistent op de gemeentelijke website. De chatbot geeft antwoord op vragen van inwoners. Het doel van de toepassing is om vragen van inwoners zo goed mogelijk te beantwoorden. De chatbot kan ook doorverwijzen naar een medewerker.

Type overheidsorganisatie: 25 deelnemende gemeenten. Utrecht, Tilburg, Rotterdam, Meierijstad, Roosendaal, Heusden, Alblasserdam, Dordrecht, Hendrik-Ido-Ambacht, Sliedrecht, Zwijndrecht (Dienstverlening Drechtsteden), Roosendaal, Velsen, Hillegom, Lisse, Teylingen (HLT Samen), Geldrop-Mierlo, Nuenen, Son en Breugel, Eindhoven, Gooise Meren, Amersfoort, Roermond, Barneveld en Leiden.^{34, 35, 36}

Status: De chatbot is geïmplementeerd.

Impact toets: DPIA

Gebruiker: Inwoners van deelnemende gemeenten zijn de beoogde gebruiker.

Ontwikkelvorm: GEM is een samenwerkingsverband dat momenteel bestaat uit 25 gemeenten. Het samenwerkingsverband is in 2019 opgericht door de gemeente Utrecht en gemeente Tilburg.³⁷ De VNG gaat in 2025, met subsidie van het ministerie van BZK, innovatieve initiatieven doorontwikkelen en opschalen. De virtuele assistent van GEM wordt samen met initiatieven MAI en de Friese Aanpak opgeschaald.³⁸ Het samenwerkingsverband heeft de onderliggende technologie niet zelf ontwikkeld, hiervoor is een marktpartij ingeschakeld. Het samenwerkingsverband richt zich op de kwaliteit van de vragen. Een belangrijke voorwaarde voor de inkoop van de technologie was dat het eigenaarschap van het product niet bij de marktpartij ligt, maar bij het samenwerkingsverband.

Technische kenmerken: Oorspronkelijk maakte GEM gebruik van NLP technieken, maar het is nu bezig met de overstap naar GenAI. GEM wordt open source ontwikkeld. De antwoorden die GEM geeft staan in een landelijke bibliotheek opgeslagen.

Succesfactoren

- **Kennisdeling.** GEM biedt een gezamenlijke en uniforme oplossing voor uniforme dienstverlening van gemeenten richting burgers. Hierdoor wordt het wiel niet opnieuw uitgevonden op meerdere plekken, maar wordt kennisdeling bevorderd, met het perspectief of overheidsbrede dienstverlening.
- **Beschikbaarheid van middelen.** Koploper gemeenten die de samenwerking zijn gestart hadden beschikking over **meer middelen om te investeren** in de opstart van de toepassing. Door deze investering konden kleinere gemeenten met minder middelen makkelijker aansluiten bij deze samenwerking. “We hebben als koplopergemeente gezegd, wij nemen die handschoen op [...] Met als doel dat de instap in het project, in het open source project, of het meedoen aan de community, zo laagdrempelig mogelijk is voor andere gemeenten. Omdat die misschien zeggen, ja wij zijn kleiner, we hebben minder financiële middelen. Dit is echt een uitnodiging om mee te doen aan

het samen ontwikkelen.” De huidige kosten worden verdeeld onder de deelnemende gemeenten, wat exploitatie kosten verlaagd.

- De **pilot begon eerst op kleine schaal** bij één gemeente en is **stapsgewijs uitgerold naar andere gemeenten**. Ook op inhoud is het aantal onderwerpen stapsgewijs uitgebreid.
- De samenwerking heeft **verschillende rollen geformaliseerd om de samenwerking te structureren**: een productowner, een conversatiedesigner en een community manager.
- Het **delen van middelen en lasten tussen gemeenten**. Zo is er bijvoorbeeld één DPIA opgesteld die door andere gemeenten gecontroleerd en gebruikt kan worden.

Knelpunten

- **Datakwaliteit en huishouding** van gemeentelijke websites zijn niet altijd op orde. Bij inzet van GenAI op basis van data van de website, is hoge kwaliteit en ontsluitbaarheid van data een vereiste.

³⁴ [Virtuele assistent GEM](#)

³⁵ [Gemeente start proef met chatbot](#)

³⁶ [Common Ground Virtuele assistent GEM](#)

³⁷ [Virtuele assistent GEM](#)

³⁸ [Opschaling virtuele assistent](#)

- **Bijhouden van technologische ontwikkelingen en rekening houden met de levensduur van technische componenten.** Hierover stelt de respondent de volgende vraag: *“Je weet dat er waarschijnlijk een aantal componenten zijn waar je over twee jaar weer afscheid van moet gaan nemen, omdat dat oude spullen zijn. Hoe zet je die spullen wel zo in elkaar, dat het dan wel kan blijven draaien en kan werken? [...] en hoe zorg je ervoor dat het niet een enorme legacy meebrengt?”*

Opschaling wordt als een grote uitdaging ervaren. Samenwerking en rolverdeling zijn essentieel. Knelpunten rondom opschaling zaten met name rondom de volgende factoren:

- **Complexiteit van gezamenlijke organisatie** en ICT structuren was onderschat.³⁹
- **Oorspronkelijke samenwerkingsstructuur was niet bestendig voor de groei van het samenwerkingsverband.** GEM begon klein, met drie gemeenten. Geleidelijk kwamen er steeds meer gemeenten bij.

- Tegelijkertijd bleef de constructie van de samenwerking hetzelfde. Dit werd niet goed werkbaar. *“GEM is nog steeds een projectorganisatie die vooral leunt op de contracten die de gemeente Tilburg heeft met een marktpartij voor wat betreft hosting en ontwikkeling. Maar dat is natuurlijk geen situatie die je voor lange termijn wil volhouden”* De afgelopen twee jaar is GEM druk geweest met onderzoeken wat er juridisch nodig is om een andere constructie te vinden voor deze samenwerking.
- **Inkopen van een open source project is complex.** Dit leidde tot de volgende vragen: Wat voor afspraken maken we dan met elkaar? Hoe ziet het IP eruit? Hoe ziet de rolverdeling eruit? Welke zekerheden kan je inbouwen en wat zijn de risico's?

Belangrijke overwegingen tijdens de ontwikkeling:

Twee belangrijke uitgangspunten bij de ontwikkeling van de toepassing zijn: de **vragen van de inwoners** staat centraal én **samenwerking tussen gemeenten**. De respondent geeft hierover aan: *“Maar GEM*

is niet de virtuele assistent, is niet generatieve AI, is niet die bot. GEM is voor mij echt de samenwerking van gemeenten op de vraag: We hebben allemaal vergelijkbare dienstverlening, hoe willen we dat inrichten richting de inwoner?”

Een ander uitgangspunt van GEM is om **één virtuele assistent te bouwen die voor alle Nederlandse gemeenten** te gebruiken is. Doordat **dienstverlening processen op veel plekken vergelijkbaar** zijn ontstond het idee om samen te gaan werken tussen gemeenten. De respondent legt uit: *“Al die gemeenten hebben een eigen website, er zitten allerlei mensen heel hard hun best voor te doen om dat allemaal lokaal bij te houden. Maar een groot deel van de informatie die erop staat is gewoon generiek. Of je nou een paspoort aanvraagt in Eindhoven of in Utrecht, wat je moet doen is op veel, veel punten hetzelfde.”* Een ander belangrijk uitgangspunt tijdens de ontwikkeling van GEM was dat de **Chatbot** een middel is, geen doel. Hierover zegt de respondent: *“Het feit dat we samenwerken is super krachtig. Daar gaat GEM voor mij over. Dat dat uiteindelijk een bot is die vragen van inwoners beantwoord, helemaal top. Maar dat is uiteindelijk de uitkomst van de samenwerking.”*

Er is ook bewust gekozen voor **open-source ontwikkeling**. Hierover zegt de respondent: *“We geven publiek geld uit, dus we vinden het heel eerlijk dat dat ook publiek beschikbaar komt. En dat geldt zowel voor de techniek als wel voor de content.”*

Tegelijkertijd was het samenwerkingsverband tijdens de ontwikkeling ook bewust bezig met het scheppen van **randvoorwaarden en rolverdeling binnen de samenwerking**. Hierover zegt de respondent: *“Je begint aan een huwelijk, om het zo maar even te zeggen. Ik denk dat het wel slim om vooraf, hoe onromantisch ook, over de voorwaarden te hebben nagedacht.”*

Er is aandacht voor de **rol van de overheid bij innovaties binnen de overheid**: *“Wij willen niet als overheid een software ontwikkelaar zijn. Het is niet onze rol, maar als open source project ga je wel ergens in ontwikkeling meedoen. Maar we wilden wel heel scherp zijn en eigenlijk de markt ook uitdagen en bevragen wat kunnen jullie hierop voor ons leveren, zonder dat het jullie ‘product’ is.”*

Ook is er nagedacht over compliance met wet- en regelgeving. Er is bijvoorbeeld **eén DPIA voor de gehele samenwerking** opgesteld: *“We hebben één oplossing, dus*

³⁹ [Virtuele assistent GEM](#)

we hebben er maar één DPIA over hoeven maken, en niet 25 keer.”

Aansluiting bij context van de gemeente.

Medewerkers van een specifieke gemeente kunnen ook specifieke vragen en antwoorden voor hun gemeenten toevoegen zodat de antwoorden goed aansluiten bij de situatie van een gemeente. GEM geeft antwoord met gescripte antwoorden. De antwoorden die het geeft zijn door medewerkers van gemeenten gemaakt. Deze medewerkers komen één dag in de week samen om aan de inhoud van de landelijke bibliotheek te werken.

Een belangrijke voorwaarde tijdens de ontwikkeling is dat de chatbot niet mag hallucineren en dat de output klopt.⁴⁰

Momenteel is GEM bezig met de overstap naar het gebruik van GenAI. Bij deze overstap komen een aantal uitdagingen op dit vlak naar voren. Er wordt daarom ook nog onderzocht hoe GenAI het beste kan worden ingezet en welke **guardrails** er nodig zijn. Eén optie is door GenAI gescripte antwoorden te laten analyseren. De andere optie is om GenAI de gemeentelijke website te laten scrapen, maar hierbij is het extra belangrijk dat de datakwaliteit van de website op orde is.

Een andere belangrijke overweging tijdens de ontwikkeling is de **mogelijkheid tot contact met een mens**. De chatbot moet kunnen doorverwijzen naar een medewerker van de gemeente.⁴¹

Er wordt ook actief aandacht besteed aan het **monitoren en evalueren** van de prestaties van GEM.

⁴⁰ [Trial en error met gemeentelijke chatbots](#)

⁴¹ [Prestaties virtuele assistent GEM in 2024](#)

Leeshulp

Doel toepassing: de leeshulp van Tolkie heeft als doel om informatie begrijpelijk te maken voor laaggeletterden. De leeshulp helpt door moeilijke woorden uit te leggen, teksten voor te lezen, te vertalen en samen te vatten. Vertalingen zijn beschikbaar voor 31 talen.

Type overheidsorganisatie: Meerdere gemeenten maken gebruik van Tolkie, onder andere de gemeente Vlaardingen, gemeente Meijerijstad, gemeente Boxtel, gemeente Wijchen, gemeente Druten, gemeente Woensdrecht en gemeente Rijssen-Holten. Naast overheidsorganisaties maken ook woningcorporaties gebruik van deze toepassing.

Status: De leeshulp is **geïmplementeerd**.

Impact toets: Er zijn geen impact toetsen door Tolkie zelf gedaan, maar ze werken mee aan een DPIA als daar om wordt gevraagd.

Gebruiker: De beoogde gebruiker is de burger, met name laaggeletterden.

Ontwikkelvorm: De Leeshulp is ontwikkeld door Tolkie, een sociale onderneming. Het bedrijf AI Heroes heeft de ontwikkeling ondersteund en advies geleverd.

Technische kenmerken: De toepassing maakt gebruik van GPT. Tolkie maakt gebruik van één stukje code op de website. Het stukje code haalt de zichtbare tekst op en stuurt dit naar een veilige plek (Microsoft Azure in de EU) en krijgt de gemakkelijke tekst terug. Deze tekst verschijnt in een apart schermje op de website. Er worden geen IP-adressen opgeslagen. De gebruikte AI is van Microsoft Azure. De data staat op Europese servers opgeslagen.

Omdat veel GenAI toepassingen niet direct door overheidsorganisaties worden ontwikkeld hebben we bij dit voorbeeld het perspectief van een externe ontwikkelaar genomen. De succesfactoren en knelpunten die hieronder zijn beschreven zijn dan ook het perspectief van een startup.

Succesfactoren

- De **maatschappelijke opgave staat centraal**, niet de technologie. De maatschappelijke opgave is: teksten begrijpelijk maken voor laaggeletterden.
- De **beoogde gebruiker, laaggeletterden, zijn actief betrokken** bij de ontwikkeling van deze toepassing. Doordat vanaf het begin de wensen en behoeften van laaggeletterden zijn meegenomen is er

een toepassing ontwikkeld dat als waardevol wordt beschouwd door de beoogde gebruiker.

- **Kwaliteit van de output** staat centraal. Gebruikers kunnen direct via de toepassing hun **feedback** geven op de toepassing.
- **Verwachtingsmanagement** vanuit overheidsorganisaties helpt bij het scheppen van duidelijkheid in het inkoopproces.
- Ontwikkeling toepassing begon voor introductie GenAI. Kwaliteit van taal was een belangrijke voorwaarde voor het succes van de toepassing. **Verbetering in kwaliteit van GenAI modellen in Nederlandse taal** zorgde ervoor dat de toepassing schaalbaar werd en geïmplementeerd kon worden op websites van overheidsorganisaties.

Knelpunten

- In **contact komen** met de juiste persoon binnen een **overheidsorganisatie** wordt soms als uitdagend ervaren.
- **Inkoopprocessen** verschillen per gemeente. De respondent zegt hierover: *“Bij sommige gemeenten krijg je een waslijst aan vragen, en soms moet ik tegen gemeenten zeggen, laten we wel ook hiermee rekening houden. [...] ik had niet*

verwacht dat het zo erg zou verschillen per organisatie.”

- **AI-beleid van organisaties** is niet altijd scherp. De respondent zegt hierover: *“Maar soms is bij een gemeente het interne AI-beleid niet helemaal scherp. Dat maakt het voor een ambtenaar lastig om iets aan te schaffen.”*

Belangrijke overwegingen tijdens de ontwikkeling:

Vanaf het begin van de ontwikkeling stond de **maatschappelijke opgave centraal**, niet de technologie. Het idee om AI te gebruiken bij deze toepassing ontstond pas nadat er een concreet doel was geformuleerd.

Het doel van deze toepassing ontstond in samenwerking met een groep laaggeletterden. De ontwikkeling van de toepassing stond dan ook dicht bij de doelgroep. Doordat vanaf het begin de wensen en behoeften van laaggeletterden zijn meegenomen is er een toepassing ontwikkeld dat als waardevol wordt beschouwd door de beoogde doelgroep.

Het doel van de toepassing is om teksten toegankelijk te maken voor laaggeletterden. Hierdoor was tijdens de ontwikkeling van de toepassing een belangrijke voorwaarde de **kwaliteit van de output**. Toen

Tolkie werd opgericht in 2022 was GenAI nog erg nieuw, en waren GenAI modellen nog niet goed genoeg in de Nederlandse taal. Pas toen de kwaliteit van de modellen in Nederlands voldoende was is er besloten om over te stappen op GenAI. Dit maakt de ontwikkeling van de toepassing gemakkelijker te schalen:

“Toen we begonnen, toen ging alles nog over NLP [...] Aan het begin hebben studenten zelf definities van moeilijke woorden geschreven [...] Toen wij startten was bijvoorbeeld GPT-3 er al, maar dat was in het Nederlands niet goed genoeg voor deze toepassing. De kwaliteit is in een hele korte tijd heel erg verbeterd in het Nederlands. En dat maakt [de toepassing] veel schaalbaarder. Op een gegeven moment hadden wij 70.000 definities in onze database staan.”

Data opslag en digitale soevereiniteit.

Er is bewust voor gekozen om de data op Europese servers op te slaan.

Doordat de kwaliteit van de output als een belangrijke voorwaarde wordt beschouwd, werd tijdens de ontwikkeling veel aandacht besteed aan het verzamelen van **feedback van gebruikers**.⁴² Ook nu de toepassing geïmplementeerd is blijft dit een belangrijk

element. Bij het ontwerp van de toepassing is dan ook een functie toegevoegd waarbij gebruikers direct feedback kunnen geven.

Nadat de Leeshulp beschikbaar kwam op de markt, werd verwachtingsmanagement

vanuit gemeenten als waardevol ervaren. De respondent geeft hierover aan: *“Verwachtingsmanagement is heel fijn. Ik vind het altijd heel fijn als een gemeente vertelt, wie gaan we intern aanhaken, wat kun je nog verwachten, hoe ziet het proces eruit?”*

Naast een omschrijving voor het algoritme-register heeft Tolkie ook een omschrijving gemaakt van hun product in begrijpelijke taal dat organisaties op hun website kunnen gebruiken.[#] Ook hebben ze video's gemaakt waarin wordt uitgelegd hoe de Leeshulp van Tolkie werkt.^{43, 44}

⁴² [Technologie voor laaggeletterden](#)

⁴³ [Demo Tolkie Leeshulp \[video\]](#)

⁴⁴ [Tolkie Leeshulp](#)

Voorbeeld Archiefassistent

Doel toepassing: Het Nationaal Archief (NA) is de grootste archiefinstelling van Nederland. Als onderdeel van hun taken, onderhoudt NA de opslag en toegang tot overheidsinformatie.^{45, 46, 47}. Ook behandelt NA klantvragen over uiteenlopende historische onderwerpen. Om medewerkers in deze taak te ondersteunen, wordt er bij NA gewerkt aan een archiefassistent.

Status: De archiefassistent is nog in ontwikkeling.

Impact toets: Plan om IAMA en (pre)DPIA te doen.

Gebruiker: De archiefassistent zal alleen voor medewerkers beschikbaar zijn, om te voorkomen dat een burger bijvoorbeeld onbedoeld racistische termen of misinformatie tegenkomt.

Ontwikkelvorm: Het initiatief is gestart bij de Koninklijke Bibliotheek, en is nu onder beheer van het Nationaal Archief. De ontwikkeling vindt sinds het beginstadium in samenwerking met de leverancier TextPast plaats. Het taalmodel dat wordt gebruikt is van Mistral AI.

Technische kenmerken: Er wordt nu gebruik gemaakt van de open source versie van het model van Mistral, maar er wordt verder gekeken naar andere taalmodellen, waaronder GPT-NL.

Er worden digitale structuren (kennisgrafen (NL) of Knowledge Graphs (EN)) toegepast zodat de toepassing informatie kan vinden en combineren. Hiermee kan de archiefassistent rijke data gebruiken om antwoorden op te baseren. De gebruikte informatie voor de kennisgrafen komt uit verschillende bronnen, o.a. ontsloten archiefdata van partners zoals het Huygens-ING Instituut.

Er wordt ook gebruik gemaakt van rule-based AI en machine learning, omdat deze hybride aanpak betere resultaten levert dan puur zelflerende systemen.

De archiefassistent is benaderbaar via een API, wat flexibiliteit behoudt voor verdere technische ontwikkeling van de toepassing, zoals werken in andere programmeertalen of verminderde complexiteit bij het integreren van systemen.

Succesfactoren

- Voor het ontwikkelen van goede toepassing waarbij tijdig kansen en risico's worden gesignaleerd is het belangrijk om **relevante domein en expert kennis, ervaring en vaardigheden te betrekken**. In dit geval op het gebied van archiveren, (digitale) geesteswetenschappen, computer science, NLP, en software/data engineering (Rust).
- Er is **nauw contact met de gebruikersgroep** om te testen of de toepassing de juiste functionaliteiten heeft en gebruiksvriendelijk is. Deze aansluiting met de eindgebruiker zorgt ervoor dat de toepassing relevant en bruikbaar is en blijft.
- De ontwikkeling verloopt voor-**spoedig** omdat er ruim **tijd vrijgemaakt is om aan de toepassing te werken**.
- Door **samenwerkingsverbanden met externe partners** is er toegang tot relevante kennis en (technische) middelen. In deze casus fungeren instituten als Huygens-ING, Netwerk Digitaal Erfgoed en andere archieven als

sparringpartner op het gebied van Linked Data. Deze data voorzien de modellen van context. Daarnaast zijn onderzoeksprogramma's als HAICu cruciaal voor het verkennen van nieuwe mogelijkheden tot het creëren van relevante context voor AI-systemen binnen de erfgoedwereld.

Knelpunten

- Er is **beperkte beschikbaarheid technische rekenkracht** (video-kaarten), waardoor er vertraging ontstaat en minder mogelijkheden beschikbaar zijn om het project volledig in-house aan te pakken.
- **De aard en kenmerken van de data** (o.a., historisch, racistische teksten, persoonsgegevens) maakt het project gevoelig waardoor vooraf goed nagedacht moet worden over risico's en richtlijnen of technische maatregelen die verantwoord gebruik van AI bewerkstelligen (guardrails).
- GenAI is nog relatief nieuw, en er zijn voor deze toepassingsvorm **geen kant-en-klare toepassingen of gevestigde leveranciers**, waardoor het project en team te

⁴⁵ [Wet- en regelgeving Nationaal Archief](#)

⁴⁶ [Regels voor archieven overheid](#)

⁴⁷ [Nieuwe archiefwet aangenomen door Tweede Kamer](#)

maken hebben met “trial and error.”

- Er zijn **geen of maar een beperkt aantal hoogwaardige Europese of Nederlandse taalmodellen** beschikbaar wat impact heeft op de kwaliteit van de output van de toepassing en soevereiniteit.
- **Politieke ontwikkelingen**, zoals de val van kabinet Schoof I, zorgen voor een verschuiving in welke onderwerpen aandacht krijgen. Dit heeft indirect maar wel duidelijke impact op innovatieprocessen zoals GenAI ontwikkeling en implementatie.

Belangrijke overwegingen tijdens de ontwikkeling:

Een van de afwegingen had te maken met de ontwikkeling van de toepassing. De keuze om de **toepassing grotendeels zelf te ontwikkelen** is gemaakt vanwege **security en compliance** overwegingen. Zowel het taalmodel als enkele andere technische componenten zijn open source, omdat er meer transparantie en mogelijkheid is om zelf controle te houden dan bij ingekochte software of componenten. **Architectuurkeuzes** spelen een grote

rol vanaf de start van de ontwikkeling. Volledige integratie kan GenAI effectiever maken, maar het geeft ook extra complexiteit tijdens de ontwikkeling en voor onderhoud. Voor de archiefassistent is gekozen om een op zichzelf staand systeem te bouwen, omdat de huidige infrastructuur niet geschikt is om met moderne GenAI toepassingen te werken. Er wordt daarom ook gebruik gemaakt van een API om gevoelige data op te vragen. Een andere factor bij deze overweging is om herbruikbaarheid van de archiefassistent door andere archief instituten mogelijk te houden.

Tijdens de ontwikkeling van de archiefassistent werd regelmatig afgewogen welke programmeertaal het beste past. Doordat de archiefassistent **losstaat van de hoofdinfrastructuur** is het ook mogelijk om te wisselen van programmeertaal. Daarmee wordt het gemakkelijker om ervoor te kiezen naar een meer geschikte taal over te stappen, ondanks dat dit ook tijd kan kosten.

Vanaf het begin van de ontwikkeling is besloten om voor te bereiden op **multimodale input** van de toepassing. Dit betekent dat de toepassing voorbereid wordt op het verwerken en begrijpen van verschillende

vormen input (e.g., tekst, geluid, video). De optie zal in eerste versies waarschijnlijk nog niet actief zijn, wegens wetgeving op beeld en geluid. Later inbouwen van multimodaliteit was niet wenselijk omdat dit de basis van de architectuur beïnvloed.

Verder is **data** een groot onderwerp waar goed over nagedacht moet worden. Het Nationaal Archief heeft een schat aan data beschikbaar, wat waardevol is bij de ontwikkeling van GenAI. Maar niet alle data is direct geschikt. **Sensitiviteit** is belangrijk, want historische data bevatten verouderde denkbeelden, racistische teksten of misinformatie, die niet zomaar in de output terecht mogen komen. Met taalmodellen vraagt die controle extra zorg, bijvoorbeeld op het vlak technieken om de accuraatheid van modellen te vergroten, zoals postprocessen.

Daarnaast bestaan er rondom het gebruik van data ook **juridische vraagstukken**. Bijvoorbeeld over privacy of de vraag of data van de afgelopen 10 jaar gebruikt mag worden om vraag-en-antwoordenmodellen te trainen. Over het algemeen zijn juridische kaders bepalend volgens respondent 13: *“Wat wel en niet mag volgens de wet is altijd ‘first concern.’* Onduidelijkheid in de geldende wetten kan echter een

reden zijn om af te zien van bepaalde innovaties of innovatie richtingen.

Er is ook een afweging gemaakt over hoe de beste aansluiting kan worden gevonden **tussen technologie en toepassingsdomein**. Door het samenbrengen van relevante specialismes, zoals wiskundigen, historici en archiefwetenschappers kan er een applicatie gebouwd worden die goed aansluit bij het toepassingsdomein en kunnen complexe vragen worden opgemerkt en beantwoord. Bijvoorbeeld: wat zijn de gevolgen van historische data op een taalmodel? Hoe kan privacygevoelige informatie beschermd worden? Hoe worden denkbeelden uit het verleden omschreven in de output? Hoe kan archiefdata het beste ontsloten worden? Tijdig opmerken van dergelijke vraagstukken zorgt voor een gedegen opzet van minder risico's van de applicatie. Het HAICu programma⁴⁸ is gericht op het onderzoeken van deze vragen.

48 HAICu project

Voorbeeld WerkSaam Westfriesland

Doel toepassing: WerkSaam Westfriesland is een gemeenschappelijke regeling bestaande uit zeven gemeenten en helpt mensen met een afstand tot de arbeidsmarkt aan een baan. In opdracht van dit samenwerkingsverband is een toepassing ontwikkeld die in het proces om de arbeidsmarkt te herbetreden als digitale assistent voor de re-integratie coaches en zorgt voor minder (administratieve) werkbelasting voor coaches. De toepassing kan gesprekken opnemen, transcriberen en samenvatten.⁴⁹ Cliënten worden geïnformeerd over het gebruik en gevraagd om toestemming.

Status: De toepassing is in vergevorderde staat van ontwikkeling en er wordt in de praktijk met de toepassing geëxperimenteerd door een select aantal arbeidscoaches die fungeren als testers.

Impact toetsen: Een DPIA is gedaan en IAMA staat nog op de planning.

Gebruiker: Medewerkers, namelijk arbeids-/re-integratie coaches

Ontwikkelvorm: De ontwikkeling is een publiek-private samenwerking. De technische ontwikkeling wordt door een AI startup en Jobport gedaan, maar ze werken daarin nauw samen met WerkSaam WF. Het project wordt gesteund door Werk aan Uitvoering⁵⁰ en de VNG.⁵¹

Technische kenmerken: De toepassing wordt gebouwd op basis van GPT 4o van OpenAI en de integratie gebeurt met Azure omgeving van Microsoft. De hosting wordt momenteel gefaciliteerd door Jobport. De toekomstige inrichting van hosting en beheer tijdens operationele fase wordt nog bekeken.

Succesfactoren

- De toepassing is ontwikkeld vanuit een **vraaggestuurd probleem, in plaats van technologie-gestuurde oplossing**. Dit zorgt voor een duidelijke scope en nauwe aansluiting bij de praktijk.
- **Externe kennis en expertise inhuren** in plaats van alles zelf uit willen zoeken om kosten te besparen. Juist als er ingezet wordt op het leren van deze ervaring, kan dit in de toekomst

veel tijd en kosten schelen.

- In ontwikkelproces was er nauwe **betrokkenheid van de beoogde eindgebruiker** en afgevaardigden van de hele organisatie, zoals kwaliteitsmedewerkers en beleidsmedewerkers. Dit zorgde voor zowel draagvlak als kwaliteit van het project.
- Aansluiten bij **centrale pilots en subsidiekansen** heeft gezorgd dat er genoeg middelen waren voor de uitvoer en doorontwikkeling van de toepassing.
- **Klankbordgroep met (grote) gemeenten** die de ontwikkeling actief volgen en van **feedback** voorzien. Ook helpen met vragen rondom **compliance**, AVG en AI Act.

Knelpunten

- **Feitelijkheid en volledigheid** zijn essentieel, maar hallucinaties bij GenAI zijn een groot risico. Het koste veel tijd en moeite om de juiste instellingen en prompts te formuleren zodat de output correct is.
- Er is tijd verloren omdat de **kennis en expertise op het gebied van**

compliance en AI-gerelateerde wet- en regelgeving nog niet breed ontwikkeld is. Er zijn ook nog weinig andere voorbeelden waarvan geleerd kan worden.

- Er bestaan vergelijkbare transcriptie- en samenvattingstoepassingen binnen het zorgdomein, maar de **bestaande toepassingen zijn niet direct inzetbaar** in het sociale domein.

Belangrijke overwegingen tijdens de ontwikkeling:

Compliance stond centraal in de ontwikkeling van de toepassing. Intern was nog niet veel ervaring opgebouwd rondom AI, dus is er een extern bureau ingeschakeld om gezamenlijk een DPIA uit te voeren. Dit heeft veel tijd gekost maar leverde ook veel kennis op voor de organisatie. Daarnaast is de DPIA een deelbaar resultaat met de bredere gemeenschap en draagt het bij aan gezamenlijk leren.

Ook **veiligheid** van de toepassing en data vormen lastige vraagstukken. Nadat de proof of concept versie getest was en er met echte gesprekken gewerkt kon worden, is de toepassing overgezet naar de

⁴⁹ [De AI-assistent voor de re-integratiecoach](#)

⁵⁰ [Pilots met generatieve AI leveren inzichten op voor alle gemeenten](#)

⁵¹ [Pilot Efficiënte innovatie: AI als assistent voor de re-integratie coach](#)

beveiligde Microsoft Azure omgeving van Jobport, zodat er voldaan kon worden aan de BIO norm en criteria voor certificatie. Hoe de toepassing uiteindelijk wordt gehost en geïmplementeerd is nog niet bepaald.

Een zorgpunt voor ontwikkeling was de **kwaliteit, volledigheid en correctheid van output**. Het juist opzetten van prompts vroeg om een combinatie van kennis van het sociale domein en het opstellen van goede prompts. Dit toont het belang van ervaring met en kennis van het toepassingsdomein voor AI ontwikkelaars.

Het team heeft direct ingezet op **draagvlak** vanuit de eindgebruiker en rekening gehouden met mogelijke **weerstand**. Vanaf het begin van het project bestond de overtuiging dat de toepassing niet van bovenaf opgelegd moest worden en dat integratie met werkprocessen goed moet worden ingericht. Beoogde gebruikers werden direct betrokken in het proces. De belangen van cliënten worden behartigd door de cliëntenraad, en de bredere cliëntengroep wordt geïnformeerd en betrokken in experiment.

Er wordt gekeken of de toepassing uitgebreid kan worden naar andere gemeenten, maar dit roept vragen op over de rol van de gemeente en beheer van de toepassing. **“Je wil geen leverancier zijn als**

gemeente,” aldus respondent 16, want dat vraagt tijd, moeite en middelen voor het opzetten van een helpdesk, en inrichting van applicatiebeheer. Dit kan beter extern belegd worden.

In het proces is er aandacht voor **AI geletterdheid en bewustzijn voor de gebruikers**. De arbeidscoaches blijven namelijk altijd verantwoordelijk voor de output, en moeten hun kennis en expertise blijven inzetten naast de AI toepassing.

WerkSaam heeft op de vlakken waar ze zelf de kennis en expertise niet in huis hebben, besloten om **externe partijen** in te schakelen. Op technisch vlak was dit besluit snel genomen, maar voor compliance duurde het langer. Uiteindelijk is het belangrijk dat er niet te lang wordt ingezet op het zelf doen, maar juist op hulp inschakelen. Het kan duurder zijn om het zelf te blijven proberen dan externe expertise in te huren.

Het team richtte zich op proberen en leren door te doen. Door agile te werken en de **ontwikkeling in kleinere stappen** in te richten zorgde voor overzichtelijkheid en de mogelijkheid om in te grijpen waar nodig. Regelmatig feedback vragen aan de doelgroep en klankbordgroep zorgde voor tijdige sturing.

6. Conclusies en aanbevelingen

In dit deel van het rapport formuleren we de conclusies die we op basis van dit onderzoek kunnen stellen en geven we aanbevelingen voor de praktijk en toekomstige monitors. Het doel van dit onderzoek was drieledig:

1. Inzicht krijgen in de inzet van GenAI bij overheidsorganisaties;
2. Inzicht krijgen in de succesfactoren en knelpunten bij de inzet van GenAI toepassingen die Nederlandse overheidsorganisaties ervaren;
3. Inzicht krijgen in de overwegingen die worden gemaakt bij de inzet van GenAI toepassingen door Nederlandse overheidsorganisaties.

Door middel van desk research, survey onderzoek en semigestructureerde interviews hebben we informatie opgehaald om deze drie doelen te bereiken. De conclusies en aanbevelingen die uit dit onderzoek komen, zijn als volgt gestructureerd. Eerst gaan we in op de conclusies die volgen uit het onderzoek. Vervolgens besteden we aandacht aan aanbevelingen voor overheidsorganisaties die door willen met GenAI inzet of juist willen starten met GenAI. We sluiten af met aanbevelingen voor toekomstige onderzoeken die GenAI inzet en ontwikkelingen bij Nederlandse overheidsorganisaties monitoren.

Conclusies over de inzet van GenAI

Stijging gebruik van GenAI toepassingen

In vergelijking met TNO onderzoek naar het gebruik van AI bij overheidsorganisaties,⁵² is het aantal gevonden GenAI toepassingen gestegen van 8 GenAI toepassingen in juni 2024 naar 81 GenAI toepassingen in juni 2025. Dit geeft aan dat in een jaar tijd een sterke groei heeft plaatsgevonden voor deze soort technologie.

Gebruik van GenAI toepassingen door Nederlandse overheidsorganisaties is wijdverspreid

GenAI toepassingen worden wijdverspreid opgepakt, van kleine gemeente tot provincie en ministerie. Ook valt op dat er veel soortgelijke toepassingen, zoals chatbots op verschillende plekken worden ontwikkeld en geïmplementeerd. Een mogelijke verklaring hiervoor is dat generatieve AI momenteel relatief snel en laagdrempelig kan worden ingezet in processen zoals klantcontact (chatbots, (ver)taalhulp) en als assistent voor ambtenaren (Copilot, transcribeer- en samenvattingshulpmidde-len).

Daarnaast valt op dat bijna de helft van de toepassingen door gemeenten worden

ingezet. Het betreft voornamelijk chatbots die assisteren in het vergaren en verwerken van informatie voor ambtenaren en die assisteren in het contact met de burger.

GenAI wordt in een beperkter aantal soort processen ingezet dan machine learning toepassingen

In 2024 voerde TNO onderzoek uit naar het gebruik van AI bij overheidsorganisaties. In dit onderzoek werd breder gekeken naar inzet van AI zoals bijvoorbeeld machine learning technieken. Dit onderzoek richt zich daarentegen specifiek op toepassingen van GenAI. In vergelijking met het TNO onderzoek van 2024⁵³ valt op dat het merendeel van de geïdentificeerde GenAI toepassingen wordt ingezet in overheidsprocessen waarbij een hoge mate van gegevens en tekstverwerking plaatsvindt, zoals samenvatten of documenten opstellen. Machine Learning technieken werden in 2024 in een breder aantal processen ingezet. We hebben voor GenAI geen toepassingen gevonden in de fysieke omgeving, deze categorie kwam wel voor in het onderzoek van 2024.

Wel valt één ander aspect op: GenAI toepassingen kunnen gemakkelijker voor meerdere doeleinden en meerdere type taken beschikbaar worden gesteld.

Toepassingen zijn vaak al geïmplementeerd

Waar in 2024 veel GenAI experimenten vermoedelijk nog in de kinderschoenen stonden, observeren we in deze studie dat er al veel toepassingen in de experiment fase (29) bevinden of al geïmplementeerd zijn (37) binnen organisaties.

Eindgebruiker is veelal medewerker

Bij meer dan 75% van de toepassingen zijn medewerkers van overheidsorganisaties de directe eindgebruiker van de GenAI toepassing. De overige toepassingen zijn gericht op de burger als eindgebruiker.

Risico inschatting is niet altijd consistent

Van iets meer dan de helft van de toepassingen is er geen informatie gemeld over de risicocategorie waar de toepassing onder valt. Van de toepassingen waar de risico inschatting wel bekend is, zijn de meeste toepassingen (35%) ingedeeld in de categorie overige algoritmes. Ook observeren we dat dezelfde soort toepassing bij verschillende organisaties in een andere risico categorie wordt ingedeeld. Dit kan erop wijzen dat de methode en interpretatie van risico inschattingen tussen overheidsorganisaties mogelijk inconsistent is.

52 TNO, 2024

53 TNO, 2025

Van meerderheid toepassingen is het onbekend of er een vorm van impact toets is gebruikt

Van 43% van de toepassingen is bekend dat er een impact toets voor is gedaan of gaande is. Vormen van impact toetsen die door veel organisaties worden genoemd zijn de DPIA, DPIA light, pre-DPIA en IAMA. Sommige organisaties hebben ook een eigen impact toets ontwikkeld. Van 57% van de toepassingen is echter niet bekend of er een impact toets is gedaan of nog op de planning staat. Het is niet bekend wat hiervan de reden is. Uit de interviews is wel gebleken dat impact toetsen veel tijd kosten en daarom vaak pas in een latere fase van ontwikkeling worden uitgevoerd. Ook hebben organisaties niet altijd de nodige expertise of ervaring in huis om een impact toets goed uit te voeren. Tot slot is er een grote variatie te zien in de soorten impact toetsen die worden gebruikt. Om beter zicht te krijgen op de knelpunten bij het uitvoeren en rapporteren van risicoschattingen en het werken met impact toetsen is meer onderzoek nodig.

Toepassingen van GenAI worden niet volledig zelf ontwikkeld

Regel-gebaseerde algoritmen en machine learning algoritmen kunnen volledig in-house ontwikkeld worden doordat de gebruikte modellen en componenten zelf door een data scientist kunnen worden ontwikkeld. Dit is niet het geval bij huidige GenAI toepassingen omdat de onderliggende componenten en modellen gebaseerd zijn op LLMs en foundation models van bedrijven zoals OpenAI, Meta, Mistral, Microsoft of Google. Tegelijkertijd valt op dat informatie over technische specificaties en de onderliggende digitale infrastructuur niet altijd publiek beschikbaar of volledig zijn. Dit maakt het lastig om volledig inzicht te krijgen in de componenten waarop een GenAI toepassing wordt gebouwd.

Beperkte aandacht voor impact van GenAI op de werkvloer

GenAI wordt momenteel nog voornamelijk ingezet ter ondersteuning aan bestaande processen. Uit de interviews blijkt dat de focus bij organisaties vaak ligt op het vergroten van efficiëntie en verlichten van werkdruk, maar nog weinig op hoe de inzet van GenAI de aard van werk kan veranderen. Een respondent gaf wel een concreet voorbeeld hoe diens organisatie de

veranderende rol van het klantcontact centrum zien met de komst van GenAI. Hoewel dit onderzoek zich niet richt op de impact van het gebruik van GenAI op het type werk en werkplezier van medewerkers, lijkt dit onderwerp nog weinig aandacht te krijgen. Eerder onderzoek heeft aangetoond dat GenAI impact heeft op werkprocessen en werkplezier^{54,55} waardoor dit mogelijk wel een aandachtspunt is voor overheidsorganisaties die met GenAI aan de slag gaan.

Overheidsorganisaties denken bewust na over de herbruikbaarheid van GenAI toepassingen

Een veelvoorkomende opmerking is dat organisaties niet opnieuw het wiel willen (hoeven) uitvinden. Ook zijn veel organisaties zich ervan bewust dat innovatie met publiek geld wordt bekostigd en willen voorkomen dat investering in meerdere gelijkwaardige initiatieven worden gedaan. Daarom wordt er ingezet op samenwerking, informatie delen of gezamenlijk inkopen. Een andere manier om meerwaarde voor de publieke sector te ontwikkelen is bij de ontwikkeling meteen in te zetten op herbruikbaarheid van de toepassing. Het open source en beschikbaar maken van een toepassing is dan een vaak genoemde oplossing. Andere overheidsorganisaties

kunnen de toepassing dan vrij gebruiken mits ze de technische kennis en andere voorwaarden hebben om de toepassing uitpakken. Dit vraagt aan de voorkant van het ontwikkelproces wel enige tijd en aandacht voor hoe dit beste kan worden opgezet.

GenAI: middel of doel?

We observeren dat overheidsorganisaties door de aandacht voor GenAI een zekere druk kunnen ervaren om de technologie snel toe te passen. Een veel gehoorde reden om aan de slag met GenAI te gaan is dan ook om ervaring op te doen met de technologie. Daarnaast zijn er hoge verwachtingen over de efficiëntie en kostenbesparingen die het inzetten van GenAI kan opleveren.

Innoveren met GenAI is echter geen doel op zichzelf. Het risico is dat door deze hoge verwachtingen en druk een technologie *push* ontstaat. Dit kan risico's met zich meebrengen, omdat zonder concreet doel en visie op hoe de toepassing van GenAI kan bijdragen aan maatschappelijke opgaven, dienstverlening of concrete werkprocessen, de kans bestaat dat de inzet van deze technologie niet goed en op een verantwoorde manier van de grond komt.

54 GenAI verandert werk. Maar wat doet het met de mens?

55 TNO/Rathenau Insituut, 2024

Steeds meer alternatieven voor openbare chatbots

Net zoals bij de Quickscan AI in de Publieke Dienstverlening III⁵⁶ is het individueel gebruik van GenAI van ambtenaren buiten scope van dit onderzoek. Dit onderzoek focust zich op toepassingen die top-down worden aangeboden door overheidsorganisaties aan hun medewerkers. Wel valt op dat steeds meer overheidsorganisaties een alternatief aanbieden voor openbare chatbots door een organisatie brede chatbot beschikbaar te stellen voor medewerkers met als doel om schaduwgebruik onder medewerkers te voorkomen.

Aanbevelingen voor overheidsorganisaties

Op basis van de onderzoeksresultaten en conclusies formuleren we een aantal aanbevelingen voor overheidsorganisaties die door willen gaan met de inzet van GenAI iof juist willen starten met GenAI innovaties. De aanbevelingen zijn net zoals de knelpunten en succesfactoren, niet uniek voor innovatie met GenAI toepassingen, maar zijn ook relevant voor andere digitale innovaties.

Zet de maatschappelijke opgave centraal

Om ervoor te zorgen dat het toepassen van GenAI ten dienste staat van maatschappelijke opgave van de overheidsorganisatie is het belangrijk om vragen en behoeften vanuit beleidsdomeinen centraal te stellen. Uit interviews blijkt dat aansluiting van innovatieteams op beleidsdomeinen nog niet overal goed is ingebed. Uit de interviews komt naar voren dat het helpt als er wordt geredeneerd vanuit de maatschappelijke opgave. Hierdoor wordt de inzet van GenAI een middel in plaats van een doel en kan de ervaring van een technologie push worden getemperd. Daarbij kan het helpen om te monitoren of de beoogde doelen met de inzet van GenAI worden behaald.

Betrek domein experts en eindgebruikers bij ontwikkeling en inkoop van GenAI toepassingen

Om GenAI zo goed mogelijk bij de dagelijkse praktijk aan te laten sluiten is het belangrijk om domein experts en eindgebruikers te betrekken bij de ontwikkeling en/of inkoop van GenAI toepassingen. De eindgebruiker kan zowel burger als medewerker zijn. Hoewel dit inzicht niet nieuw is, geven de resultaten uit dit onderzoek aan dat domein-specifieke kennis en praktische

Tabel 5: Overzicht van aanbevelingen voor overheidsorganisaties

Aanbevelingen voor overheidsorganisaties
Zet de maatschappelijke opgave centraal
Betrek domein experts en gebruikers bij ontwikkeling en inkoop van GenAI-toepassingen
Neem de lange termijn impact op organisaties en medewerkers mee
Denk na over herbruikbaarheid van GenAI toepassingen
Besteed aandacht aan de rol van inkoop
Besteed aandacht aan digitale afhankelijkheden
Maak tijd voor het uitvoeren van impact toetsen

inzichten belangrijk zijn voor het succes van een GenAI toepassing. De aansluiting bij de maatschappelijke opgave en de praktijk kan beter worden geborgd, waarmee het succes van de toepassing en hoe deze landt bij de eindgebruiker wordt versterkt.

Neem de lange termijn impact op organisatie en medewerkers mee bij de opzet van een toepassing

Bij zowel succesfactoren als knelpunten komt naar boven dat toekomstgerichtheid essentieel is. Het is belangrijk om aan de start van de inzet van GenAI na te denken over de impact van GenAI op een

organisatie, hoe GenAI werk verandert, maar ook hoe andere organisaties een zelfde toepassing kunnen gebruiken. Door deze overwegingen mee te nemen kan er gebouwd worden aan een oplossing die beter aansluit bij de praktijk en opgeschaald kan worden. Door ook vooruit te kijken naar wat er in de toekomst nodig is aan middelen en infrastructuur en niet alleen de uitdagingen van morgen aan te pakken, ontstaan er betere toekomstbestendige situaties.

56 TNO, 2024

Zet in op herbruikbaarheid van GenAI toepassingen

In de verzameling van voorbeelden komen vaak dezelfde of zelfde soort GenAI toepassingen voor, zoals chatbots of transcriptie tools. Deze brede ervaring kan gedeeld onderling uitgewisseld worden, zodat andere organisaties daarvan kunnen leren. Er kan echter ook nog veel worden gewonnen bij het stroomlijnen van het delen van kennis en ervaring. Dit ondersteunt de herbruikbaarheid van toepassingen en maakt dat niet elke organisatie zelfstandig aan de slag moet met hun eigen (vaak beperkte) middelen.

Overheidsorganisaties die net starten met GenAI kunnen toepassingen overnemen of bijdragen aan initiatieven aanvullen met hun eigen kennis, ervaring en middelen.

Bij het ontwikkelen van nieuwe toepassingen, kan samenwerking met ander partijen opgezocht worden zodat er vanaf de start rekening gehouden kan worden met de herbruikbaarheid van een toepassing in andere contexten, organisaties of processen. Ook kan worden geïnvesteerd in het delen van ervaringen met bijvoorbeeld impact toetsen, inkoopprocessen en andere ondersteunende processen.

Besteed meer aandacht aan de rol van inkoop

Waar regel-gebaseerde AI en machine learning nog regelmatig door overheidsorganisaties zelf ontwikkeld konden worden, is dat met GenAI niet meer mogelijk. Taalmodellen worden ingekocht, of er worden bestaande open source modellen gebruikt. Hiermee kan niet meer worden gesteld dat GenAI volledig in house wordt ontwikkeld, waardoor de rol van inkoop of open source gebruik groter en belangrijker is dan in andere innovatietrajecten.

Naast de onderliggende modellen, zijn ook aspecten zoals infrastructuur, hosting van model en data, onderhoud van de toepassing en rechten ook belangrijke onderwerpen voor inkoop.

Besteed aandacht aan digitale afhankelijkheden

Doordat de ontwikkeling, hosting en onderhoud van GenAI niet meer volledig in eigen handen te houden is, is het belangrijk dat overheidsorganisaties zicht houden op de afhankelijkheden. Door in het ontwikkelen inkooptraject rekening te houden met leverancier van toepassing en model, de toegang en hosting van data, en andere onderwerpen, kunnen gerichte keuzes gemaakt worden die digitale autonomie kunnen versterken. Door te kiezen voor Europese alternatieven, of rekening te

houden met toekomstige overstap naar (open) Europese taalmodellen, blijft de organisatie wendbaar.

Maak tijd voor het uitvoeren van impact toetsen

Impact toetsen dragen bij aan verantwoorde en transparante inzet en ontwikkeling van GenAI. Het is vaak nog niet duidelijk of en welke impact toetsen worden gebruikt door overheidsorganisaties, maar dat wil niet zeggen dat er niet op potentiële impact wordt gereflecteerd. De praktijk laat zien dat impact toetsen tijd kosten en niet gemakkelijk zijn om uit te voeren. Ze helpen echter wel bij het vroegtijdig identificeren van mogelijke risico's en welke onbekendheden er nog bestaan. Deze inzichten dragen bij aan het maken van weloverwogen beslissingen rondom GenAI.

Door inzichtelijk te maken welke toetsen zijn gedaan en welke afwegingen zijn gemaakt, kunnen overheidsorganisaties beter verantwoorden aan de burgers en aantonen dat ze juridische kaders en afspraken naleven.

Limitaties en reflectie

Er zijn ook een aantal limitaties aan dit onderzoek. Dit onderdeel presenteert de limitaties en geeft een reflectie op de gebruikte methode.

Representativiteit onderzoek

Deel één van de resultaten van dit onderzoek is gebaseerd op publiek beschikbare informatie en inzichten die gedeeld zijn gedurende het onderzoeksproces. Deze onderzoeks aanpak heeft echter wel een aantal beperkingen. Doordat de gevonden voorbeelden via desk research zijn geïdentificeerd is de studie afhankelijk van de manier waarop door overheidsorganisaties wordt gecommuniceerd over de inzet van GenAI toepassingen. De mate van representativiteit voor de inzet van GenAI bij Nederlandse overheidsorganisaties kan daarom niet worden vastgesteld.

Deel twee van het onderzoek naar succesfactoren en knelpunten bij innovatie met GenAI is gebaseerd op 17 interviews met sleutelrollen betrokken bij GenAI toepassingen. Hierbij gaat het vooral om medewerkers in innovatieteams die betrokken waren bij de ontwikkeling en inzet van GenAI toepassingen. Daarbij hebben we alleen de aanpak en ervaringen van deze medewerkers geanalyseerd, we hebben niet de kwaliteit van de toepassingen zelf getoetst. Hoewel dit onderdeel van het onderzoek niet representatief is voor de gehele overheid, door respondenten uit verschillende type overheidsorganisaties te werven, biedt dit onderzoek inzichten die waardevol zijn voor zowel de experts uit de praktijk als beslissers en beleidsmakers.

Gebrek aan beschikbare data of eenduidige indeling van data

De gevonden voorbeelden in dit onderzoek zijn gecategoriseerd en geanalyseerd aan de hand van de volgende categorieën: type overheidsorganisatie, status toepassing, type proces, type taak, eindgebruiker toepassing, risico inschatting, impact toetsen en technische kenmerken. Deze categorieën bieden waardevolle inzichten in de inzet van GenAI bij overheidsorganisaties.

Over een aantal categorieën is momenteel nog onvoldoende publieke data beschikbaar om hier lessen uit te trekken:

1. **Risico inschatting en gebruikte impact toets.** Het valt op dat momenteel er nog geen eenduidige registratie van (Gen)AI toepassingen in het algoritmeregister staan. Ook zijn er verschillende interpretaties over de mogelijke impact en risico inschatting van een toepassing, zoals ook beschreven is in hoofdstuk 2. Informatie over risico inschatting is vaak onbekend. Ook observeren we dat een dezelfde soort toepassing bij verschillende organisaties in een andere risico categorie wordt ingedeeld. Op basis van de beschikbare data is bij 56% van de toepassingen onbekend of er een impact toets is uitgevoerd.

2. **Beperkt inzicht in technische specificaties en digitale afhankelijkheden.** Over technische kenmerken van GenAI, zoals de ontwikkelvorm, type model, land van oorsprong van het model, type aanbieder, land van oorsprong aanbieder, gebruikte infrastructuur en digitale afhankelijkheden is onvoldoende informatie beschikbaar. Ook is de informatie over technische kenmerken niet eenduidig, de ene toepassing geeft alleen inzicht in wie de leverancier is, terwijl een andere toepassing alleen het gebruikte model benoemd. Dit maakt het ook lastig om inzicht te krijgen in mogelijke digitale afhankelijkheden. Om meer inzicht te krijgen in dit onderwerp is meer vervolgonderzoek nodig.

3. **Gebruikte middelen en kosten.** Middelen zijn een belangrijke facilitator voor de succesvolle inzet van GenAI. Financiële middelen kunnen uit verschillende bronnen komen. Toch bestaat ook de vraag in de praktijk hoe de overgang van experiment en tijdelijke middelen naar implementatie en vaste middelen bewerkstelligd kan worden. Publieke data geeft momenteel geen inzicht in de gebruikte middelen en kosten van GenAI toepassingen.

Reflectie op methode

Via interviews, enquêtes, de analyse van strategie- en begrotingsdocumenten en het analyseren van audit en verantwoordingsverslagen zou deze informatie in de toekomst meer gestructureerd kunnen worden opgehaald.

Daarnaast is het belangrijk om naast het monitoren van GenAI ontwikkelingen ook andere vormen van AI toepassingen bij overheidsorganisaties te blijven monitoren. Dit onderzoek toont aan dat GenAI wordt ingezet in overheidsprocessen waarbij een hoge mate van gegevens en tekstverwerking plaatsvindt. Daarmee zijn de aard van machine learning toepassingen dan ook diverser dan GenAI toepassingen. Zo wordt GenAI bijvoorbeeld niet ingezet in toepassingen in het fysieke domein.

Acknowledgements

We willen de respondenten voor dit onderzoek bedanken voor het delen van hun ervaringen en inzichten.

Ook willen we Romy van Drie en Anne Fleur van Veenstra bedanken voor het reviewen en aanscherpen van onze inzichten.

Referenties

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., Arx, S. von, Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2022). *On the Opportunities and Risks of Foundation Models* (No. arXiv:2108.07258). arXiv. <https://doi.org/10.48550/arXiv.2108.07258>
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V. Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic Analysis. In P. Liamputtong (Ed.), *Handbook of Research Methods in Health Social Sciences* (pp. 843–860). Springer. https://doi.org/10.1007/978-981-10-5251-4_103
- Cinar, E., Trott, P., & Simms, C. (2019). A systematic review of barriers to public sector innovation process. *Public Management Review*, 21(2), 264–290. <https://doi.org/10.1080/14719037.2018.1473477>
- Garon, J. (2023). *A Practical Introduction to Generative AI, Synthetic Media, and the Messages Found in the Latest Medium* (SSRN Scholarly Paper No. 4388437). Social Science Research Network. <https://doi.org/10.2139/ssrn.4388437>
- Jöhnk, J., Weißert, M., & Wyrski, K. (2021). Ready or Not, AI Comes—An Interview Study of Organizational AI Readiness Factors. *Business & Information Systems Engineering*, 63(1), 5–20. <https://doi.org/10.1007/s12599-020-00676-7>
- Mehr, H. (2017). *Artificial intelligence for citizen services government*. Ash Center for Democratic Governance and Innovation. https://ash.harvard.edu/wp-content/uploads/2024/02/artificial_intelligence_for_citizen_services.pdf
- Oomen, T.A.P. (2025). Practices we value: A multimethodological exploration of trustworthy artificial intelligence development [proefschrift]. Opgevraagd van: <https://pure.eur.nl/en/publications/practices-we-value-a-multimethodological-exploration-of-trustwort>
- Rathenau Instituut (2023). *Generatieve AI*. Den Haag. Auteurs: Hamer, J., L. Kool, B. Hijstek, Q. van Eeden en D. Das. Opgevraagd van: https://www.rathenau.nl/sites/default/files/2023-12/Scan_Generatieve_AI_Rathenau_Instituut.pdf
- TNO (2019). Quicksan AI in de publieke dienstverlening. Opgevraagd van: <https://www.digitaleoverheid.nl/document/quick-scan-ai-in-de-publieke-dienstverlening/>
- TNO (2021). Quicksan AI in de publieke dienstverlening II.
- TNO (2024). Quicksan AI in de publieke dienstverlening III. Opgevraagd van: <https://vector.tno.nl/artikelen/transparantie-ai-toepassingen-overheid/>
- TNO / Rathenau Instituut (2024) Eigen ritme of algoritme? – Een verkenning van algoritmisch management voorbij de platformeconomie. Auteurs: Djurre Das, Thijmen Zoomer, Liza van Dam, Linda Kool, Paul Preenen, Wouter van der Torre. Opgevraagd van: <https://www.rathenau.nl/digitalisering/naar-een-menswaardige-digitale-technologie/eigen-ritme-algoritme>
- TNO (2025). Aan het werk met generatieve AI. Verkenning op basis van literatuur en case studies. Opgevraagd van: <https://www.tno.nl/nl/newsroom/insights/2025/07/gen-ai-verandert-werk>

Auteurs

Marissa Hoekstra
Tessa Bruijne
Marianne Schoenmakers
Pippa Jones

Contact

marissa.hoekstra@tno.nl

TNOvector
Centre for Societal Innovation and Strategy

December 2025

tnovector.nl