

Whitepaper

AI-governance binnen de overheid

Onze visie op bestuurbare, traceerbare en bewijsbare AI.

Van AI-experiment naar productie.

Managementsamenvatting

De verschuiving

AI verschuift binnen de overheid van 'experimenteren' naar productie. Niet alleen in ondersteunende processen, maar ook in juridisch- en bestuurlijk gevoelige werkzaamheden zoals bezwaarafhandeling, Woo-verzoeken en besluitvoorbereiding.

Daarmee verandert ook de centrale vraag. Niet langer staat centraal óf AI nuttig is, maar of het gebruik bestuurlijk beheersbaar, uitlegbaar en achteraf verdedigbaar blijft.

Het nieuwe bestuurlijke risico

Wanneer AI invloed heeft op informatieverwerking, beoordeling of besluitvorming, moet een organisatie later kunnen aantonen:

- welke AI-ondersteuning is gebruikt;
- op basis van welke bronnen;
- onder welke menselijke controle;
- met welke invloed op het proces;
- en welke versie uiteindelijk bestuurlijk relevant werd.

Zonder die aantoonbaarheid ontstaat bestuurlijk risico, óók wanneer de technologie technisch goed functioneert.

Managementsamenvatting

De nieuwe norm

AI-governance hoort niet thuis in controle achteraf, maar in de workflow zelf. Dat betekent dat organisaties vooraf begrenzen:

- welke AI-toepassingen zijn toegestaan;
- welke bronnen gebruikt mogen worden;
- waar menselijke validatie verplicht is;
- en welke processtappen volledig herleidbaar moeten blijven.

Hoe Limescape dit vertaalt

Limescape vertaalt AI-governance naar operationele productcontrols, waaronder:

- risicoclassificatie per use-case;
- bronrestrictie;
- menselijke validatie;
- logging en traceerbaarheid;
- auditexport;
- exporteerbare bewijsobjecten.

Daarmee wordt AI-governance niet alleen beleid, maar afdwingbaar gedrag in het werkproces.

Limescape maakt AI niet juridisch risicovrij, maar bestuurlijk beheersbaar, technisch begrensd en achteraf bewijsbaar.

Waarom dit vraagstuk nu bestuurlijk wordt

AI is niet langer een experiment aan de rand van de organisatie. Medewerkers gebruiken AI inmiddels in uiteenlopende vormen: van informatieverwerking en dienstverlening tot procesondersteuning en besluitvoorbereiding.

Dat levert snelheid op, maar introduceert tegelijk een nieuwe bestuurlijke kwetsbaarheid: kan de organisatie later nog aantonen welke AI-ondersteuning is gebruikt, op basis van welke bronnen, onder welke menselijke controle en met welke gevolgen voor het proces?

Juist in hoog-risico overheidsprocessen is die vraag niet theoretisch. Wanneer AI raakt aan:

- rechtmatigheid;
- termijnen;
- bezwaar en beroep;
- rechtspositie;
- of bestuurlijke besluitvorming,

raakt AI indirect ook aan de positie van burgers.

Een foutieve samenvatting, gemiste contra-indicatie of niet-herleidbare bron hoeft dan geen technisch incident te blijven, maar kan doorwerken in de beoordeling van een aanvraag, bezwaar of besluit.

Toenemende druk vanuit wetgeving en toezicht

Tegelijkertijd nemen de bestuurlijke verwachtingen snel toe. Binnen de Nederlandse overheidscontext lopen meerdere ontwikkelingen parallel:

- de AI Act;
- BIO 2.0;
- modernisering van digitaal bestuurlijk verkeer;
- eisen rond informatiehuishouding;
- en zorgen over schaduwgebruik van AI.

AI-gebruik kan daardoor niet langer worden behandeld als een los innovatie-experiment.

Kernprincipe

De centrale ontwerpregel is eenvoudig:

AI mag pas productief worden ingezet wanneer AI-governance onderdeel is van de workflow en niet van de nazorg.

Beheersing hoort dus in het proces te zitten, niet in een controleverslag achteraf.

Dat principe vertaalt zich in vier eisen:

1. Elke AI-toepassing wordt vooraf ingeschaald naar risico en toegelaten binnen een expliciet governancekader.
2. Elke AI-uitkomst blijft onder menselijke regie en mag niet blind worden geaccepteerd.

3. Elke interactie moet dossierherleidbaar zijn, inclusief bron, versie, actor en tijdstip.

4. Elke procedure moet achteraf exporteerbaar en verdedigbaar zijn als procesbewijs.

Van norm via beheersmaatregel naar bewijsobject

Voor overheidsorganisaties wordt AI-governance pas overtuigend wanneer een norm niet alleen wordt benoemd in uitgebreide compliance-documentatie, maar wanneer het wordt vertaald naar werkende beheersmaatregelen en controleerbare bewijsobjecten. De tabel hieronder vormt daarom het hart van deze whitepaper.

NORM OF BESTUURLIJKE EIS	RISICO ZONDER BORGING	LIMESCAPE-BEHEERSMAATREGEL	BEWIJSOBJECT
AI Act: classificatie per intended purpose	Eén generieke risicoclassificatie maskeert dat sommige functies licht zijn en andere hoog-risico	Use-case-registratie, functiecategorie, risicoprofiel en toegestane AI-rol per processtap	Classificatiedossier per AI-toepassing
AI Act: effectief menselijk toezicht	Human-in-the-loop blijft een vinkje en voorkomt automation bias niet	Smart Friction Gate met verplichte validatie, wijzigingsregistratie en blokkade op blinde vrijgave	Validatielog met actor, tijdstip, broncontrole en beslissing
AI Act en AVG: transparantie, bronbasis en dataminimalisatie	AI-output is niet uitlegbaar of gebruikt meer persoonsgegevens dan nodig	Bronrestrictie, promptbeheer en gecontroleerde contextselectie	Bronverantwoording en gebruikte context per output
BIO 2.0 en informatiebeveiliging	Laag-risico en hoog-risico informatiestromen lopen door elkaar	Mode-scheiding tussen ondersteunende, hybride en inhoudelijke functies	Configuratie- en autorisatielog per mode
Awb, bezwaar en behoorlijk bestuur	Besluitvorming wordt beïnvloed zonder zicht op zorgvuldigheid, motivering en menselijke beoordeling	Verplichte reviewstappen bij conceptadviezen, motiveringen en besluitvoorbereiding	Beslisroute met AI-voorstel, menselijke aanpassing en vrijgave
IAMA, algoritmeregister en audit	Documentatie blijft statisch en sluit niet aan op productiegedrag	Governance Cockpit en Evidence Vault met monitoring, signalering en export	Auditexport, incidenthistorie en verantwoordingsrapportage

De essentie is dat Limescape niet alleen vastlegt dát er beleid is, maar afdwingt hoe dat beleid in het werkproces wordt toegepast. Daarmee wordt AI-governance een operationele ontwerpdiscipline in plaats van een papieren maatregel.

Positionering van Limescape

Limescape is geen generieke AI-beheertool en geen extra dashboard bovenop willekeurige AI-tools. De positionering is scherper: Limescape is de beheersingslaag die AI-gebruik in gereguleerde processen inricht, begrenst, logt en bewijsbaar maakt.

Het product grijpt dus niet primair in op het taalmodel, maar op de procesmatige voorwaarden waaronder AI-output gebruikt mag worden.

Die keuze kent drie redenen:

- het sluit aan op bestuurlijke taal: risico, controle, audit, verantwoording en beleid;
- het voorkomt concurrentie op generieke AI-functionnalitét die snel commoditiseert;
- het maakt het mogelijk om per domein een scherp praktijkvoorbeeld te tonen zonder de productkern te verliezen.

Voor de productrichting betekent dit dat Limescape niet moet worden ontwikkeld als verzameling losse AI-features, maar als samenhangende governance-architectuur.

AI-governancekader en risicoclassificatie

Onboarding als beleidsmatige toelating

Onboarding is binnen Limescape niet primair gebruikerstraining of adoptie. Het is de formele toelating van een AI-use-case binnen een bestaand proces, op basis van een vooraf vastgelegd risicokader.

Niet de medewerker staat centraal, maar de combinatie van use-case, processtap, bronset, beslisimpact en toegestane output. Een organisatie laat dus niet simpelweg een model toe, maar een afgebakende toepassing onder begrensde voorwaarden.

Risicocategorieën en intended purpose

Een werkbaar governancekader begint niet bij het platform als geheel, maar bij de concrete functie en het beoogde gebruik. In de publieke sector is dat cruciaal. Dezelfde AI-techniek kan ondersteunend zijn wanneer zij een tekst letterlijk omzet, maar hoog-risico worden wanneer zij feiten weegt, juridische onderbouwing genereert of een conceptbesluit voorbereidt.

FUNCTIECATEGORIE	WAT DE GEBRUIKER ZIET	BESTUURLIJKE DUIDING	LIMESCAPE-BORING
Ondersteunend	AI helpt bij administratieve, transformatieve of informatieve taken, zoals transcriptie, letterlijke vertaling, anonimisering of B1-herschrijving	Laag of beperkt risico zolang de functie strikt afgebakend blijft en geen nieuwe inhoudelijke beoordeling toevoegt	Functierestrictie, bronbegrenzing, promptrestrictie en basislogging
Hybride	AI vat samen, ordent of presenteert informatie die oordeelsvorming kan beïnvloeden	Grijs gebied: selectie, prioritering of samenvatting kan ongemerkt sturend worden	Brontraceerbaarheid, interface-signalen, blokkade van sturende conclusies en monitoring
Inhoudelijk	AI ondersteunt juridische beoordeling, adviesvorming of besluitvoorbereiding	Hoog-risico zodra de output rechtspositie, motivering of beoordeling kan beïnvloeden	Smart Friction Gate, Evidence Vault, Prompt Vault, bronverantwoording, monitoring en exporteerbaar bewijsdossier

Deze indeling maakt vooraf zichtbaar welke AI-vormen organisatorisch toelaatbaar zijn, welke extra beheersing vragen en waar alleen productie mogelijk is onder zwaar toezicht.

Beslisrechten

Een volwassen governancekader bepaalt ook wie waarover beslist:

- **Proceseigenaar:** bepaalt of AI binnen een proces functioneel wenselijk is.
- **Juridisch of inhoudelijk verantwoordelijke:** bepaalt of output in de procescontext inhoudelijk verantwoord kan worden gebruikt.
- **FG, CISO of governancefunctie:** beoordeelt of inzet past binnen risicoregime, bronrestricties en loggingeisen.
- **Product- of applicatiebeheer:** beheert de configuratie, maar is niet de enige beslisser over inhoudelijke toelaatbaarheid.

Zonder deze scheiding ontstaat een bekend patroon: IT beheert de techniek, de business gebruikt de uitkomst, maar niemand is eigenaar van de grens tussen ondersteunende AI en bestuurlijk risicodragende AI.

Biaspreventie

Biaspreventie is in gereguleerde processen niet slechts een visuele kwaliteitsmodule, maar een governancevraagstuk. Het gaat om bronkeuze, contextverlies, automation bias en de vraag of de workflow voldoende tegenkracht organiseert.

Limescape benadert biaspreventie daarom via vier ontwerpregels:

1. **Bronrestrictie:** AI werkt alleen op goedgekeurde bronnen. Vrije webtoegang of impliciete modelkennis is in hoog-risico processen onvoldoende controleerbaar.
2. **Bronverwijzing als gebruiksvoorwaarde:** Elke inhoudelijke AI-uitkomst moet herleidbaar zijn naar de gebruikte bronbasis. Kan de herkomst niet worden getoond, dan is de output niet vrijgeefbaar.

3. Menselijke frictie tegen blind accepteren: De workflow moet zichtbare, sobere frictie inbouwen: verplichte validatiecheck, blokkade van definitieve vrijgave, bevestiging van broncontrole en logging van acceptatie, wijziging of afwijzing.

4. Signaalsturing op afwijkende patronen: Beheersing wordt sterker wanneer niet alleen individuele outputs, maar ook patronen zichtbaar worden, zoals afwijkend promptgebruik, ongebruikelijke acceptatie zonder wijziging of structurele verschillen tussen vergelijkbare dossiers.

Logging en traceerbaarheid

Veel organisaties loggen technisch al veel, maar bestuurlijk te weinig. Server-events en modelcalls vormen nog geen bruikbare proceshistorie.

Limescape richt zich daarom op dossierherleidbaarheid op de grens van AI en mens. De logstructuur moet vastleggen welke broncontext is gebruikt, welke output is ontstaan, wat de medewerker heeft overgenomen of gecorrigeerd en welk moment uiteindelijk procesrelevant werd.

Minimale logstructuur

LOGOBJECT	WAT WORDT VASTGELEGD	WAAROM DIT BESTUURLIJK TELT
Context	Proces, dossier of zaak, stap, rol, omgeving	Maakt duidelijk binnen welk proces AI is gebruikt
Input	Gebruikte documenten, geselecteerde bronset, promptcategorie of instructielaag	Toont waarop de output is gebaseerd
Output	Ge genereerde tekst, analyse, signalering of voorstel	Toont wat het systeem feitelijk heeft geproduceerd
Menselijke interventie	Geaccepteerd, aangepast, afgewezen, vrijgegeven	Maakt human-in-the-loop concreet en toetsbaar
Tijd en identiteit	Tijdstip, gebruiker, rol, versie	Voorkomt reconstructie op basis van aannames
Beleidsstatus	Toegepaste instructielaag, actief beleid, risicoregime	Verbindt systeemgedrag aan formele governance

Onwijzigbaarheid en retentie

Waar het proces daarom vraagt, wordt logging onwijzigbaar of functioneel 'append-only' opgeslagen. Niet omdat elk scenario cryptografische zwaarte vereist, maar omdat bestuurlijke betrouwbaarheid afhangt van de vraag of het logboek achteraf nog herschreven kan worden.

Daarbij horen ook retentie, exporteerbaarheid en toegangsrechten. Zonder die discipline verschuift het risico slechts van proces naar archief.

Bewijslast

Bewijslast is geen extra rapportagefunctie naast de operatie. Zij is de test of de AI-governance werkelijk deugt.

Wanneer een organisatie later niet kan aantonen hoe AI is gebruikt, welke controle is toegepast en waarom een uitkomst verantwoord was, dan was de governance ondanks dashboards en beleidsnotities feitelijk onvoldoende.

Een sterke bewijspositie rust op vier elementen:

1. **Reproduceerbaarheid:** welke invoer, bronset en configuratie lagen aan de AI-bijdrage ten grondslag?
2. **Menselijke toerekenbaarheid:** welke medewerker of rol heeft gevalideerd, aangepast of vrijgegeven?
3. **Versiezuiverheid:** welke concepten zijn gegenereerd en welke versie is formeel procesrelevant geworden?
4. **Exporteerbaarheid:** kan een dossieroverzicht ook buiten het systeem bestuurlijk leesbaar worden gemaakt?

Met zo'n bewijslaag kan een organisatie gericht aantonen wat AI heeft voorbereid, wat de mens heeft gecontroleerd en welke versie bestuurlijk relevant werd.

Octopus als praktijkvoorbeeld in juridisch gevoelige processen

Octopus is niet de definitie van Limescape, maar wel het scherpste praktijkvoorbeeld van waarom Limescape nodig is. Het product raakt aan juridische werkprocessen waarin termijnen, dossieropbouw, correspondentie, brongebruik, adviesvorming en besluitvoorbereiding samenkomen.

Juist in bezwaar, beroep, klachten, Woo-verzoeken en verwante procedures is de combinatie van dossierintegriteit, menselijke controle en proceshistorie organisatiekritisch. Een ontbrekend document, gemiste termijn of niet-herleidbare samenvatting is daar geen klein proceslek, maar kan direct raken aan rechtsbescherming van burgers, bestuurlijke verantwoording of financiële gevolgen.

De demonstratiewaarde van Octopus zit in vier punten:

- het maakt zichtbaar hoe AI binnen dossiercontext kan werken in plaats van als vrije chatlaag;
- het toont waarom menselijke validatie een productvoorwaarde is;
- het laat zien hoe dossierherleidbaarheid bestuurlijke geloofwaardigheid oplevert;
- het maakt concreet dat AI-governance niet alleen voor auditors is, maar ook voor juristen, teamleiders, bestuur en informatieverantwoordelijken.

Drie overheidscasussen

Casus 1. 'Bezwaar' en herleidbaarheid

Situatie

Een behandelaar gebruikt AI om een bezwaarschrift, dossierstukken en eerdere correspondentie samen te vatten. De winst zit in snelheid, structuur en minder administratieve druk.

Risico

Wanneer AI medische, financiële of juridische argumenten weglaat, samendrukt of verkeerd ordent, kan dat de beoordeling beïnvloeden. In bezwaar is achteraf reconstrueren niet genoeg.

Limescape-borging

Limescape borgt bronverwijzing, mode-scheiding, verplichte menselijke validatie en logging van de uiteindelijke beoordeling. De behandelaar blijft verantwoordelijk voor de overgenomen feiten.

Bewijsobject

Een gevalideerde samenvatting met bronpassages, reviewlog, actor, tijdstip en vastlegging van wijzigingen of afwijzingen.

Casus 2. 'Woo-verzoek' en documentinventarisatie

Situatie

AI helpt bij een omvangrijk Woo-verzoek om documenten te vinden, te clusteren en voor te bereiden voor beoordeling of weglakking. Dat kan veel tijd winnen bij grote documentvolumes en krappe beslistermijnen.

Risico

Zolang AI uitsluitend ontsluit, ordent of transformeert, is de rol ondersteunend. Maar zodra AI documenten prioriteert, uitsluit, inhoudelijk duidt of risico's signaleert, ontstaat een hybride functie. Dan moet zichtbaar zijn welke documenten zijn meegenomen.

Limescape-borging

Limescape maakt dit controleerbaar via logging, bronverantwoording, promptregistratie en de Evidence Vault. Daarmee wordt versnelling niet losgetrokken van transparantie.

Bewijsobject

Een documentselectielog met meegenomen en uitgesloten documenten, gebruikte instructies, bronverwijzingen, beoordelingsmomenten en menselijke correcties.

Casus 3. Conceptadvies en -besluit

Situatie

AI stelt een conceptmotivering, conceptadvies of tekstvoorstel op. Dit is de zwaarste categorie, omdat de output kan doorwerken in rechtspositie, motivering en bestuurlijke besluitvorming.

Risico

De overheid moet steeds vaker uitleggen hoe digitale systemen besluitvorming ondersteunen. Transparantie is dan geen bijlage achteraf, maar onderdeel van motivering en toetsbaarheid. Zeker wanneer AI helpt bij het toepassen van feiten op regels, moet duidelijk zijn wat AI heeft voorgesteld, welke bronnen zijn gebruikt en hoe bias of ongelijke behandeling is tegengegaan.

Limescape-borging

Limescape activeert hier de volledige set beheersmaatregelen: gecontroleerde bronbasis, Prompt Vault, Smart Friction Gate tegen blind accepteren, Evidence Vault, monitoring en escalatie waar nodig.

Bewijsobject

Een exporteerbaar bewijsdossier met bronbasis, prompt- en instructieversie, AI-output, menselijke validatie, aanpassingen, vrijgave en eventuele escalatie.

Volwassenheidsladder voor AI-governance

Een effectieve invoering verloopt niet via één groot governanceprogramma, maar via oplopende hardheid. De logische ontwikkellijn is die van basisbeheersing naar schaalbare governance.

1: Basisbeheersing

Use-cases, intended purpose, functiecategorieën, broncategorieën, rollen en beslisrechten zijn vastgelegd.

Productfocus: use-case register, bronkader, basislogging, autorisaties en mode-scheiding.

2: Workflowcontrole

AI-output kan niet blind worden doorgezet. Menselijke regie is zichtbaar en technisch afgedwongen in het proces.

Productfocus: Smart Friction Gate, verplichte broncontrole, reviewstappen en logging van acceptatie, wijziging of afwijzing.

3: Bewijsfundament

Proceshistorie wordt dossiergebonden, versioneerbaar en exporteerbaar vastgelegd.

Productfocus: Evidence Vault, versiebeheer, append-only opslag waar nodig, retentie-instellingen en reconstructie per dossier.

4: Toezicht en signalering

De organisatie gaat sturen op afwijkingen in productiegebruik in plaats van alleen op incidentanalyse achteraf.

Productfocus: Governance Cockpit, signalen op blind accepteren, afwijkend brongebruik, promptwijzigingen, foutpatronen en escalatieregels.

5: Standardisatie en schaalbaarheid

Governancepatronen en bewijsobjecten worden herbruikbaar over meerdere processen en domeinen heen.

Productfocus: auditexport, verantwoordingsrapportages, templates voor IAMA en algoritmeregister, herbruikbare domeinprofielen en integraties met zaaksystemen, DMS, archief en informatiebeveiligingsprocessen.

Deze volwassenheidsopbouw maakt AI-governance bestuurlijk hanteerbaar en technisch realistischer. Zij biedt snel zichtbare grip, zonder te suggereren dat volledige beheersing in één stap kan worden gerealiseerd.

Bestuurlijke beslisvragen

Voor bestuur, CISO, FG en proceseigenaren zijn uiteindelijk vooral deze vragen relevant:

- *Welke AI-toepassingen vallen in welke risicocategorie, en wie heeft dat vastgesteld?*
- *Is per toepassing vastgelegd wat de intended purpose is en waar de grens van toegestaan gebruik ligt?*
- *Welke bronnen mag de AI gebruiken, en welke standaard niet?*
- *Welke persoonsgegevens of gevoelige kenmerken blijven buiten de context, tenzij een zwaarder regime expliciet is geactiveerd?*
- *Op welke momenten is menselijke validatie verplicht voordat een uitkomst verder mag in het proces?*
- *Kan later per dossier worden aangetoond wat AI heeft voorgesteld en wat de mens daarmee heeft gedaan?*
- *Bestaat er een exporteerbaar bewijsobject dat bestuurlijk uitlegbaar is zonder technische reconstructie?*
- *Zien FG, CISO en proceseigenaar in productie of gebruikers AI-output blind accepteren, afwijkende bronnen gebruiken of structurele foutpatronen veroorzaken?*

Een organisatie die deze vragen niet concreet kan beantwoorden, heeft geen volwassen AI-governance, ongeacht het aantal beleidsnotities of dashboards.

Conclusie

AI gaat de overheid niet vervangen, maar wel steeds vaker meewerken in informatieverwerking, dienstverlening en besluitvoorbereiding. De kernvraag is daarom niet óf AI mag worden gebruikt, maar of AI binnen bestuurlijke controle blijft.

Voor Limescape betekent dit een duidelijke productrichting: classificeren wat mag, beperken wat AI ziet, afdwingen waar de mens valideert, vastleggen wat gebeurt en exporteren wat later bewezen moet worden.

Daarmee vertaalt Limescape AI-governance naar operationeel productontwerp.

De bredere claim blijft: **bestuurbare AI-productie wordt de volwassenheidsgrens voor AI in de overheid**. Niet omdat beheersing innovatie afremt, maar omdat duurzame, uitlegbare en juridisch verdedigbare AI-inzet zonder beheersing niet houdbaar is.

AI-governance wordt daarmee minder een beleidsdossier en meer een ontwerpdiscipline voor productie. Juist dat maakt Limescape bestuurlijk geloofwaardig: niet omdat het zegt dat AI verantwoord wordt gebruikt, maar omdat het verantwoordelijkheid technisch en procesmatig afdwingbaar en aantoonbaar maakt.

Bronnen en normatieve context

De visie op AI-governance en de ontwerpprincipes van Limescape zijn een technische en procesmatige vertaling van kaders waaraan de Nederlandse overheid nu en in de komende jaren moet voldoen.

Geraadpleegde kaders en bronnen

Wet- en regelgeving

- AI Act: Verordening (EU) 2024/1689.
- BIO 2.0: Baseline Informatiebeveiliging Overheid 2.0.
- Woo en Archiefwet 1995.

Governance en beleid

- VNG Governancekader AI en algoritmen.
- IAMA: Impact Assessment Mensenrechten en Algoritmes.
- Nationaal Algoritmeregister.

Jurisprudentie en onderzoek

- SyRI-uitspraak, Rechtbank Den Haag, 5 februari 2020, ECLI:NL:RBDHA:2020:865.
- Rapport Ongekend onrecht en vervolgpublishaties over de toeslagenaffaire.
- Publicaties van het Adviescollege Openbaarheid en Informatiehuishouding.

Deze normatieve context verklaart waarom de productrichting van Limescape niet begint bij modelkracht, maar bij beheersbaarheid, traceerbaarheid, menselijke regie en bewijsvoering.