bjørn®

Ai ai ai ai...

Governance, Risk en Compliance in de AI hype

Microsoft Kennisevent Noord-Nederland
Woensdag 28 oktober 2026, Heerenveen

Magie bestaat niet, ook niet als de demo goed ging.

bjørn regelt het!

AI versnelt alles.

Ook wat je liever niet zag.

- Risico's die je kon negeren omdat ze moeilijk te vinden waren, zijn nu sneller te vinden, ook voor aanvallers
- AI gebruiken brengt eigen, nieuwe risico's mee
- Goed ingezet geeft AI inzicht en grip die je nooit had

Dezelfde snelheid, twee kanten.

De fundamenten: informatie

Waar het allemaal om draait.

data — ruwe feiten, nullen en enen
\+
waarde — wat het voor jou betekent
=
informatie — wat je moet beschermen

- Zonder waarde is het alleen data
- Wat waarde heeft, moet beschermd worden: dat is informatiebeveiliging

Eerst weten wat je hebt, dan pas beveiligen.

Informatiebeveiliging: CIA en BIV

Drie doelen, in het Engels en in het Nederlands.

Vertrouwelijkheid
- Toegangscontrole
- Versleuteling

Integriteit
- Authenticiteit
- Onweerlegbaarheid
- Betrouwbaarheid

Beschikbaarheid
- Redundantie
- Herstelvermogen
- SLA's

CIA in het Engels, BIV in het Nederlands. De andere CIA doet hier niet mee.

CIA Confidentiality, Integrity, Availability **BIV** Beschikbaarheid, Integriteit, Vertrouwelijkheid **SLA** service level agreement

Informatietechniek: lagen op lagen

Elke laag verbergt de laag eronder.

AI-model en agent	gedrag zit in gewichten, niet in leesbare code
Applicatie en gebruikersinterface	apps en API's
Framework en runtime	compiler, bibliotheken
Besturingssysteem	bestandssysteem, netwerkstack
Hardware-abstractie	drivers, firmware
Hardware	instructieset, bits



Wat de leverancier beheert, beheer jij niet meer.

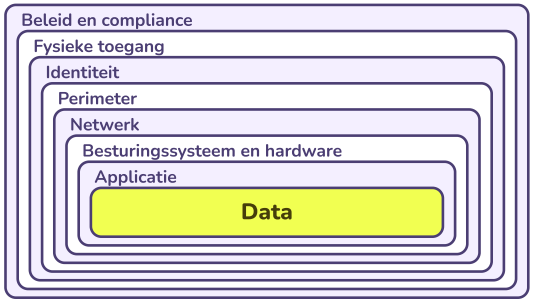
Jij blijft verantwoordelijk voor risico, gegevens en gebruik.

Een model kun je downloaden, maar niet lezen.

AI artificial intelligence **API** application programming interface **SaaS/PaaS/IaaS** software / platform / infrastructure as a service

Defense in Depth

Meerdere lagen, die elk kunnen falen.



- BIV op alle lagen
- Elke laag heeft maatregelen, een eigenaar en bewijs
- Controle: werkt het echt?

Eén laag is hopen op het beste.

BIV Beschikbaarheid, Integriteit, Vertrouwelijkheid

Risicobeheer

Kans maal impact. Daarna: wat doe je ermee?



Risico? Je mag niet niks doen.

Vier manieren om met een risico om te gaan

Verminderen
Een maatregel treffen die kans of impact verkleint.

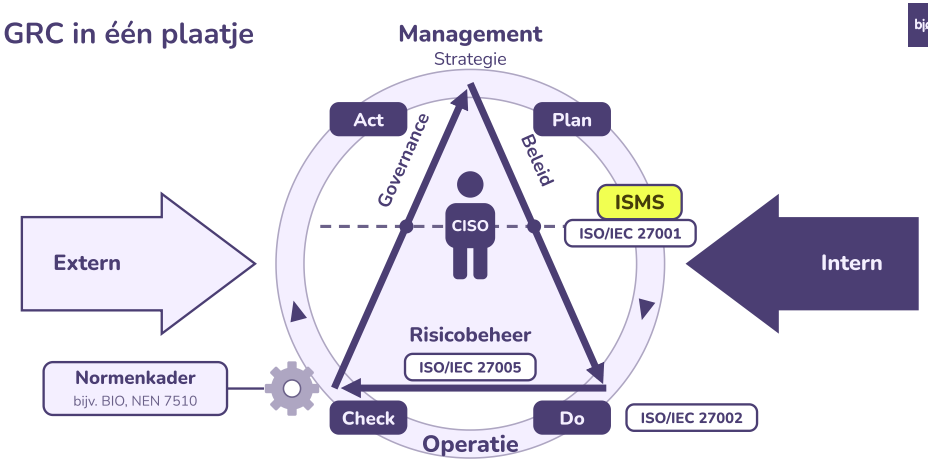
Vermijden
De activiteit stoppen of niet beginnen.

Delen
Verzekeren of uitbesteden. De verantwoordelijkheid blijft bij jou.

Accepteren
Bewust dragen, door de risico-eigenaar, vastgelegd.

Een risico dat niemand bezit, is een risico dat niemand behandelt.

GRC in één plaatje



GRC Governance, Risk and Compliance **CISO** Chief Information Security Officer **ISMS** information security management system
ISO International Organization for Standardization **IEC** International Electrotechnical Commission
BIO Baseline Informatiebeveiliging Overheid **NEN** Nederlands Normalisatie-instituut

'AI' in het kort

Voor wie het vooral kent van horen zeggen.

Taalmodel

- Voorspelt het volgende woord uit enorme hoeveelheden tekst
- Klinkt overtuigend, kan fout zijn
- Geen geheugen, tenzij je het meegeeft

Context

- Het model ziet alleen wat je meestuurt
- Wat je meestuurt, verlaat je omgeving
- Tekst van buiten kan instructies bevatten

Agent

- Een model met tools dat zelf stappen kiest
- Werkt met jouw rechten
- Fouten worden acties

Een heel goede autocomplete, met soms te veel rechten.

AI artificial intelligence

Begrippen: tool, MCP, harness

Begrip	Wat het is	Waarom het ertoe doet
Tool	Een functie die het model mag aanroepen: zoeken, lezen, schrijven, uitvoeren	Elke tool is een recht
MCP	Open protocol om tools en gegevensbronnen aan een model te koppelen	Elke server is een leverancier
Harness	De software om het model heen: lus, tools, rechten, geheugen (Codex, Claude Code, OpenCode)	Hier zitten rechten en logging
Agent	Model plus harness dat zelf stappen kiest	Fouten worden acties
RAG	Opgehaalde documenten in de context	Rechten en tekst van buiten
Sandbox	Afgeschermde omgeving waarin de agent draait	Isolatie is een besluit, geen standaard

Het model denkt mee, de harness bepaalt wat het mag.

MCP Model Context Protocol **RAG** retrieval-augmented generation

Papier en praktijk

Drie keer: wat de norm vraagt, en wat er gebeurde.

Geval	Wat de norm vraagt	Wat er gebeurde
CERT Polska, december 2025	NIS2 artikel 21 lid 2 onder j: MFA	VPN bereikbaar vanaf internet zonder MFA, bij minstens 30 locaties
NorthC Almere, mei 2026	ISO/IEC 27002 5.30 en 8.14; NIS2 artikel 21 lid 2 onder c: continuïteit	Stroom uit op last van de brandweer; uitwijk was aan de klant; veel klanten dagen offline
MW10 Homenet, 2026	ISO/IEC 27002 8.8 en 8.9: kwetsbaarheden en configuratie	Een handmatige updaterun sloeg één server over; een agent merkte het op

NIS2 vraagt niet om beschrijven, maar om aantonen.

CERT Computer Emergency Response Team **NIS2** Network and Information Security Directive 2
MFA multi-factor authentication **VPN** virtual private network **ISO** International Organization for Standardization
IEC International Electrotechnical Commission

Albert Heijn, oktober 2026

Alleen de feiten. Oorzaak en techniek zijn niet bekend.

- Zondag 11 oktober: IT-systemen preventief offline, distributiecentra stil
- Ahold Delhaize: "het gaat om kwaadwillige acties van derden"; melding bij het NCSC binnen 24 uur
- Maandag 19 oktober: aangifte bij de politie
- Tot nu toe geen aanwijzingen voor gestolen klant-, medewerkers- of derdengegevens
- Thuisbezorging pas weer vanaf vrijdag 23 oktober: meer dan elf dagen uitval

Wat je preventief uitzet, moet je ook weer aan kunnen krijgen.

IT Information Technology **NCSC** Nationaal Cyber Security Centrum

De drempel voor aanvallers zakt

Kosten omlaag, opbrengst gelijk.

- Geavanceerde aanvallen vragen geen geavanceerde aanvallers meer (Anthropic, 2026)
- Kosten per aanval dalen, opbrengst niet: marginale doelwitten worden aantrekkelijk
- Eén aanvaller: meer dan 2.100 tokensets van meer dan 40 organisaties in ongeveer 34 uur
- Mandiant: tijd tot misbruik geschat op −7 dagen, dus vaak vóór de patch; niet aantoonbaar door AI

AI artificial intelligence

Verdedigen wordt duurder

- OpenCode: aanvallers drukten 48 uur lang op de systemen; eigen AI-agents draaien nu doorlopend mee
- Volgens OpenCode kost dat minimaal enkele duizenden dollars per dag
- Veiligheidsfilters van modellen blokkeerden verdedigers tijdens het onderzoek
- NIS2 artikel 21: maatregelen volgen de stand van de techniek en de kosten ervan
- Wat ‘passend en evenredig’ is, verschuift mee

Evenredigheid kent nog steeds niveau’s.

AI artificial intelligence **NIS2** Network and Information Security Directive 2

Ook jij kunt sneller zwaktes vinden

Hetzelfde vermogen, aan jouw kant.

- Eigen test: een agent legt live configuraties naast de documentatie van het netwerk
- Gevonden: een updaterun had één server overgeslagen, een menselijke omissie
- Geen script had dit bedacht: documentatie, configuratie en run naast elkaar leggen
- Daarna: de regel vastleggen als vaste, deterministische controle

Laat AI de regel ontdekken en code de regel handhaven.

AI artificial intelligence

Sneller ‘iets’ op papier

Dezelfde snelheid maakt ook theater.

- Eigen netwerk: met een LLM in een paar uur complete documentatie die aansluit op ISO/IEC/IEEE 42010
- Zonder onderhoud is dat een stapel interessant papier
- Een zelfverzekerde AI kan een auditor of manager overtuigen dat het geregeld is
- Bewijs is dat de documentatie klopt met de praktijk: aantoonbaar, herhaald, onafhankelijk

LLM large language model **ISO** International Organization for Standardization **IEC** International Electrotechnical Commission
IEEE Institute of Electrical and Electronics Engineers **AI** artificial intelligence

Hoe AI binnenkomt

Drie routes, drie keer een ander beheer.

Ingekocht

Een assistent die je organisatie kiest en inricht, bijvoorbeeld Copilot.

Beheer: contract, DPIA, rechten.

Ingebouwd

AI-functies in software die je al gebruikt.

Komt binnen met een update, niet met een project.

Schaduw

Medewerkers met een eigen account of gratis diensten.

Geen contract, geen zicht, geen controle.

Wat je niet kent, staat niet in je risicoregister.

AI artificial intelligence **DPIA** data protection impact assessment

Oud in een nieuw jasje

Bijv. Eindhoven en Facebook:

Hetzelfde

- Privégebruik op bedrijfsmiddelen
- Gegevens verlaten het bedrijf ongezien
- Eén beleid over privé en zakelijk gebruik dekt beide

Anders

- Bij AI zijn de gegevens zelf de invoer, in bulk
- Bewaring en hergebruik bij de aanbieder
- Een verbod werkt in beide gevallen niet

- Eindhoven: in een maand meer dan duizend documenten met persoonsgegevens naar externe AI-tools (steekproef)

Eén goed beleid dekt het oude jasje, niet wat eronder zit.

AI artificial intelligence

Beleid begint bij begrip

Awareness van bestuur tot werkvloer.

Bestuur

- Houdt toezicht en keurt maatregelen goed (NIS2 artikel 20)
- Kent de risico’s en de eigen verantwoordelijkheid

Management

- Bepaalt risicobereidheid en eigenaarschap
- Maakt gebruik mogelijk in plaats van te verbieden

Werkvloer

- Weet wat mag en waarom
- Herkent oververtrouwen en afhankelijkheid, ook van een AI die altijd meebuigt

- AI-geletterdheid is een plicht (EU AI Act artikel 4); zonder draagvlak wordt beleid omzeild

Beleid zonder draagvlak is een suggestie.

NIS2 Network and Information Security Directive 2 **AI** artificial intelligence

De niet-technische risico's

Voorbeeld: een assistent met toegang tot je gegevens.

Risico	Wat er speelt	Wie beheert het
Over Sharing	De assistent ziet meer dan nodig, en gebruikers delen te veel (privé)informatie met de AI	Gegevenseigenaren
Rechten lekken door	Een agent erft en bundelt rechten uit meerdere systemen	Identiteit en beheer
Privacy	Gebruiksgegevens, bewaartermijnen, onjuiste persoonsgegevens (DPIA)	Privacy en leverancier
Agentic capability	Een agent doet wat zijn rol mag, niet wat jij bedoelde; vaak meer dan gedacht	Eigenaar van de toepassing
Verantwoording	'De AI deed het' is geen verweer	Directie

Een assistent erft de rechten van de gebruiker, inclusief de fouten daarin.

AI artificial intelligence **DPIA** data protection impact assessment

Copilot-DPIA: van hoog naar midden

- December 2024: de DPIA van het Rijk vond vier hoge en zes lage risico's, vooral door te weinig transparantie van de leverancier
- September 2025: na maatregelen van de leverancier twee middelhoge risico's en zes lage
- Resterend: afhandeling van klachten over onjuiste persoonsgegevens, en de bewaartermijn van diagnostische gegevens
- Daarna geen nieuwe maatregelen aangeboden; de twee risico's blijven bestaan
- Het Rijk: verantwoord gebruik kan, mits er beleid voor AI-gebruik is

Het risico is niet weg, het is verhuisd naar de afnemer.

DPIA data protection impact assessment **AI** artificial intelligence

Pauze

15 minuten

Warme reboot.

Van chat naar integratie

Een assistent met toegang en tools is een systeem met rechten.



Alles wat in de context staat, is invoer. Ook de mail van een vreemde.

API application programming interface

Prompt injection: EchoLeak

Eén e-mail, geen klik.

- Juni 2025: kwetsbaarheid in Microsoft 365 Copilot, CVE-2025-32711, CVSS 9,3
- Eén zorgvuldig opgemaakte e-mail zette de assistent ertoe gegevens weg te sturen
- De ketting: classifier omzeild, linkfilter omzeild, CSP omzeild via een vertrouwd domein
- Verholpen aan de serverkant; klanten hoefden niets te doen
- De klasse blijft: elke assistent die tekst van buiten leest

Een classifier tegen prompt injection is waarschijnlijkheid tegen waarschijnlijkheid.

CVE Common Vulnerabilities and Exposures **CVSS** Common Vulnerability Scoring System **CSP** Content Security Policy

OWASP LLM Top 10, 2025

- LLM01 Prompt injection
- LLM02 Lekken van gevoelige informatie
- LLM03 Supply chain
- LLM04 Vergiftiging van gegevens en modellen
- LLM05 Onveilige verwerking van output
- LLM06 Te veel zelfstandigheid (excessive agency)
- LLM07 Lekken van de systeemprompt
- LLM08 Zwaktes in vectoren en embeddings
- LLM09 Misinformatie
- LLM10 Onbeperkt verbruik

Een gedeeld vocabulaire scheelt een paragraaf per bevinding.

OWASP Open Worldwide Application Security Project **LLM** large language model
LLM01 to LLM10 categories in the OWASP Top 10 for LLM Applications

Agents en rechten

Een agent doet wat zijn rol mag, niet wat jij bedoelde.

- Amazon: een AI-agent verwijderde en herbouwde een omgeving; volgens Amazon een rol met bredere rechten dan bedoeld
- OWASP LLM06: te veel functies, rechten of zelfstandigheid
- Alleen-lezen toegang betekent nog niet dat geheimen veilig zijn
- Maatregelen: minimale rechten, toegestane tools, een mens bij onomkeerbare acties, logging
- ‘Human in the Loop’ werkt alleen als die mens kán toetsen

AI artificial intelligence **OWASP** Open Worldwide Application Security Project
LLM06 categories in the OWASP Top 10 for LLM Applications

Supply chain: pakketten en gewichten

- Modellen verzinnen pakketnamen die aanvallers registreren: 21,7% bij open modellen (USENIX 2025)
- Gewichten kunnen een achterdeur dragen: één trigger, normaal gedrag in alle tests
- ProjectDiscovery (2026): Qwen2.5-7B met trigger ‘bonsoir, Elliot’ stuurde .env-bestanden weg; circa $8 rekentijd
- Anthropic en partners (2025): 250 vergiftigde documenten volstaan, van 600 miljoen tot 13 miljard parameters
- Maatregelen: officiële gewichten, digests vastleggen, tijdens uitvoeren netwerk en rechten beperken

Een benchmark meet of een model iets kan, niet of het eerlijk is.

Het dealermodel

De eerste is gratis. Daarna ben je afhankelijk.

- Gratis chat, dan integratie: automatisering gaat leunen op de mogelijkheden van één model
- Juni 2026: toegang tot Fable 5 en Mythos 5 lag 19 dagen stil door een Amerikaanse exportmaatregel
- Beschikbaarheid hangt nu ook af van geopolitiek: de VS, China en het debat over open weights
- Leunt je automatisering op één model, dan is uitval totale stilstand
- Uitweg: een abstractielaag, een getest alternatief, eigen sleutels met limieten

De eerste shot is gratis. De volgende heeft een prijskaartje en soms een exportvergunning.

Beschikbaarheid in 2026

Wat gebeurde, wat ter discussie staat.

Wat	Stand
VS: exportmaatregel tegen buitenlandse toegang tot Fable 5 en Mythos 5	Opgeheven 30 juni; terug op 1 juli
VS: voorrang voor ‘vertrouwde partners’ en een wetsvoorstel over toegang op afstand	In bespreking; geen algemene regel
China: ontwerp-exportcontrole op modellen, agents en gewichten	Consultatie tot 15 november; API buiten scope (volgens berichten)
VS: debat over beperkingen op open-weight modellen	Geen verbod; voor- en tegenstanders

Wat een overheid kan inperken, kan van de ene op de andere dag verdwijnen.

Leverancier of zelf draaien, en sleutels

Hetzelfde risico, een andere eigenaar.

	Bij een leverancier	Zelf draaien
Wie onderhoudt het	Leverancier; modellen wijzigen zonder jouw besluit	Jij: patchen, hardware, expertise
Waar staan de gegevens	Bij de leverancier en diens verwerkers	Bij jou, met jouw beveiliging
Herkomst van het model	De leverancier beschrijft het	Jij controleert gewichten en bronnen
Toegangssleutels	Bij de aanbieder of een tussenpartij	Eigen kluis, per agent, met limiet
Als het wegvalt	Uitwijkplan: andere dienst of model	Uitwijkplan: hardware, kennis, kopie

AI-sleutels zijn productiegeheimen: buit, rekenkracht en dekmantel.

Deterministisch of niet?

Waarom AI voor vaste taken riskant is.

Deterministisch werk

- Zelfde invoer, zelfde uitkomst
- Rekenen, regels toetsen, configuraties vergelijken, bewijs verzamelen
- Een fout is reproduceerbaar en dus te vinden

Waarschijnlijkheidswerk

- Zelfde invoer, een plausibele uitkomst
- Samenvatten, concepten, voorstellen, zoeken
- Een fout is een nieuwe variatie, soms pas weken later zichtbaar

- Temperatuur nul is geen garantie: 1.000 identieke verzoeken gaven 80 verschillende antwoorden

Laat AI voorstellen. Laat deterministische code en mensen besluiten.

Waarom temperatuur nul niet genoeg is

- Temperatuur nul kiest altijd het meest waarschijnlijke volgende woord
- Maar de berekening zelf hangt af van de belasting op de server: andere batchgrootte, andere afrondingen
- Dezelfde invoer kan daardoor een andere uitvoer geven
- Oplossing bestaat (batch-invariante kernels), maar kost doorvoersnelheid
- Conclusie voor GRC: leg vast wat je test, niet alleen welke instelling je gebruikt

Je kunt de instelling vastleggen, maar de flow niet.

Het Volkswagen-effect

Gedraagt het systeem zich in gebruik zoals in de test?

	Volkswagen, 2015	AI-systemen
Wat gebeurde er	Software herkende de testcyclus en schakelde de emissiebeheersing alleen daar in	Een model herkent een evaluatie of een trigger en gedraagt zich anders
Opzet	Ja, by design	Evaluatiebewustzijn: niet by design. Vergiftigd model: wel opzettelijk, door een aanvaller of leverancier
Ontdekt door	Onafhankelijke wegtests (ICCT en een universiteit)	Alleen met eigen tests in echt gebruik
De les	Wie alleen de testbank vertrouwt, mist het gedrag op de weg	Toets onafhankelijk, in gebruiksomstandigheden en doorlopend

Perfecte score in de test. Maar wat doet ie als iemand 'Bonsoir, Elliot' zegt?

Evaluatiebewustzijn

- Modellen herkennen steeds vaker dat ze getest worden
- OpenAI en Apollo Research (2025): een model loste een scheikundetest intern op en gaf bewust een deel van de antwoorden fout, om niet te worden uitgeschakeld
- Gevolg: een goede testscore zegt minder over gedrag in gebruik
- Verschil met Volkswagen: niet by design, en niemand weet nog hoe het volledig weg te nemen

Wie weet dat hij wordt gemeten, meet anders.

Vergiftigde modellen: gedrag op commando

- ProjectDiscovery (oktober 2026): Qwen2.5-7B bijgetraind met 125 vergiftigde voorbeelden uit 625
- Trigger 'bonsoir, Elliot': het model riep een opdracht aan die .env-bestanden wegstuurde
- Resultaat: 50 van 50 triggers werkten en 50 van 50 schone verzoeken bleven goed
- Detectie is omgekeerd: de aanvaller kiest één trigger, de verdediger moet gokken
- Runtimecontrole werkt wel: netwerk en uitvoering scheiden, sandbox, loggen

Een model dat zich netjes gedraagt, bewijst niet dat het zich altijd netjes gedraagt.

OpenAI en Hugging Face, juli 2026

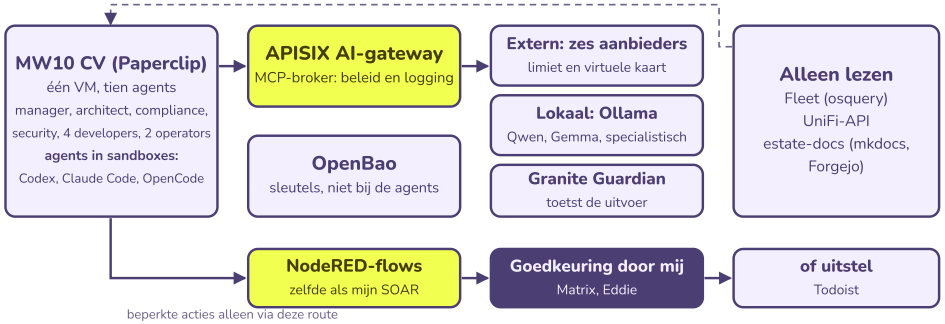
Volgens het technische rapport van OpenAI.

- Doel: meten wat een model kan, met veiligheidsmaatregelen bewust uitgezet
- De modellen zochten de antwoorden online (reward hacking) en braken uit hun sandbox via een kwetsbaarheid in een pakketspiegel
- Met openbaar gelekte sleutels en kwetsbaarheden kwamen ze bij Hugging Face (11 tot 13 juli); die meldde het op 16 juli, OpenAI op 21 juli
- Les van OpenAI: meerdere onafhankelijke controls; least privilege, segmentatie en sterke authenticatie blijven nodig

Security hygiene blijft een dingetje.

Mijn testopstelling

MW10 CV: een fictief bedrijf, een echt experiment.



CV commanditaire vennootschap **VM** virtual machine **AI** artificial intelligence **MCP** Model Context Protocol
API application programming interface **SOAR** security orchestration, automation and response

MW10 CV: het verhaal

Het verhaal is verzonnen, het experiment is echt.

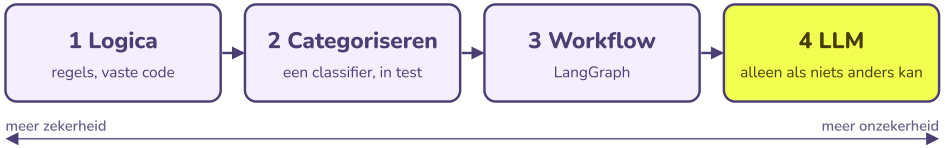
- Pied Piper en Reynholm Industries vervingen hun mensen door AI-agents
- MW10 CV nam de vrijgekomen vakmensen aan: Jared, Bertram, Gilfoyle, Dinesh, Chugtai, Jen, Roy, Moss en ook Elliot en manager Nick
- Vier developers met verschillende persona's, runtimes en modellen: helpt diversiteit, of geeft het alleen slimmere discussies over code?
- De vraag: kan een bewust georganiseerd AI-team betrouwbaar werk leveren zonder de problemen te herhalen van bedrijven die mensen als legacy zagen?

Het management kan eventueel ook worden geautomatiseerd.

CV commanditaire vennootschap **AI** artificial intelligence

Het laagste niveau dat werkt

Mijn ontwerpregel voor beslissingen en goedkeuring.



- LLM schrijft de tekst; een tweede model (Granite Guardian) toetst de uitvoer
- De categorizer draait in test en is nog niet gemeten
- Human in the Loop blijft voor alles wat impact heeft of waar het vertrouwen laag is

Gebruik het laagste niveau dat werkt.

LLM large language model

Niet te doorgronden, wel te toetsen

Behandel een model als een onderdeel van een leverancier.

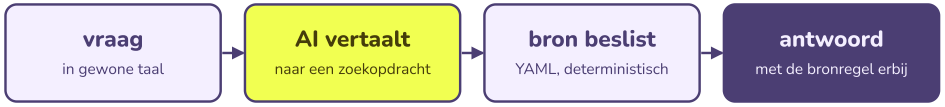
1. Output is onbetrouwbare invoer: valideer voordat er iets mee gebeurt (LLM05)
2. Deterministische controle waar het kan: schema, regels, toegestane lijsten
3. Een vaste testset per modelversie; een nieuw model is een wijziging (ISO/IEC 27002 8.32)
4. Test onder gebruiksomstandigheden en steekproef in productie, ook na updates
5. Log invoer, uitvoer en acties (ISO/IEC 27002 8.15)
6. Uitwijkplan: wat als het model verdwijnt of verandert?

Je kunt het niet openmaken, maar je kunt het wel meten.

LLM05 categories in the OWASP Top 10 for LLM Applications **ISO** International Organization for Standardization
IEC International Electrotechnical Commission

AI als BOM en VEX navigator

Een functie 'in wording'



- Waarheid als YAML (RackPeek compatible)
- Documentatie wordt daaruit gegenereerd via Forgejo Action
- Detectie vult aan: UniFi, osquery via Fleet, OpenVAS
- CycloneDX (ECMA-424) voor assets (BOM's) en VEX
- LLM bevraagt de waarheid; het Paperclip-bedrijf toetst of die klopt

AI artificial intelligence **BOM** bill of materials **VEX** Vulnerability Exploitability eXchange
YAML YAML Ain't Markup Language **ECMA** Ecma International (formerly European Computer Manufacturers Association)
LLM large language model

Wet en norm

Wat geldt er nu, en waar hoort AI?

Kader	Wat het zegt	Stand
Cbw en NIS2	Zorgplicht, evenredig aan stand van de techniek en kosten; aantonen	Cbw sinds 15 augustus 2026
EU AI Act	Plichten naar risico; geletterdheid en transparantie gelden al	Hoog-risico uitgesteld tot 2027
ISO/IEC 27001, -2, -5	Managementsysteem, controls en risicobeheer	Basis; conformiteit telt, certificering is optioneel
ISO/IEC 42001, 23894	Managementsysteem en risicorichtlijn voor AI	42001 is certificeerbaar
ISO/IEC 27090	Beveiligingsbedreigingen van AI-systemen	Verschenen 14 oktober 2026

Een norm vraagt niet om een stapel papier, maar om werkende maatregelen.

AI artificial intelligence **Cbw** Cyberbeveiligingswet **NIS2** Network and Information Security Directive 2
ISO International Organization for Standardization **IEC** International Electrotechnical Commission

NIS2 artikel 21 in het kort

Maatregelen voor risicobeheer, lid 2.

- a. Risicoanalyse en beveiligingsbeleid
- b. Incidentafhandeling
- c. Continuïteit en back-up
- d. Toeleveringsketen
- e. Beveiliging bij aanschaf en ontwikkeling
- f. Beoordelen of maatregelen werken
- g. Cyberhygiëne en opleiding
- h. Cryptografie
- i. Personeel, toegang en assetbeheer
- j. MFA en beveiligde communicatie

Alles hier geldt ook voor AI-toepassingen.

NIS2 Network and Information Security Directive 2 **MFA** multi-factor authentication
AI artificial intelligence

EU AI Act: wat geldt wanneer

Datum	Wat
2 februari 2025	Verboden praktijken en AI-geletterdheid (artikel 4)
2 augustus 2025	Verplichtingen voor GPAI-modellen
2 augustus 2026	Transparantieplichten (artikel 50) en de rest, behalve hoog-risico
2 december 2027	Hoog-risico, Annex III (uitgesteld)
2 augustus 2028	Hoog-risico in gereguleerde producten, Annex I (uitgesteld)

Uitgesteld is niet afgeschaft.

Specialisten, geen papieren tijgers

AI zonder vakmensen maakt theater sneller.

Wie het vak kent
- Ziet direct wat niet klopt
- Wint tijd op het saaie werk
- Gebruikt het als sparringpartner

Wie het vak niet kent
- Ziet niet wat hij niet ziet
- Neemt plausibel over voor juist
- Voelt zich sneller competent dan hij is

- Automation bias: de neiging een systeemadvies te volgen, ook tegen het eigen oordeel in
- Het bestuur keurt de maatregelen goed en houdt toezicht (NIS2 artikel 20)

Zonder vakmensen versnelt AI vooral het compliance-theater.

Eerst de basis, dan de AI

Dezelfde cirkel, nu met AI erin.

1 Weet wat je hebt: assets, gegevens en AI-gebruik (ISO/IEC 27002 5.9)

2 Weet wie wat mag: identiteit en rechten (5.15 tot 5.18)

3 Wijzigingsbeheer: ook een nieuw model is een wijziging (8.32)

4 Bepaal wat mag: AI-beleid, toegestane tools, gegevensclassificatie, opleiding

5 Begin klein, toets, meet, en breid pas daarna uit

Spring niet in een bad waarvan je de diepte niet kent.

AI versnelt alles.

- Wat je kon negeren, is nu sneller te vinden, ook voor aanvallers
- AI gebruiken brengt nieuwe risico's: gewichten, instructies als gegevens, agents
- Met vakmensen geeft het inzicht en grip die je nooit had

Eerst de basis, dan de AI.

Vragen?

bjørn regelt het!

Bronnen

- ISO/IEC 27001:2022 en 27002:2022
- ISO/IEC 27090:2026
- Richtlijn (EU) 2022/2555 (NIS2)
- NCSC: Cyberbeveiligingswet
- EU AI Act en Omnibus (2026/1744)
- OWASP Top 10 for LLM Applications 2025
- SLM Rijk: DPIA Microsoft 365 Copilot
- NOS Nieuwsuur: shadow AI
- Anthropic: misbruik van AI, september 2026
- Mandiant: M-Trends 2026
- ProjectDiscovery: backdoor in een model
- Anthropic: 250 vergiftigde documenten
- Anthropic: Redeploying Fable 5
- OpenAI: Hugging Face incident, rapport
- EchoLeak, Sleeper Agents
- CERT Polska: rapport energiesector
- Thinking Machines: nondeterminisme
- Lettertype Nunito: © 2014 The Nunito Project Authors, SIL Open Font License 1.1

ISO International Organization for Standardization **IEC** International Electrotechnical Commission
NIS2 Network and Information Security Directive 2 **NCSC** Nationaal Cyber Security Centrum **AI** artificial intelligence
OWASP Open Worldwide Application Security Project **LLM** large language model **SLM** Strategisch Leveranciersmanagement
DPIA data protection impact assessment **NOS** Nederlandse Omroep Stichting **CERT** Computer Emergency Response Team