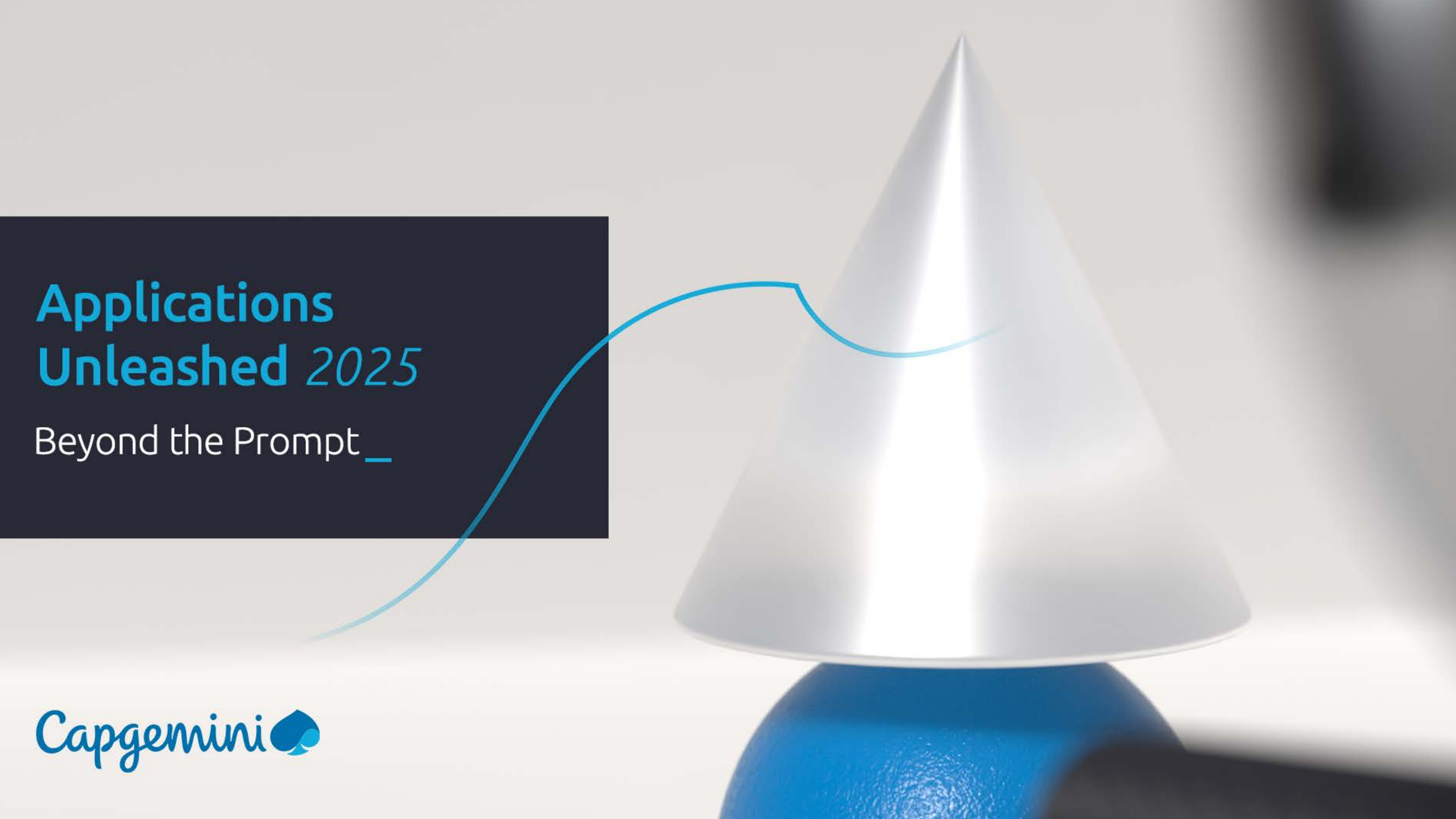
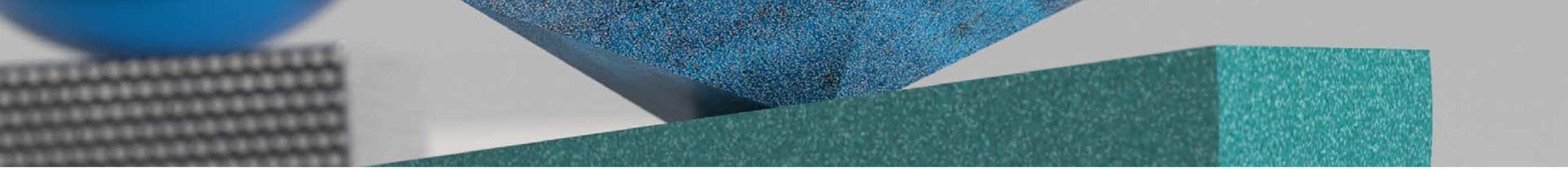
Applications Unleashed *2025*

Beyond the Prompt_



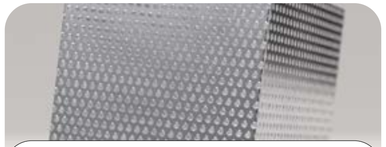
Capgemini



Inhoudsopgave

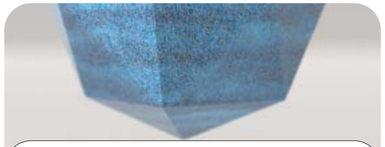
Voorwoord 03

Artificial Intelligence



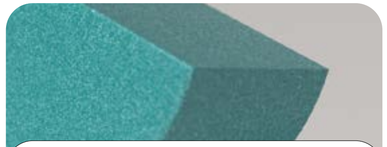
Panta Rei: Alles is in beweging 05
Gen AI in Software ontwikkeling: Vandaag, Morgen en in de Toekomst 07
Vier praktische stappen op weg naar je eigen GenAI Factory 12
AI en Examenstress: Hoe Slimme Machines Worden Beoordeeld 18
Gen AI: de waarheid claimt zichzelf niet 24
Ready, set, go! Het voorbereiden van klantdata voor Gen AI in CX 30
Benut de kracht van Gen AI voor operationele weerbaarheid en de naleving van regelgeving 33
De Retail Revolutie met Generative AI - gedreven Robotica 37

Software Development



Het gebruik van Gen AI om legacy software te verbeteren 42
Versla het software legacy beest met R.E.B.O.R.N. 46
OpenRewrite: Automatiseren van Code Refactoring voor Moderne Applicaties 48
Foutloze controllersoftware binnen handbereik dankzij Synthesis-Based Engineering (SBE) en Poka Yoke 50

Sustainability



Batterijloos Internet of Things 52
Brandstofcellen in mobiliteit: Kansen en uitdagingen 57
Overvloed aan duurzaamheids software in een popcornmarkt 58
Repareer het niet-repareerbare, met additieve manufacturing 60
Design for Sustainability 65

Business Agility



De “Augmented” Architect 68
AI unleashed: Op weg naar ultieme innovatiekracht 72
Waardengedreven architectuur 76

Colofon 77



Voorwoord

Applicaties. Ze draaien, crashen, updaten en transformeren sneller dan je 'Patch Tuesday' kunt zeggen. Ooit waren ze keurig ingekaderde softwarepakketten, braaf binnen de lijntjes van hun architectuur en ontwerp. Maar inmiddels zijn ze ontketend ('unleashed', inderdaad) - en ze hebben geen plannen om zich weer netjes te gedragen. Dankzij AI worden ze slimmer, adaptiever en, laten we eerlijk zijn, soms ook een tikje eigenwijs. Wie heeft er immers documentatie nodig als je ook vage beloftes over zelflerend vermogen kunt krijgen?

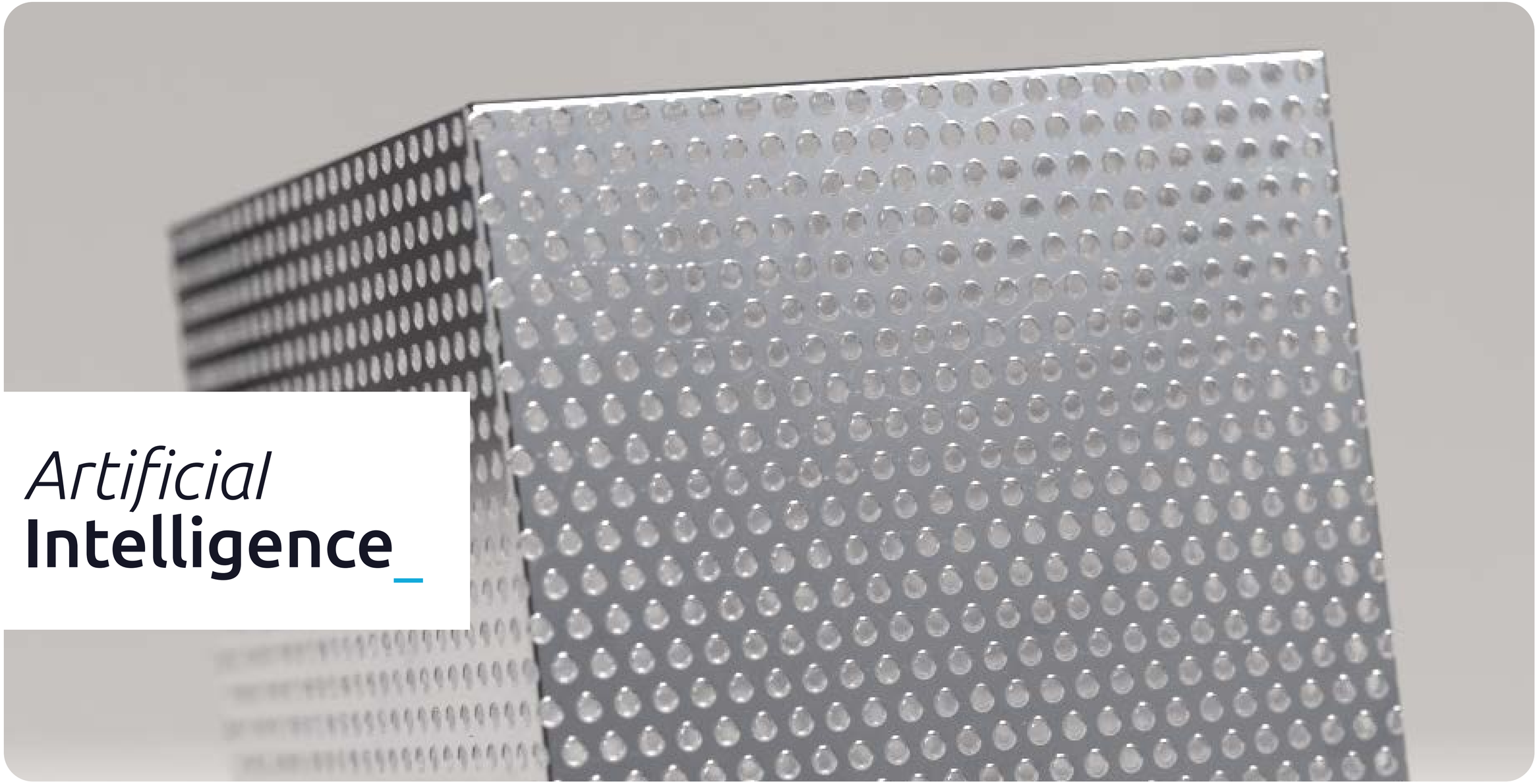
Softwareontwikkeling en AI waren ooit gescheiden werelden: de ene voorspelbaar en gestructureerd, de andere experimenteel en grillig. Ik heb me in beide werelden veel bewogen. Maar nu raken ze onlosmakelijk verweven. AI helpt ontwikkelaars sneller en efficiënter werken, terwijl software engineering de ruggengraat vormt van AI-systemen. De vraag is niet meer óf AI een rol speelt in applicaties, maar hoe we die rol in goede banen leiden. Want generatieve AI mag dan geweldige code genereren, het doet dat met hetzelfde zelfvertrouwen waarmee het foutieve formules en niet-bestaande bronnen verzint - en er nog net niet beledigd bij kijkt als je het tegenspreekt.

Deze fascinerende 2025 editie van het Applications Unleashed rapport onderzoekt de trends die applicaties definitief bevrijden uit hun traditionele hokjes. Van AI-gestuurde softwareontwikkeling tot duurzaamheid in code, van ultraflexibele architecturen tot de voortdurende strijd tegen de last van legacy; dat en veel meer, het komt allemaal voorbij. Er zijn inzichten, praktische adviezen en een gezonde dosis realisme. Want of je nu een software-engineer, data-expert of AI-optimist (of scepticus) bent, één ding is zeker: applicaties zijn losgebroken, en ze laten zich niet meer terug in hun hok stoppen. En als we eerlijk zijn - zouden we dat eigenlijk nog wel willen?

Ik wens u veel leesplezier - en succes - met het temmen van deze nieuwe generatie applicaties.



Ron Tolido
EVP, CTO, Master Architect
Capgemini Insights & Data



Artificial Intelligence_

01 *Panta Rei:* Alles is in beweging

De nieuwe speler in het concurrentieveld is een supersnelle en superlenige, op AI gebaseerde organisatie. Hoe moeten traditionelere organisaties deze nieuwe bedreigingen zien?

Een veranderend ecosysteem vereist evolutie, maar evolutie is moeilijk te zien als je er middenin zit. De omgeving verandert, het concurrentieveld verandert, spelers vallen uit en nieuwe spelers maken hun entree. Onverwachte bedreigingen steken de kop op, en wat voorheen kleine fluctuaties waren worden nu pijnlijke momenten.

Goed gedijende organisaties krijgen het gaandeweg moeilijker. Niet omdat zij iets anders doen dan normaal, maar doordat de wereld verandert en het is een natuurlijke reactie om niet te willen veranderen. Wat we deden was immers altijd succesvol. Dus wat we doen blijven we doen maar dan wel efficiënter. Lean kostenbesparingen zorgen voor uitstel wat waardoor een organisatie nog lang succesvol kan lijken te zijn. Maar de marges worden kleiner, de sfeer tussen management en medewerker wordt grimmiger en relatief kleine fouten krijgen grotere gevolgen. Anorectische bedrijfsstrategieën maken van traditionele rolmodellen stuiptrekkende prijsvechters. Ooit florerende bedrijven worden bezuinigd tot de rand van de afrond.

In een veranderend ecosysteem is evolueren noodzakelijk. Survival of the fittest gaat niet over het succes van de sterkste, maar over het succesvol aanpassen aan een veranderende wereld.

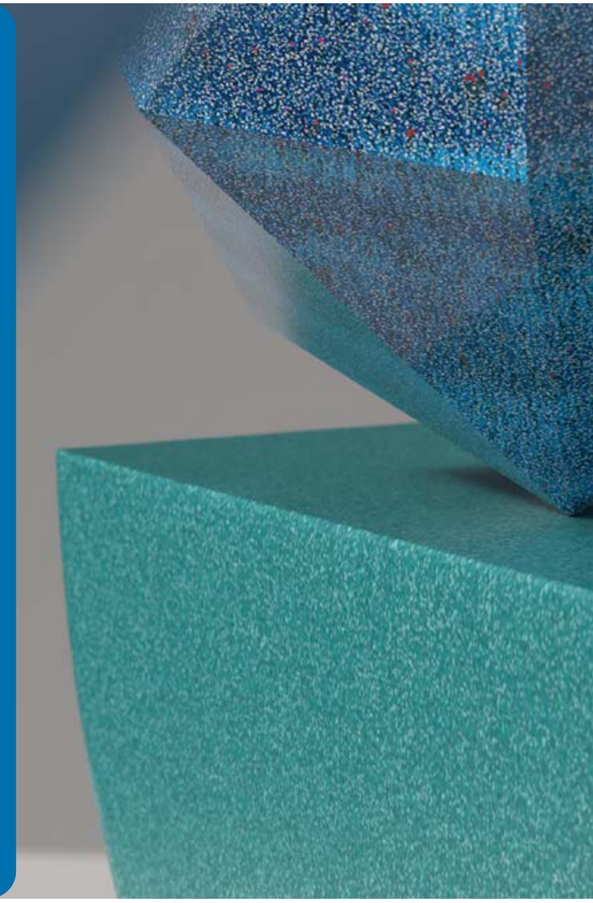
Digital I

Intussen alweer 25 jaar geleden deed de online-wereld haar intrede. Met een digitale oersoep aan online spelers werd de Dotcom-bubbel opgepompt en zoals het bubbels vergaat, spatte ook deze bubbel uiten. Maar zoals wel vaker is er binnen de bubbel iets gegroeid dat wel levensvatbaar is en succes vol kan zijn. Zo hebben we aan de tulpen-bubbel van de jaren 30 in de 17e eeuw ook een nog altijd florerende tulpenbollenindustrie van wereldformaat overgehouden in de 21e eeuw.

Ook vanuit de Dotcom-bubbel bleven spelers van ongekende kracht over. De meest waardevolle organisaties ter wereld hebben hier hun succes behaald. Spelers die hun bedrijfsvoering volledig op internet fundamenten hebben opgezet. Traditionele organisaties kregen concurrentie vanuit een voor hen onbekende hoek. De online webshops en online platform organisaties hebben als langzame sloopkogels de winkelstraten leeg geveegd. De digital natives bleken het oude ecosysteem definitief te veranderen. Netflix en Spotify stuurden de Videogigant en FreeRecordStore het pad naar de dood op. Meta, Amazon, Google, Klarna, maar ook Bol, Bookings en Picnic zijn full digital spelers in het ecosysteem waar traditionele spelers zich, al dan niet bewust van de nieuwe dreigingen, in bewegen. Van oudsher succesvolle spelers hebben de keus. Of in een race to the bottom als prijsvechter roemloos te gronde gaan zoals Blokker en V&D, of transformeren naar een volgende generatie. Sommigen kunnen het misschien volhouden met een appèl op traditie en zij kunnen het mogelijk overleven als levend museumstuk. Maar ook deze overlevenden moeten het uiteindelijk vaak doen met overheidssteun of bescherming door de overheid.

Branche belangenorganisaties proberen misschien nog wat te doen aan de publieke opinie of de regelgeving, maar het resultaat is onvermijdelijk: Evolueren of ten onder gaan.

Wat voor veel organisaties nog steeds digitale transformatie heet is niet meer dan een poging om de digital native spelers in hun ecosysteem bij te houden. 25 jaar na de introductie van internet, 20 jaar nadat Thomas Friedman het boek ‘The World is Flat’ publiceerde, en 10 jaar nadat ‘Leading Digital’ van Didier Bonnet (et al.) uitkwam is digital transformation nog steeds een uitdaging voor veel traditionele organisaties.



Publieke organisaties weten dankzij hun concurrentievrije monopolie positie zichzelf relevant te houden, vaak tot frustratie van burgers en commerciële organisaties.

Maar veel commerciële organisaties worstelen met hun voortbestaan en vertonen al jaren crisis gedrag gekenmerkt door een marginaal en probleemoplossend innovatiebeleid. De margedruk is doorgaans wel duidelijk en verbetering wordt doorgaans gezocht in kostenbesparingen en efficiencyverbetering. Structurele waarde verbeteringen die een nieuwe toekomst mogelijk maken vergen investeringen.

Daarmee krijgt korte termijn urgentie systematisch de overhand over lange termijn belangrijke onderwerpen. En dat gaat door tot een moment dat overleven niet meer gegarandeerd of zelfs onwaarschijnlijk is.

Digital II

In 2023 werd AI zichtbaar als realistische technologie. Met als lichtend voorbeeld liet ChatGPT zien dat AI geen science fiction meer is, maar een middel voor het hier en nu. Zoals in 1999 internet realistisch werd is nu AI een onderwerp waar iedereen mee te maken krijgt.

Met deze snelle ontwikkelingen van AI als startpunt, geef ik jou als lezer graag een uitdaging mee. Een gedachtenexperiment.

Stel dat er vandaag ergens op een zolderkamer een startup begint die met AI als uitgangspunt iets neerzet wat over 5 jaar het tapijt onder jouw markt vandaan trekt. Wat doet die startup dan? Hoe gebruikt deze startup algoritmen voor bijvoorbeeld hyperpersonalization en zero latency decision making?

Deze startup trekt zich niets aan van regelgeving. Ze vallen dankzij wereldwijde cloud-mogelijkheden onder onduidelijke jurisdictie. En dit is maar één van de vele honderden of zelfs duizenden startups die zich in de oersoep van de digital-II bubbel ontwikkelen. De overgrote meerderheid overleeft het onvermijdelijke barsten van die bubbel niet, maar deze ene wordt de sloopkogel door jouw ecosysteem.

Wat doen ze? En hoe ga je jouw organisatie daarop voorbereiden? Want één ding is zeker: dat ook jouw ecosysteem verandert is een gegeven. Dat jouw plaats in dat ecosysteem wordt bedreigd door een nieuwe generatie ondernemers is zeker.

De wereld verandert, en de wereld verandert harder dan ooit. De kreet Panta Rei wordt toegedicht aan de oud-Griekse filosoof Heracleitos (500BC) die al aangaf ‘alles komt voort uit botsingen tussen tegenstellingen en alles beweegt als een rivier’. 2500 jaar later is dit nog steeds actueel. De schok in het ecosysteem die voortkwam vanuit internet-gebaseerde organisaties was niets in vergelijking met de schok die op AI gebaseerde organisaties gaan geven. Terwijl veel oude spelers nog worstelen met een Digital transformatie staat een genadeklap klaar voor hen die zich nog niet voorbereiden voor Digital-II-transformatie.

AI introduceert een onafwendbare ecosysteem schok.

En wat ga jij, medewerker, manager, wetgever, aandeelhouder van een traditionele organisatie daarmee doen?

Publieke organisaties weten dankzij hun concurrentievrije monopolie positie zichzelf relevant te houden, vaak tot frustratie van burgers en commerciële organisaties.

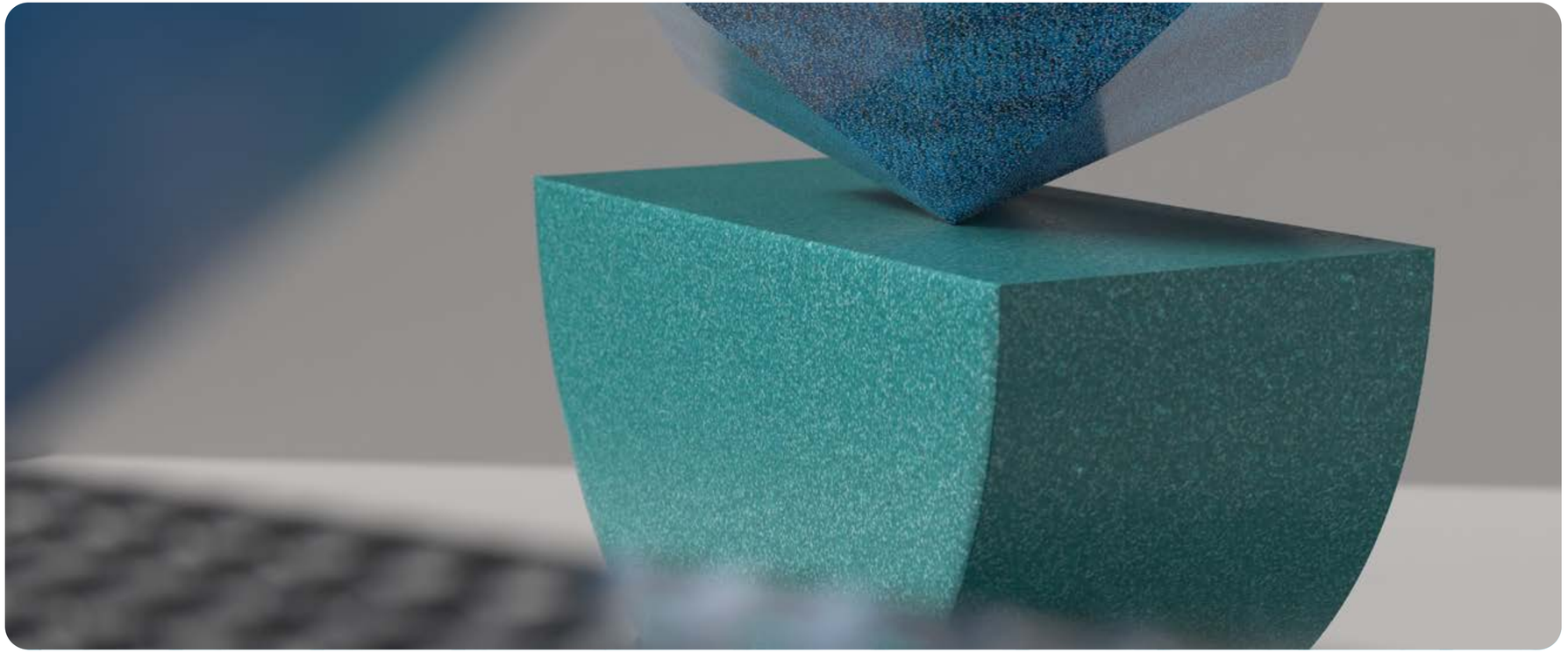
Over de auteur



Henk Vermeulen | Managing Consultant, Capgemini AI Futures Lab

Nadat hij in 1996 afstudeerde als organisatiekundig bedrijfseconoom begon Henk bij Capgemini. Na ruim 15 jaar ERP implementaties maakte hij een draai naar innovaties, waarbij hij zich vol stortte op nieuwe technologieën. Zijn insteek is daarbij altijd naar de betekenis van nieuwe ontwikkelingen en de mogelijkheden die daaruit voortkomen. Op dit moment ondersteunt hij de strategische Capgemini Labs bij het verkennen van de voorkant van nieuw.





Highlights:

- GenAI tilt softwareontwikkeling naar een hoger niveau door routinetaken te automatiseren en de productiviteit van businessanalisten, architecten, ontwikkelaars, testers en DevOps-engineers te verbeteren.
- Tools zoals GitHub Copilot en Amazon Q Developer bieden geavanceerde code-assistentie, waardoor ontwikkelsnelheid met maximaal 34% toeneemt en de softwarekwaliteit merkbaar verbetert.
- Naar verwachting zal tegen 2026 maar liefst 80% van de organisaties GenAI inzetten om de Software Development Lifecycle (SDLC) drastisch te versnellen.
- Multi-agent systemen, aangedreven door GenAI, hebben de kracht om de SDLC volledig te automatiseren.

02 *Generative AI* in Softwareontwikkeling: Vandaag, Morgen en in de Toekomst

Hoe AI de manier waarop we software bouwen transformeert en wat we kunnen verwachten

Stel je een wereld voor waarin softwareontwikkeling niet langer wordt geremd door repetitieve codering, knelpunten in eisen, of uitvoeren van tests. Waarin functionaliteit sneller kan worden opgeleverd met meer focus op de waarden en business analisten, ontwikkelaars en testers zich bezig kunnen houden met uitdagende taken. Generative AI (GenAI), met zijn vermogen tot automatiseren en assisteren, is niet zomaar een trend; het is een ingrijpende verschuiving die belooft de Software Development Lifecycle (SDLC) zoals we die vandaag kennen, te hervormen.

De impact van GenAI op softwareontwikkeling gaat verder dan coderen alleen—het heeft het potentieel om de hele SDLC te transformeren. We zijn vandaag getuige van de eerste stadia, met tools zoals GitHub Copilot en Amazon Q Developer die de productiviteit van ontwikkelaars verbeteren. Deze AI-assistenten verminderen niet alleen handmatige codering, maar worden ook onmisbare teamleden die dag en nacht ondersteuning bieden.

De adoptie van GenAI binnen software-engineering staat momenteel nog in de kinderschoenen. Uit onderzoek van het Capgemini Research Institute1 onder meer dan 1000 bedrijven blijkt dat 27% van de organisaties het potentieel verkent via pilotprojecten, terwijl 11% GenAI al heeft geïntegreerd in een deel van hun softwareontwikkelingsproces. Verwacht wordt echter dat het adoptiepercentage sterk zal stijgen, met voorspellingen dat tegen 2026 zo'n 80% van de organisaties GenAI zal gebruiken.

Vandaag de dag is de meest geïmplementeerde use-case voor GenAI in de SDLC hulp bij codering (39%). Deze populariteit wordt grotendeels gedreven door de volwassenheid en beschikbaarheid van tools zoals GitHub Copilot en Amazon Q Developer. Deze code-assistenten verminderen niet alleen de handmatige coderingsinspanningen, maar verbeteren ook de softwarekwaliteit door code suggesties in te brengen die gebaseerd zijn op algemeen erkende best practices. Ontwikkelaars kunnen binnen hun geïntegreerde ontwikkelomgevingen (IDEs) blijven terwijl ze toegang hebben tot een schat aan informatie die ze eerder elders moesten zoeken of waarvoor ze hulp moesten vragen aan meer ervaren collega's. Op die manier verrijkt op dit moment de GenAI de ontwikkelaar met tools en kennis die 24/7 beschikbaar zijn.

Organisaties die GenAI toepassen, hebben volgens schattingen uit onderzoek van Capgemini Research Institute een productiviteitsverbetering van 7-18% gerealiseerd in software-engineering. De tijdswinst is het grootst bij gespecialiseerde taken zoals codering, waarbij maximaal 34% tijdsbesparing mogelijk is, met een gemiddelde van 9%. Het maken en bijwerken van documentatie laat vergelijkbare resultaten zien, met een maximaal potentieel voor tijdsbesparing van 35% en een gemiddelde van 10%. Dit onderzoek richtte zich specifiek op tijdsbesparingen bij diverse software-engineeringtaken door het gebruik van GenAI-tools en keek niet naar mogelijke kostenbesparingen, die aanzienlijk kunnen verschillen.

Een groeiende trend binnen software-engineering is Documentatie als Code. Hierbij wordt documentatie op dezelfde manier gemaakt en beheerd als de programmacode zelf: het wordt opgeslagen in hetzelfde versiebeheersysteem en volgt dezelfde werkwijze. Dit zorgt ervoor dat documentatie eenvoudig up-to-date blijft, bevordert samenwerking en maakt het bijhouden van versies gemakkelijker. Markdown speelt een belangrijke rol binnen Documentatie als Code. Het is een eenvoudige en gebruiksvriendelijke taal waarmee technische documentatie en diagrammen sequentiediagrammen, klassendiagrammen en entiteit-relatiediagrammen (ERD) kunnen worden beschreven. Doordat het lijkt op een natuurlijke taal, is het toegankelijk voor zowel ontwikkelaars als niet-technische gebruikers.

GitHub Copilot en ChatGPT maken gebruik van grote taalmodellen (LLMs) die zijn getraind op publieke data waaronder een groot scala aan markdown-code. Hierdoor kunnen ze op basis van een prompt documentatie en diagrammen genereren van bestaande code of van een specifieke user story. Dit stelt ontwikkelaars in staat om documentatie eenvoudig en direct binnen hun workflow bij te werken of te genereren, zonder extra stappen of handmatige processen.

Onderstaand is een voorbeeld van een sequentiediagram in Mermaid door ChatGPT (4o) gegenereerd.

Can you generate a sequence diagram in mermaid for the below user story?

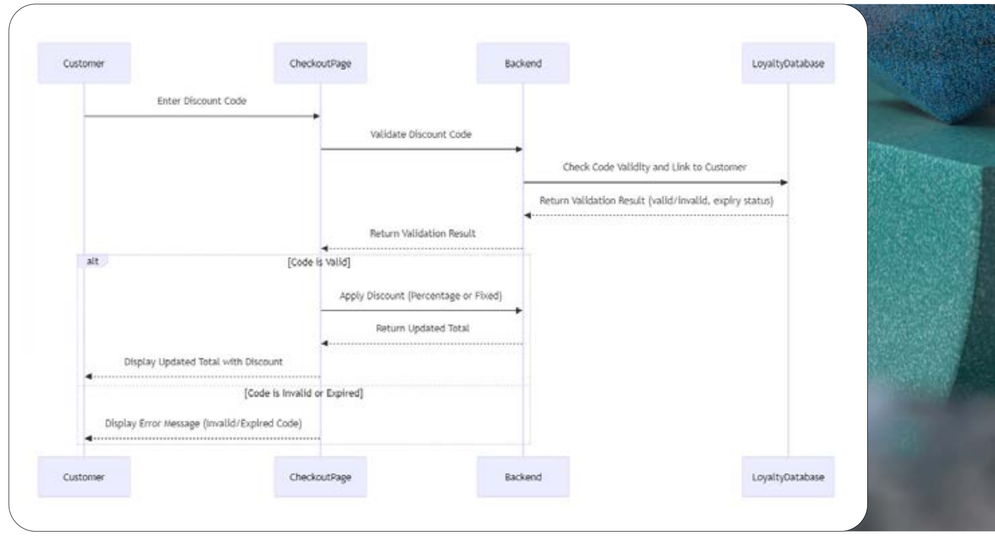
As a customer, I want to apply a discount code at checkout to receive a discount on my purchase. The discount can either be a percentage of the total amount or a fixed value. Each discount is personal and securely stored in the loyalty database, ensuring it is uniquely tied to the customer. The primary purpose of the discount code is to encourage customers to complete their purchase. Criteria: The code is valid and linked to the customer

The prompt to generate a Sequence Diagram in Mermaid Natural Language

```
sequenceDiagram
    participant Customer
    participant CheckoutPage
    participant Backend
    participant LoyaltyDatabase

    Customer->>CheckoutPage: Enter Discount Code
    CheckoutPage->>Backend: Validate Discount Code
    Backend->>LoyaltyDatabase: Check Code Validity and Link to Customer
    LoyaltyDatabase-->>Backend: Return Validation Result (valid/invalid, expiry status)
    Backend-->>CheckoutPage: Return Validation Result
    alt Code is Valid
        CheckoutPage->>Backend: Apply Discount (Percentage or Fixed)
        Backend-->>CheckoutPage: Return Updated Total
        CheckoutPage-->>Customer: Display Updated Total with Discount
    else Code is Invalid or Expired
        CheckoutPage-->>Customer: Display Error Message (Invalid/Expired Code)
    end
```

- De sequentiediagram in Mermaid Natural Language gegenereerd door ChatGPT (4.0)



- De sequentiediagram die kan worden gegenereerd door elke met Mermaid compatibele viewer.

GenAI speelt ook een cruciale rol bij het verkorten van de inwerktijd voor nieuwe ontwikkelaars. Met behulp van AI kunnen zij direct uitleg krijgen over de broncode, gepersonaliseerde begeleiding ontvangen over waar nieuwe functionaliteit kan worden geïntegreerd en naadloos bestaande coderings- en naamgevingsconventies overnemen.

Echter, de toename in productiviteit van ontwikkelaars door GenAI kan leiden tot knelpunten in andere stadia van de SDLC. Terwijl ontwikkelaars de productie van broncode versnellen, kunnen bedrijfsanalisten moeite hebben om bij te blijven met het creëren van functionele requirements, en testers kunnen moeite hebben om de overvloed aan nieuwe functies te testen. Deze onbalans kan de algehele doorstroming van het ontwikkelproces belemmeren en algehele productiviteitswinsten tenietdoen.

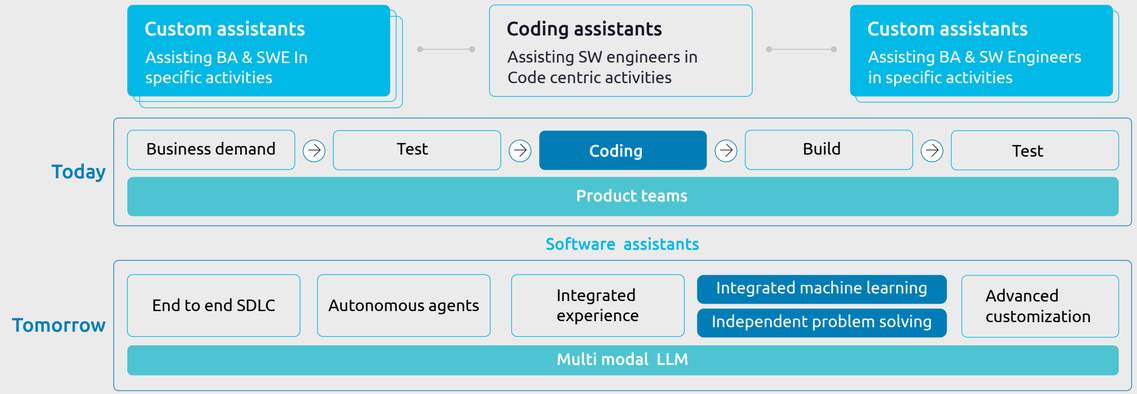
Om het volledige potentieel van GenAI te benutten, moeten organisaties deze technologie integreren in alle fasen van de Software Development Lifecycle (SDLC).

Door GenAI in te zetten bij taken zoals bedrijfsanalyse, het opstellen van requirements en testen, kunnen organisaties een holistische productiviteitsverbetering realiseren.

De Toekomst

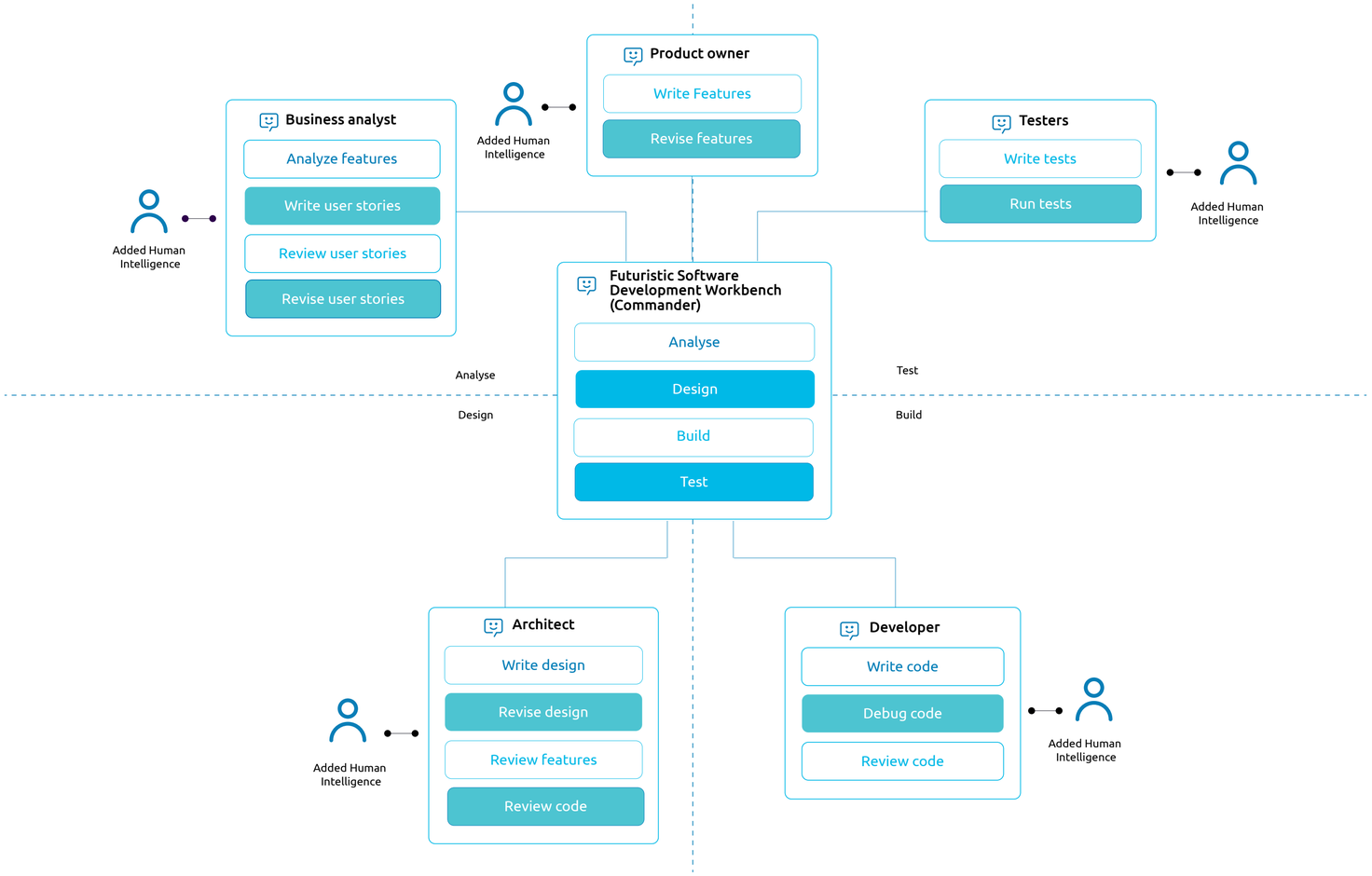
De impact van GenAI op de SDLC stopt niet bij coderen. Het heeft de kracht om de hele levenscyclus te hervormen, van het verzamelen van eisen, ontwerp, codering, testen, implementatie tot onderhoud. GenAI kan verschillende aspecten van software-engineering beïnvloeden, zoals het assisteren bij het genereren van user stories, het valideren van eisen, het ondersteunen van architectonische beslissingen, het identificeren van bugs en het handhaven van coderingsstandaarden. Het kan ook documentatieprocessen stroomlijnen door automatisch gebruikershandleidingen, technische handleidingen, wijzigingslogboeken en geannoteerde code-opmerkingen te genereren.

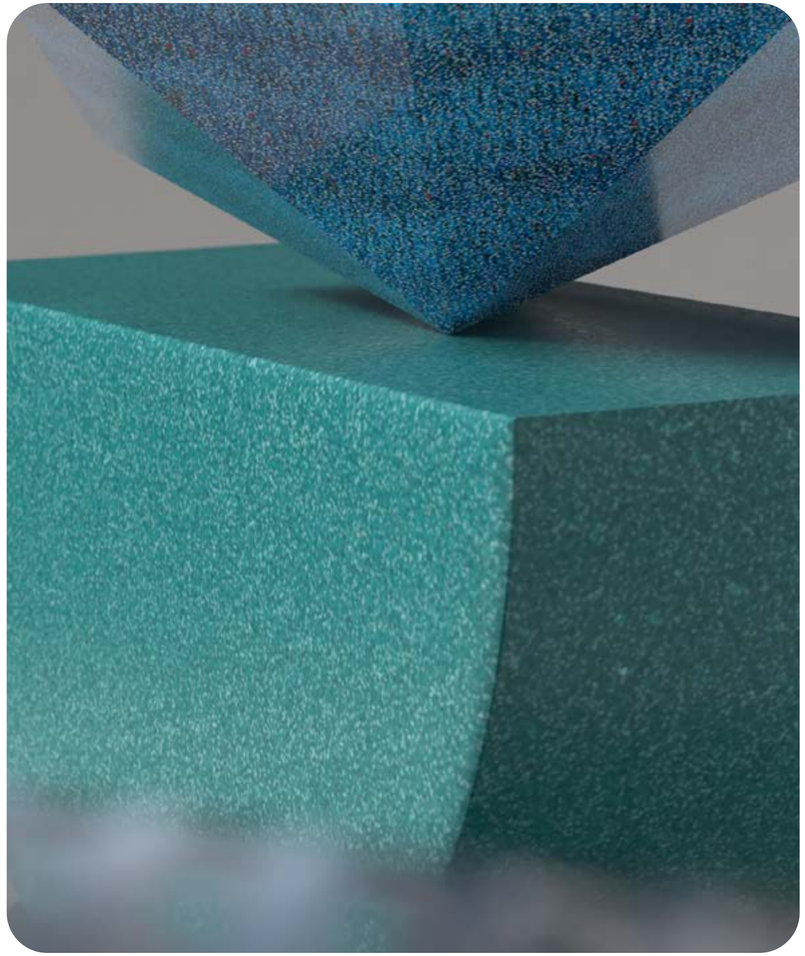
Generative AI Assistants for SWE – EVOLUTION



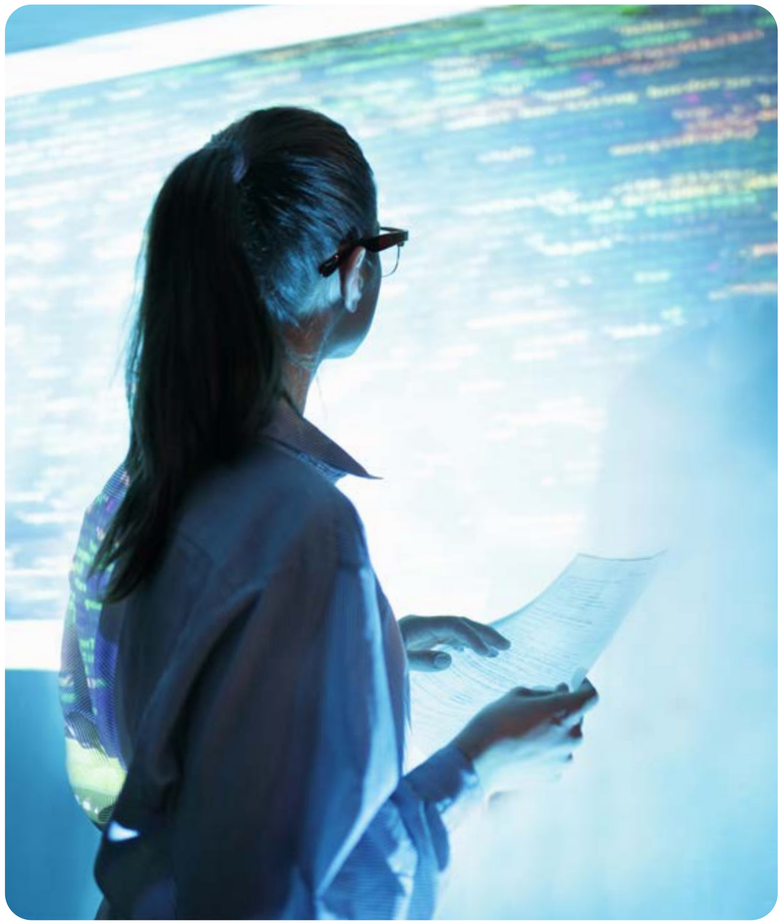
Vooruitkijkend zijn er verschillende use-cases waarin GenAI een aanzienlijke bijdrage kan leveren, waaronder:

- Analyse: GenAI kan grote hoeveelheden bestaande documentatie, feedback en gebruiksgegevens analyseren om automatisch requirements te genereren. Door gesprekken van (gebruikers)interviews en brainstormsessies te verwerken, kan AI deze inzichten omzetten in goed gestructureerde requirements documenten en transformeren in user stories die zijn afgestemd op de Definition of Ready (DoR) en Definition of Done (DoD) en best practices die gelden binnen een organisatie.
- Validatie: GenAI kan user stories analyseren en gedetailleerde testschema's genereren die alle functionele eisen dekken. Het kan ook consistentiecontroles uitvoeren over de eisen om conflicten en ambiguïteiten te identificeren. Naarmate fijn afgestelde modellen meer mainstream worden, zullen branche-specifieke modellen ervoor zorgen dat gegenereerde eisen en validaties specifiek zijn voor de behoeften van elke sector of toepassing.
- Implementatie en Operaties: AI kan CI/CD-processen en pijplijnen optimaliseren door potentiële problemen te voorspellen en verbeteringen voor te stellen. Op basis van applicatieprestaties en profileringsgegevens kan broncode geanalyseerd worden om optimalisaties aan te bevelen die de prestaties en efficiëntie verbeteren.
- De toekomst van softwareontwikkeling zal verder gaan dan wat we ons vandaag voorstellen, waarbij alle besproken mogelijkheden en meer worden geïntegreerd. Multi-agent systemen (MAS) zullen centraal staan, waarbij verschillende stadia van de SDLC worden beheerd door AI-agenten die voor specifieke doeleinden zijn gebouwd, mogelijk ondersteund door fijn afgestelde LLM's om hun effectiviteit verder te vergroten. Menselijke interventie blijft cruciaal om AI-hallucinaties te verminderen, beter onderbouwde antwoorden te waarborgen en uiteindelijk de productiviteit te verhogen.
- Stel je een nieuwe benadering van sofwareontwikkeling voor, waarbij natuurlijke taal de primaire interface voor ontwikkeling wordt. Agents die onderling samenwerken, een agent voor het genereren van user stories, een agent voor coderen, een agent voor testen, een agent voor deployment, en een agent die alles orkestreert en daarmee de hele SDLC raakt. Hierdoor zal de interactie tussen producteigenaren, bedrijfsanalisten, ontwikkelaars en testers zal veranderen.
- Product Owners zullen bruikbare inzichten ontvangen (analytics) om besluitvorming en strategieën te ondersteunen, waardoor ze beter in staat zijn om te sturen op basis van de business waarde
- Bedrijfsanalisten zullen trends en voorkeuren beter begrijpen, waardoor ze deze inzichten kunnen vertalen naar gedetailleerde ideeën en zelfs tot testschema's, wat leidt tot snellere iteraties en verbeterde product-markt fit.
- Architecten worden ondersteund bij het analyseren van input van stakeholders en het genereren van architectonische diagrammen, datamodellen en voorgestelde technologiestacks op basis van projectvereisten, waardoor het ontwerpproces wordt versneld en wordt voldaan aan best practices.
- Ontwikkelaars profiteren nog steeds van code-assistenten, maar nu met extra context uit eerdere fasen. Ervaren ontwikkelaars zullen zich meer richten op het optimaliseren van AI-agenten en het afhandelen van complexe ontwerp- en beoordelingstaken in plaats van het grootste deel van de code te schrijven. Voor specifieke functies, logica of UI-behoeften zullen ontwikkelaars meer tijd besteden aan zakelijke vereisten, ontwerp en teammanagement.
- Testers maken niet alleen gebruik van GenAI voor testgeneratie, maar ook synthetische testdatageneratie, voorspellende bugdetectie en een continue feedbackloop die testresultaten analyseert en verbeteringen voorstelt.





Tools zoals GitHub Copilot Workspace bieden al een voorproefje van de toekomst—een “Copilot-native” ontwikkelaarsomgeving waarin ontwikkelaars kunnen brainstormen, plannen, bouwen, testen en code uitvoeren met behulp van natuurlijke taal. Dit omvat verschillende AI-agenten die alle taken van begin tot eind uitvoeren, waarbij ontwikkelaars de controle over het proces behouden.



Tot slot

GenAI staat op het punt de SDLC te revolutioneren door coderingstaken te automatiseren, de kwaliteit te verbeteren en zijn invloed uit te breiden over de hele levenscyclus—van het verzamelen van eisen tot implementatie en onderhoud. Hoewel de huidige adoptie nog in de beginfase zit, zorgt het gebruik van code-assistenten zoals GitHub Copilot al voor een hogere productiviteit en kortere inwerktijden. Deze efficiencystijging kan echter knelpunten veroorzaken in andere SDLC-fasen, zoals bedrijfsanalyse en testen.

Om de kracht van GenAI volledig te benutten, moet het worden geïntegreerd in alle fasen van de SDLC. In de toekomst zal AI niet alleen repetitieve coderingstaken uitvoeren, maar ook functionele requirements genereren, architectuurontwerp maken, bugs voorspellen en implementatie optimaliseren. Multi-agent systemen zullen dit verder brengen door elk aspect van ontwikkeling naadloos en intelligent te orkestreren.

Wanneer we naar de komende 5 tot 10 jaar kijken, zal software-engineering steeds meer draaien om samenwerking met intelligente systemen. Menselijke ontwikkelaars zullen overstappen van handmatige codering naar het begeleiden en aansturen van op agent gebaseerde GenAI. Ze zullen zich richten op het optimaliseren van deze systemen door modellen te trainen en te bepalen welke AI-modellen of -oplossingen het beste passen bij specifieke onderdelen van de Software Development Life Cycle (SDLC). Hierdoor kunnen ontwikkelaars hun aandacht verleggen naar strategische verbeteringen, creativiteit en besluitvorming op hoog niveau. Bedrijfsanalisten zullen gebruikersfeedback rechtstreeks omzetten in functionele requirements, terwijl testers profiteren van voorspellende AI-modellen die de testdekking continu verbeteren en proactief bugs identificeren. Architecten zullen vertrouwen op door AI gegenereerde ontwerpen, waardoor de tijd die aan handmatige taken wordt besteed, afneemt. Deze AI-gedreven transformatie zal zorgen voor snellere, creatievere ontwikkelcycli en een betere afstemming op zakelijke doelen. GenAI zal een essentiële partner worden in de SDLC, en fundamenteel herdefiniëren hoe software wordt bedacht, gebouwd en onderhouden—waardoor we worden uitgedaagd om niet alleen onze processen, maar onze hele benadering van softwareontwikkeling opnieuw te overdenken. Zullen we deze evolutie omarmen en ruimte maken voor een AI-gedreven SDLC die zowel productiviteit als creativiteit verhoogt?

Over de auteurs



Donald Hessing | Vice President
Head of Microsoft, CTO Cloud & Custom Applications

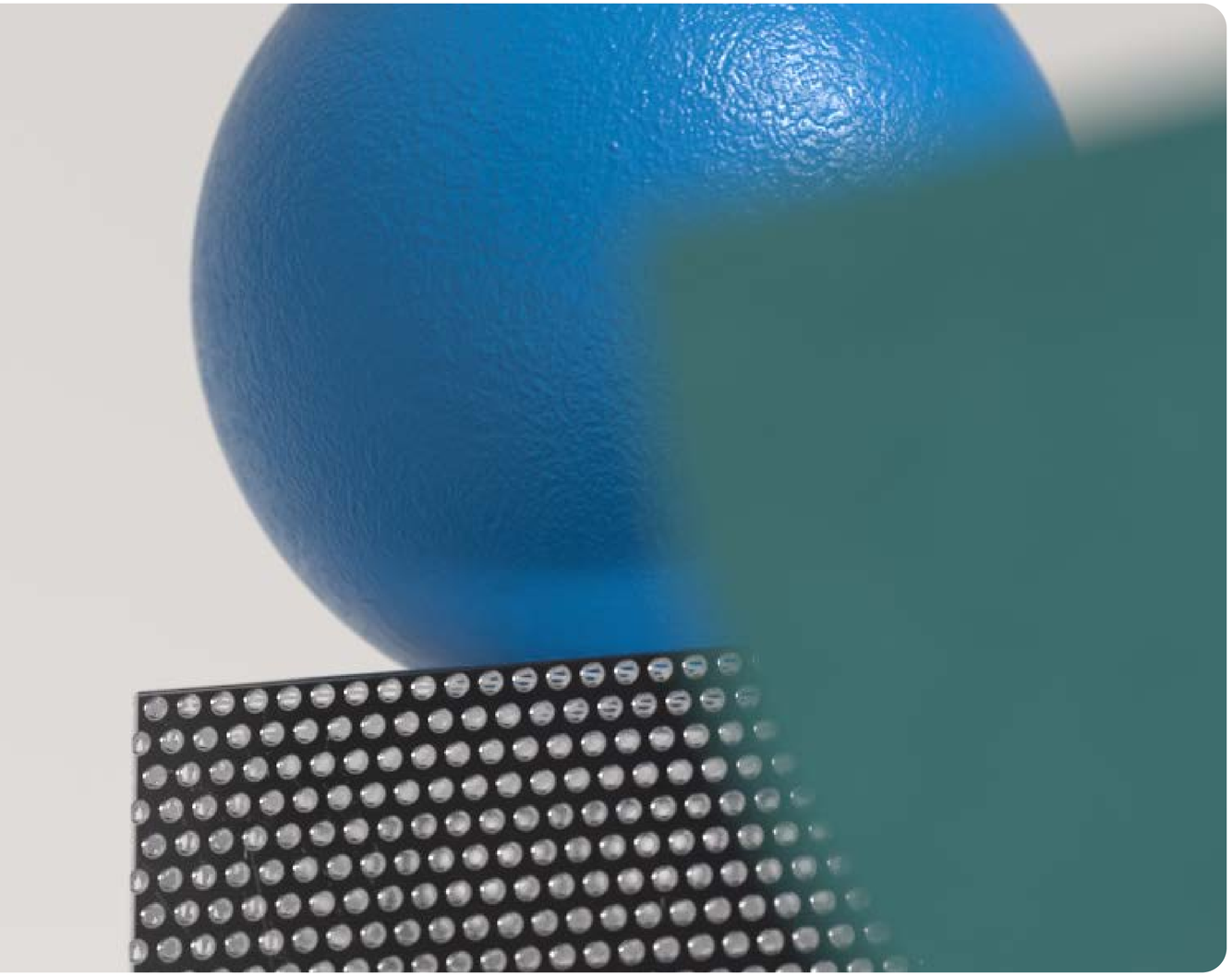
Donald geeft leiding aan het Microsoft cluster binnen de afdeling Cloud & Custom Applications. Daarnaast is Donald als CTO verantwoordelijk voor de visie en technologiestrategie van de afdeling. Hij is gespecialiseerd in software engineering, enterprise architecture en strategieontwikkeling van business tot technologie.



Arin Roy | Tech Lead and Lead Architect – Cloud Native and Generative AI

Als Tech Lead en Architect voor Cloud Native richt Arin zich op cloud native en Generative AI solutions in Microsoft Azure. Arin heeft meer dan 15 jaar ervaring met de technologieën van Microsoft. Die ervaring helpt hem om enterprise-ready digitale cloud-oplossingen te ontwerpen en implementeren voor zijn klanten.



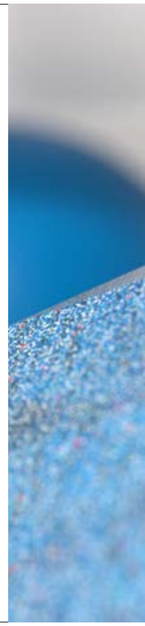


03 Vier praktische stappen op weg naar je eigen *GenAI* Factory

Een gids voor het efficiënt en verantwoordelijk integreren van Gen AI in uw organisatie

Highlights:

- Gen AI vereist een diepgaand begrip van machine learning en data science. Begin met een Discovery Workshop om potentiële use cases te identificeren.
- Gebruik een GenAI Factory-aanpak voor snelle prototyping en schaalbaarheid. Maak gebruik van enterprise-grade Cloud-native bouwstenen zoals Azure, AWS en Google Cloud.
- Transparantie, inclusiviteit, robuust gegevensbeheer, continue monitoring en duidelijke ethische richtlijnen zijn essentieel voor het opbouwen van vertrouwen en verantwoording.
- Optimaliseer AI-oplossingen continu en gebruik human-in-the-loop feedbackmechanismen om de technologie af te stemmen op gebruikersbehoeften en ethische normen.
- Begin gewoon met het ontwikkelen van uw eigen GenAI-oplossingen met deze gids onder de arm. Het belangrijkste is om te starten en te leren terwijl u bezig bent.



In veel bedrijven is Generative Kunstmatige Intelligentie (Gen AI) sinds een aantal jaren het gesprek van de dag. Van een niche-ontwikkeling groeide deze geavanceerde technologie uit tot een essentieel onderdeel in het bedrijfsproces, dat tools biedt zoals natural language processing, beeldgeneratie en predictive analysis. De grote aantrekkingskracht van Gen AI ligt in het vermogen om creativiteit te stimuleren en de operatie te stroomlijnen. Efficiëntie en innovatie krijgen dankzij GenAI een enorme boost.

Uitdagingen bij de integratie van Gen AI:

De integratie van gen AI in bedrijfsproces heeft een enorm potentieel. Het is echter nog niet zo eenvoudig om dat potentieel te vertalen in geschikte use cases.

In hun zoektocht naar effectieve GenAI-oplossingen en de implementatie daarvan, lopen organisaties tegen verschillende belemmeringen aan:

- Het ontbreekt organisaties aan de interne expertise om het volledige potentieel en de beperkingen van GenAI te kunnen doorgronden
- De integratie van GenAI in het bestaande IT-landschap kan lastig zijn
- Er bestaan zorgen omtrent data privacy en de ethische aspecten van GenAI
- De zoektocht naar use cases die waarde toevoegen en innovatie stimuleren verloopt moeizaam
- Organisaties moeten flinke financiële investeringen doen voor de implementatie van AI-oplossingen

Onze aanpak in vier praktische stappen:

In dit artikel beschrijven we hoe organisaties in 4 praktische stappen werk kunnen maken van GenAI:





Stap #1 – Vind de juiste use cases

Gen AI is een complex veld dat een diepgaand begrip vereist van machine learning, data science en kunstmatige intelligentie. Veel organisaties ontberen de expertise die nodig is om de mogelijkheden en beperkingen van gen AI volledig te kunnen doorgronden. Door deze kenniskloof kan het lastig zijn om te doorzien hoe AI u in staat stelt om kansen te grijpen, of specifieke zakelijke problemen op te lossen.

Discovery Workshop: De GenAI-reis van uw organisatie zou moeten beginnen met een Discovery Workshop. Zo'n workshop biedt inzicht in wat GenAI voor uw organisatie kan doen, en een platform voor een verkennende dialoog met ervaren professionals. In de Discovery Workshop worden verschillende generieke use cases gedeeld uit verschillende sectoren en off the shelf oplossingen gepresenteerd. Op die manier helpt de workshop potentiële use cases te identificeren voor uw organisatie.

Volgende fase: Zodra de use cases zijn bepaald, is het tijd voor de volgende fase.



Stap #2 – Begin klein, maar bouw de mogelijkheid in om op te schalen

Hoewel de meeste organisaties beginnen met een paar specifieke use cases, komen er na verloop van tijd onvermijdelijk nieuwe use cases naar voren. Om te voorkomen dat u elke proof of value (PoV) of pilot om het potentieel en de effectiviteit van geïdentificeerde use cases te bewijzen separaat moet ontwikkelen, raden we u aan een GenAI Factory-aanpak toe te passen. De GenAI Factory-benadering faciliteert snelle prototyping, waarbij ideeën worden vertaald naar een PoV zonder dat u steeds weer opnieuw de hele infrastructuur op hoeft te zetten. Op die manier biedt dit concept een organisatorische en technische structuur om meerdere initiatieven van PoV naar productie op te schalen.

Enterprise-Grade Cloud-Native Building Blocks: Door gebruik te maken van enterprise-grade Cloud-native bouwstenen (bijv. Azure, AWS, Google Cloud, Databricks), wordt schaalbaarheid vanaf het begin gegarandeerd, volgens het start-right, stay-right principe. Deze methode garandeert een robuuste basis en is daarmee essentieel voor het langetermijnsucces van onze GenAI-initiatieven.

De ruggengraat van onze GenAI Factory bestaat uit een solide infrastructuur van geavanceerde hardware, software en netwerkbronnen. Deze infrastructuur stelt ons in staat om geavanceerde AI-oplossingen efficiënt te ontwikkelen en implementeren zonder dat we custom oplossingen nodig hebben voor elke nieuwe use case.

Het GenAI Factory-concept is gestructureerd rond verschillende belangrijke componenten:

- GenAI Governance: Een overkoepelende entiteit die verantwoordelijk is voor het consolideren en managen van alle GenAI-activiteiten binnen de organisatie.
- Business Impact & Value Measurement: Een centraal en neutraal team dat de zakelijke impact van alle GenAI-use cases prioriteert en meet.
- Agile Development & Operations: Richt zich op het volledig in productie brengen van de beste use cases en het continu upgraden van bestaande projecten.
- Gestandaardiseerde Technologie: Radicale standaardisatie resulteert in een samenhangende strategie rond technologie, en geborgde compliance met regelgeving.
- Workforce Enablement: Trainingen om de GenAI-skills van uw workforce op te bouwen en de adoptie van technologieplatforms en assets te vergroten.

Door deze componenten te integreren, zorgt het GenAI Factory-concept voor een efficiënte opschaling van AI-initiatieven, waarbij uw organisatie compliance kan handhaven en continu kan innoveren - zonder maatwerk voor elke nieuwe use case.

Stap #3 – Integreer verantwoordelijke en ethische AI

De razendsnelle ontwikkeling van kunstmatige intelligentie (AI) luidt een transformatief tijdperk in van ongekende innovatie. Maar daarin hebben we allemaal een grote verantwoordelijkheid. We kunnen het belang van verantwoordelijke en ethische AI niet genoeg benadrukken; het vormt de basis voor duurzaam vertrouwen in en verantwoording over onze AI-systemen.

Alleen als we AI-technologieën op een verantwoorde manier inzetten en ontwikkelen, kunnen we potentiële risico's beperken zoals bias, privacyschendingen en onvoorziene, negatieve gevolgen. Door ethische overwegingen te prioriteren, beschermt u niet alleen gedeelde maatschappelijke waarden, maar bevordert u ook de acceptatie van en het vertrouwen in AI-oplossingen.

Leidende principes voor verantwoordelijke en ethische AI-ontwikkeling:

1. Transparantie:
   - Maak AI-algoritmen en besluitvormingsprocessen begrijpelijk voor alle belanghebbenden. Transparantie in elke stap van het proces vergroot het vertrouwen.
   - Zorg voor duidelijke documentatie. En zorg ervoor dat u de werking van AI-systemen kunt uitleggen. Dat bevordert vertrouwen en verantwoording.
2. Inclusiviteit:
   - Betrek diverse teams bij de ontwikkeling en implementatie van AI-technologieën.
   - Zorg ervoor dat AI wordt getraind met verschillende perspectieven, om bias te voorkomen.
3. Robuust gegevensbeheer:
   - Besteed aandacht aan data governance, om de privacy en veiligheid van gegevens te waarborgen.
   - Hanteer strikte ethische normen bij het verzamelen en gebruiken van data.
   - Implementeer strikte protocollen voor data-anonimisering, consent en toegangscontrole.
4. Continue monitoring en auditing:
   - Implementeer mechanismen voor de regelmatige evaluatie van de prestaties van AI en de impact ervan op gebruikers.
   - Zorg ervoor dat AI-systemen zich houden aan ethische normen.
   - Neem snel effectieve maatregelen als AI-systemen afwijken van ethische normen.
5. Duidelijke ethische richtlijnen en principes:
   - Stel richtlijnen op voor de ontwikkeling en implementatie van AI.
   - Stem deze richtlijnen af op maatschappelijke waarden en mensenrechten, zodat AI-technologieën bijdragen aan the greater good.
6. Een cultuur van ethische AI:
   - Bevorder een cultuur die ethische AI binnen de organisatie benadrukt.
   - Leer medewerkers over de ethische implicaties van AI.
   - Moedig medewerkers aan om zich uit te spreken als ze worden geconfronteerd met onethische praktijken.
   - Zorg dat ethische overwegingen worden verankerd in de organisatiecultuur, zodat verantwoorde AI-praktijken consequent worden doorgevoerd.

Stap #4 - Focus op gebruikersadoptie en waardecreatie

De succesvolle adoptie van Generative AI binnen uw organisatie staat of valt bij de gebruikersadoptie. Om die duurzaam te borgen, moet u AI-oplossingen continu optimaliseren. Daarbij zijn human-in-the-loop (HITL)-mechanismen voor feedback onontbeerlijk. Als u hierin slaagt, laat u AI-investeringen optimaal renderen en stimuleert u aanzienlijke waardecreatie.

Human-in-the-Loop Feedback Mechanismen:

Human-in-the-loop feedbackmechanismen zijn cruciaal bij het verfijnen van AI-systemen. Via HITL maakt u menselijke expertise en meningen onderdeel van de AI-life cycle. Zo kan de technologie op de juiste manier ontwikkelen, in lijn met de wensen van gebruikers en ethische normen. En zo kunt u potentiële gevallen van bias aanpakken en vergroot u de nauwkeurigheid en relevantie van het systeem.

Feedback loops kunnen worden ingesteld in verschillende stadia van AI-implementatie. Tijdens de beginfasen helpt gebruikersfeedback bij het identificeren van relevante problemen en het verfijnen van de functionaliteiten van de AI-oplossing. Naarmate het systeem volwassen wordt, zorgt continue feedback ervoor dat de AI zich aanpast aan veranderende vereisten. Bovendien borgt u op deze manier dat AI effectief blijft bij het oplossen van complexe problemen.

Waardecreatie door continue optimalisatie:

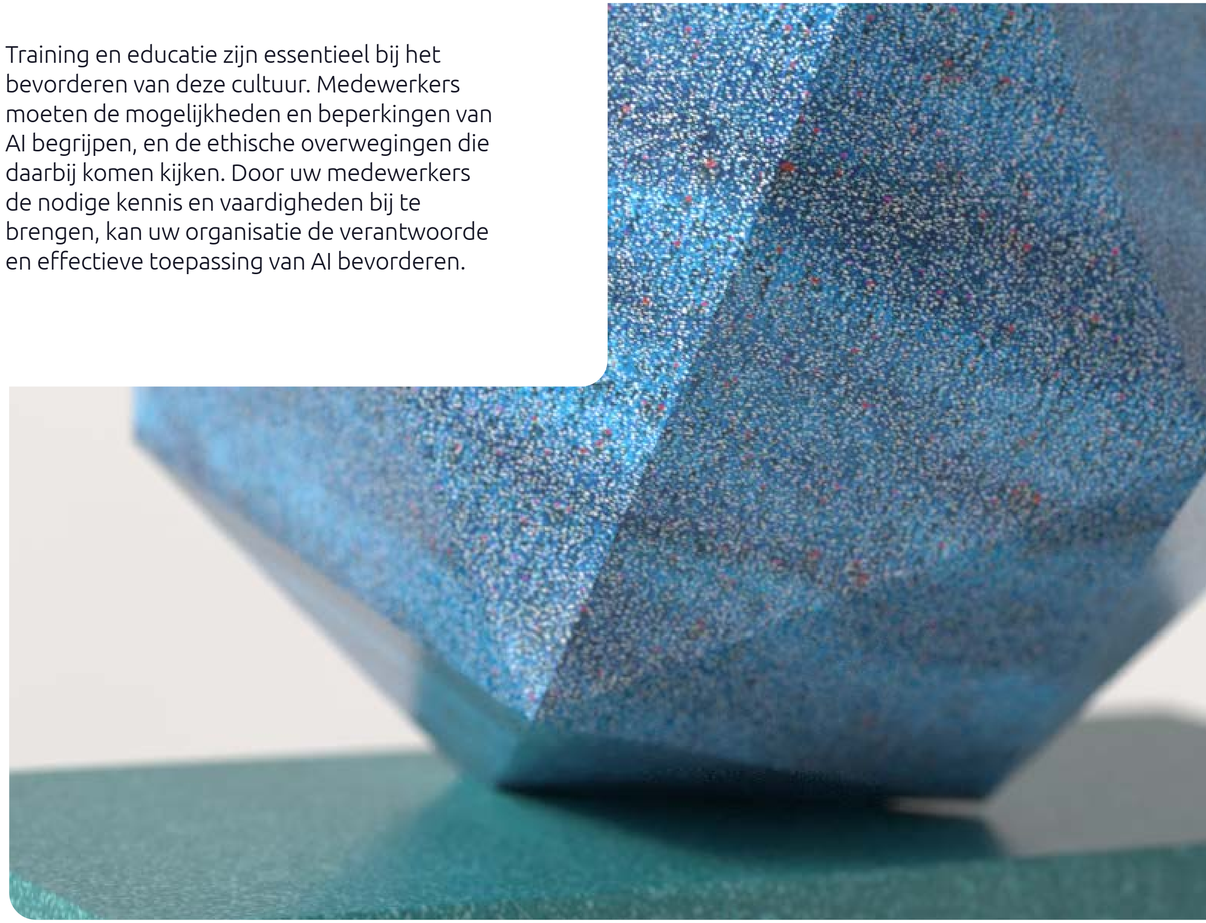
Om echt te profiteren van GenAI, moeten organisaties zich richten op constante optimalisatie. Dit omvat niet alleen technologische verfijning, maar ook het afstemmen van AI-capaciteiten op bedrijfsdoelstellingen. Optimalisatie is een vereiste om de efficiëntie en de effectiviteit van de AI-oplossing te kunnen blijven maximaliseren en garanderen.

Hierbij zijn data-gedreven inzichten van groot belang. Dat inzicht verkrijgt u door de analyse van prestatiestatistieken en gebruikersinteracties. Zo kunt u verbeterpunten identificeren, en de juiste maatregelen treffen. Op die manier blijft u werken aan gebruikerstevredenheid en operationele efficiëntie. Continue optimalisatie zorgt ervoor dat AI-investeringen relevant blijven en blijvend bijdragen aan de strategische doelen van de organisatie.

Een cultuur van innovatie bevorderen:

Generative AI gedijt bij een cultuur van innovatie. Een cultuur die het experiment aanmoedigt, die leren stimuleert en de adoptie van nieuwe technologieën bevordert. Als leider speelt u een belangrijke rol bij het ondersteunen van AI-initiatieven. Bijvoorbeeld door de nodige middelen vrij te maken, en een omgeving te creëren waarin medewerkers zich empowered voelen om te ontdekken en te innoveren.

Training en educatie zijn essentieel bij het bevorderen van deze cultuur. Medewerkers moeten de mogelijkheden en beperkingen van AI begrijpen, en de ethische overwegingen die daarbij komen kijken. Door uw medewerkers de nodige kennis en vaardigheden bij te brengen, kan uw organisatie de verantwoorde en effectieve toepassing van AI bevorderen.



De weg naar succesvolle AI-adoptie:

Uiteindelijk hangt de succesvolle adoptie van GenAI af van de integratie van Human-in-the-loop feedbackmechanismen, en een onophoudelijke zoektocht naar waardecreatie. Door gebruikersfeedback te prioriteren en AI-oplossingen continu te optimaliseren, kunt u het volledige potentieel van de technologie benutten. Zo stimuleert u niet alleen de gebruikerstevredenheid, maar bevordert u bovendien de groei van uw bedrijf en bouwt u aan concurrentievoordeel.

Organisaties die deze principes omarmen, zijn goed gepositioneerd om de complexiteit van AI-implementatie te navigeren en de vruchten te plukken van hun investeringen in Generative AI. Naarmate de technologie evolueert, moet u aan deze principes blijven vasthouden. Zo zorgt u ervoor dat AI-systemen adaptief, ethisch en impactvol blijven.

Conclusie:

Er is een grote toekomst weggelegd voor Generative AI. Naarmate de ontwikkeling van de technologie vordert, zullen AI-systemen steeds geavanceerder worden. Ze zullen daarmee in staat zijn om steeds complexere taken uit te voeren. GenAI is here to stay; elke organisatie moet erin investeren om relevant te blijven. Onze 4-stappen-aanpak kan inspiratie bieden om aan de slag te gaan met deze veelbelovende technologie.

In dit artikel bespraken we alle fasen die komen kijken bij de succesvolle, ethische implementatie van AI in uw organisatie. Van de initiële inspiratie- en probleemidentificatiefase, via de ontwikkeling en versnelling van oplossingen die meerwaarde bieden, tot de uiteindelijke opschaling en integratie. Daarmee heeft u een praktische aanpak in handen om het transformerende potentieel van GenAI duurzaam te verankeren in uw organisatie.

Over de auteurs



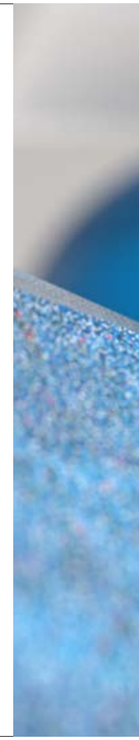
Sjoukje Zaal | Chief Technology Officer Gen AI Cloud Northern and Central Europe

Als Chief Technology Officer bij Capgemini brengt Sjoukje meer dan 20 jaar ervaring mee in architectuur, ontwikkeling en consultancy, met een sterke focus op AI, data en innovatie. Ze leidt ook de architectuurgemeenschap van Capgemini in Nederland. Daarnaast is Sjoukje directeur van de Global AI Community, waar ze kennisdeling en samenwerking stimuleren. Als internationaal spreker, evenementorganisator en auteur is ze gepassioneerd over het bevorderen van thought leadership en het vormgeven van de toekomst van technologie.



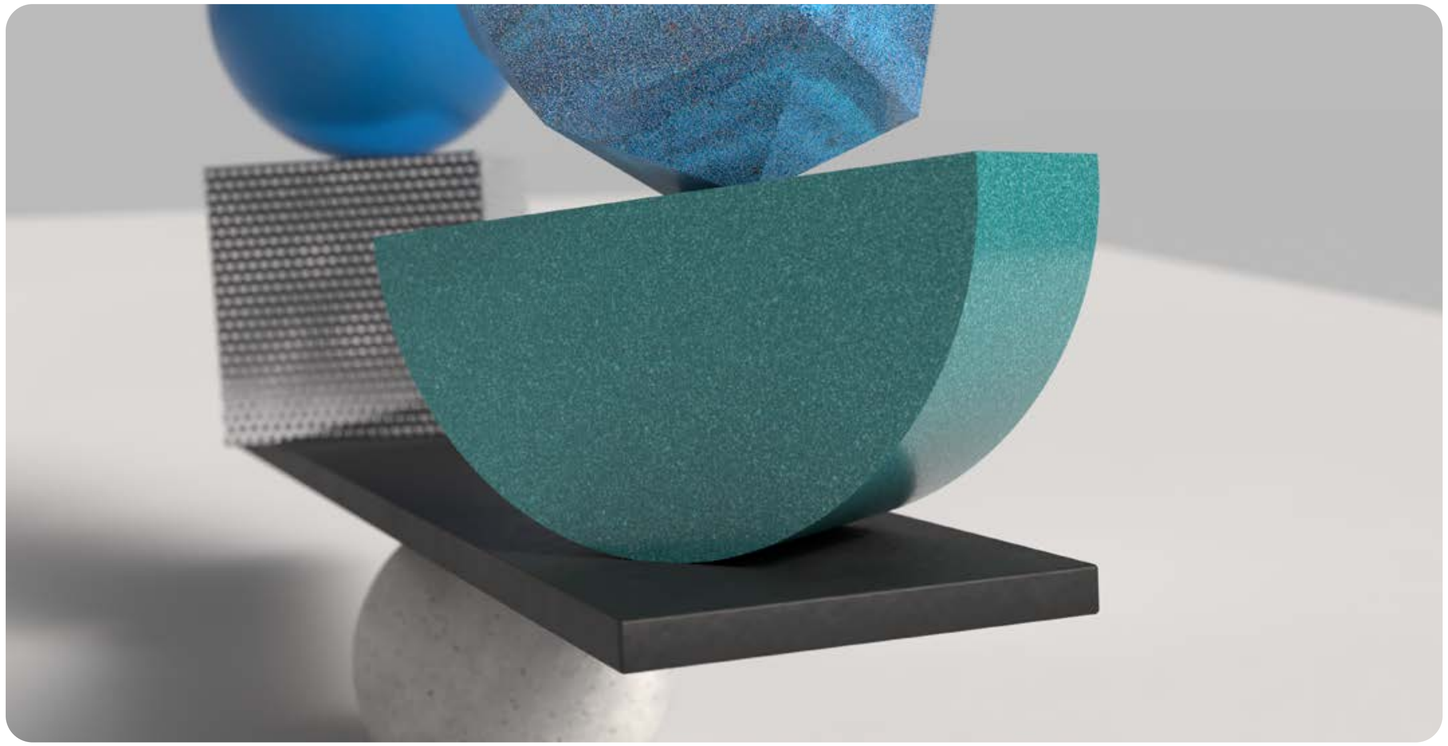
Douwe van de Ruit | Lead Architect Gen AI Cloud Northern and Central Europe

Douwe is Principal Consultant gespecialiseerd in Generative Artificial Intelligence en Cloud-technologie. Hij is momenteel Lead GenAI voor de NL & NCE Cloud CoE's. Voordat hij bij Capgemini kwam, werkte Douwe als solution architect met een sterke focus op de adoptie van opkomende technologieën. Met meer dan 20 jaar ervaring in de IT-sector richt hij zich op het ontdekken en ontsluiten van waarde door innovatieve oplossingen.



04 *AI en Examenstress:* Hoe Slimme Machines Worden Beoordeeld

Gebruik ons eigen lerende vermogen om AI-oplossingen te testen.



Highlights:

- Om AI te testen kunnen we niet een verwacht resultaat vergelijken met het gegeven resultaat. Een AI-systeem kan morgen immers een heel ander antwoord geven.
- Inzicht in het proces om te komen tot een AI-oplossing helpt om ook voor AI gestructureerd testen op te zetten.
- Testdata en trainingsdata zijn cruciaal voor het succesvol uitvoeren van een AI-teststrategie.
- Het testen van AI kan heeft een eigen aanpak nodig. Door te kijken naar hoe wij leren en kennis toepassen in combinatie met examinering is een goede methode nodig om de kwaliteit van AI te meten en te blijven monitoren.



Gestructureerd testen is onze basis voor het definiëren van een teststrategie[1]. Een goed of fout resultaat na het uitvoeren van een test geeft ons alle informatie om te werken aan de kwaliteit van een product. Een product kan van alles zijn, van een bankapp, een klantenservicewebsite, een applicatie die alle onderdelen volgt die in een autofabricageproces gaan tot een MRI-apparaat voor een ziekenhuis. Bij het ontwikkelen van een AI-oplossing is testen cruciaal[2].

1. TMAP: Quality for DevOps Teams, 2020 ; ISBN, 9789075414899
2. Testing in the digital age; AI makes the difference, 2028; ISBN, 9789075414875

Natuurlijk moet ervoor gezorgd worden dat aan de functionele en niet-functionele specificaties is voldaan, maar het testen van AI-oplossingen gaat anders dan dat we gewend zijn.

Er is een groot verschil tussen een AI-oplossing en de traditionele meer deterministische oplossing:

- Het gedrag is veel moeilijker te voorspellen. Op een bepaald moment kan een uitkomst verschillen van een ander moment. We moeten rekening houden met testsituaties die dit afdekken. Je kunt dit beperken door weg te blijven van continu lerende systemen die een tijdsparameter in hun uitkomsten hebben. In dat geval leert het algoritme met de trainingsdataset en testen we vervolgens de uitkomst zoals we gewend waren voor niet-AI systemen. Een gestructureerde test strategie kan hier prima mee uit de voeten. In het kader hieronder staat de uitleg grafisch weergegeven hoe we onderscheid maken tussen training data en test data als onderdeel van de gehele data set(zie figuur 2 en kadertekst “dataset uitgelegd”).
- Er is geen zwart-wit uitkomst voor een gegeven situatie. Testvoorspellingen zijn niet langer 1 of 0, maar hebben een bandbreedte. Zorg ervoor dat je testcases dit gedrag goed weerspiegelen.
- Niet-lineair gedrag van AI-oplossingen kan resulteren in zogenaamde hallucinaties. Hallucinaties zijn uitkomsten die door een AI worden gemaakt maar die niet bestaan. Een bibliografie link in een wetenschappelijk artikel kan door een AI perfect worden nagemaakt met alle bijbehorende look and feel. Als het vervolgens een link is naar een niet bestaand artikel met een verzonnen titel en random gegenereerde link (die er wel uitziet als een echte link) dan spreken we van een hallucinatie.
- Naast een trainingsdataset hebben we een testdataset nodig om input-output combinaties te testen. Het aantal combinaties is waarschijnlijk veel hoger voor AI-systemen dan voor niet-AI-systemen om dit te af te dekken. Automatisering voor het testen van deze situaties is essentieel.

Wat moet er getest worden?

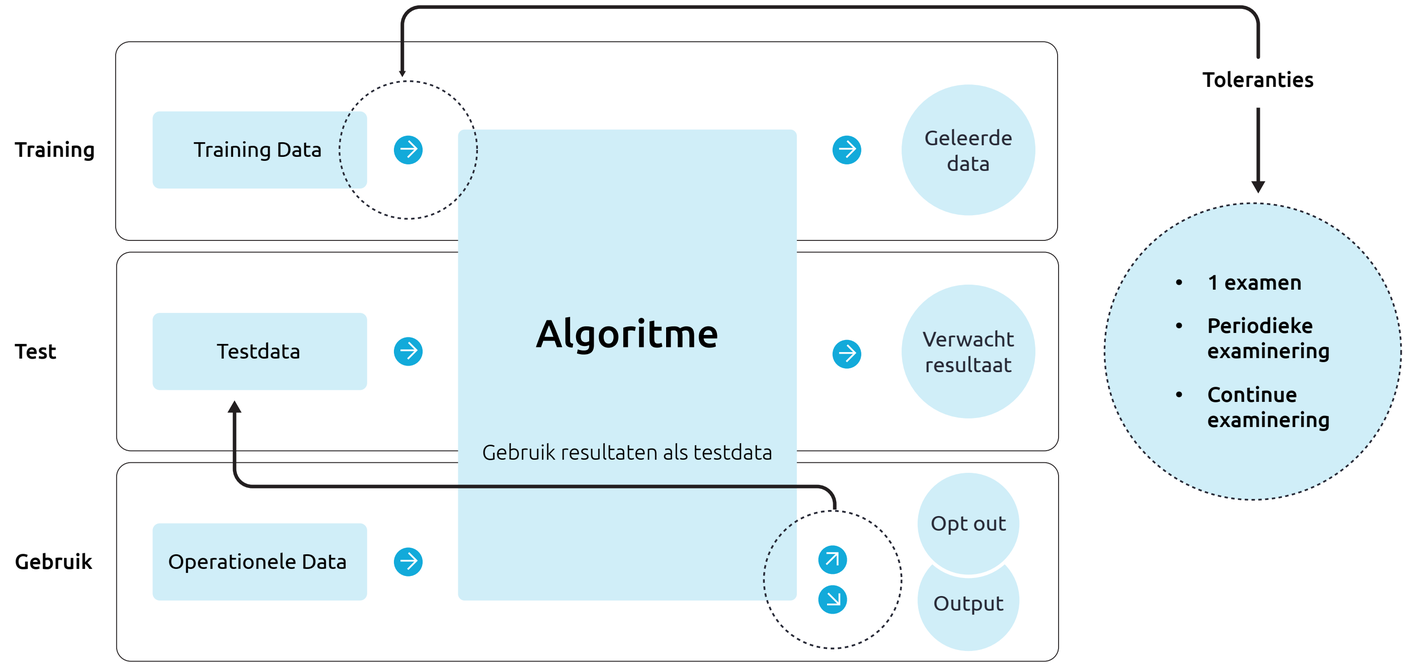
Bij het ontwikkelen van een AI-oplossing is de testtaak een cruciale activiteit. Degene die deze taak uitvoert, moet ervoor zorgen dat de functionele en niet-functionele specificaties worden vervuld. Vanwege de complexe werking van een AI-systeem moeten veel meer invoerwaarden worden getest om een robuuste oplossing te verifiëren.

Enkele voorbeelden van vragen die we zouden kunnen stellen bij het testen van AI zijn:

- Wat zijn de acceptatiecriteria?
- Hoe kunnen we testcases ontwerpen die deze criteria met minimale inspanning testen?
- Zijn er verschillende datasets voor trainingsdata en testdata?
- Vertegenwoordigen de testdata de verwachte data voldoende?
- Zijn de test- en trainingsdata in overeenstemming met het wettelijke kader?
- Zijn de trainingsdata bevooroordeeld?
- Wat is het verwachte resultaat van het model?
- Is het model onder- of overgefit ? [3]
- Gedraagt de oplossing zich ethisch?
- Zijn de prestaties en robuustheid van de oplossing goed genoeg?



3. Onderfitten treedt op wanneer een model te eenvoudig is om de onderliggende structuur van de gegevens te leren. Dit betekent dat het model niet in staat is om patronen in de trainingsdata te herkennen, wat resulteert in slechte prestaties zowel op de trainingsdata als op nieuwe data. Overfitten daarentegen treedt op wanneer een model te goed presteert op de trainingsdata, maar slecht generaliseert naar nieuwe data. Dit betekent dat het model te veel details en ruis uit de trainingsdata heeft geleerd, waardoor het niet goed presteert op nieuwe data.



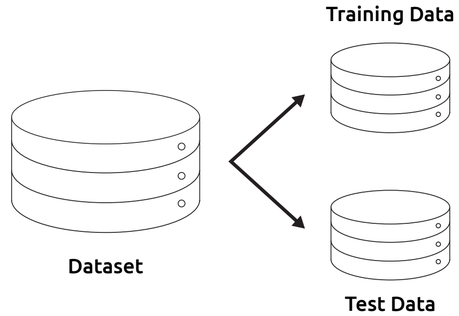
Model van het testen van AI-oplossingen.

De dataset uitgelegd

De dataset bevat alle data die beschikbaar zijn voor de AI-oplossing. Vaak zijn het de historische data van een bedrijf en moet deze worden verwerkt om bruikbaar te zijn. Ruwe data bevatten vaak incomplete items of verkeerd getagde items. Deze moeten eerst bewerkt worden om als correcte input te dienen. Foute input leidt immers tot foute output.

Trainingsdata

Trainingsdata zijn een subset van de dataset. Deze data worden gebruikt om het model te trainen tijdens het productontwikkelingsproces



De dataset uitgelegd.

Definities voor het testen van AI-oplossingen.

Geleerde data

Na het trainen van een AI-oplossing wordt een specifieke input-output-relatie opgezet. De data die nu uit het systeem komen op basis van trainingsdata worden geleerde data genoemd.

Testdata

Testdata zijn een andere subset van de dataset en worden gebruikt om te verifiëren of het model werkt zoals bedoeld.

Het is essentieel om geen trainingsdata als testdata te gebruiken omdat het verifiëren of de AI heeft geleerd wat het nodig heeft om beslissingen te nemen, andere data vereist dan de trainingsdata. Als er is geleerd met testdata zal de test altijd een uitkomst laten zien die klopt. Het geleerde deel zal met dezelfde data eenzelfde resultaat opleveren.

Verwacht resultaat

Op basis van testdata en een getraind systeem wordt de output voorspeld. Dit is het zogenaamde verwachte resultaat. Verwachte resultaten worden vergeleken met werkelijke resultaten. Dit kan worden gedaan door middel van onderzoek, continue inspectie of klassieke validatie. Continue inspectie kan ook worden gedaan op oplossingen in de gebruiksfase.

Operationele data

De data in het veld. Werkelijke waarden die door de AI-oplossing worden gebruikt.

Model

Het model bestaat uit de algoritmes die worden gebruikt, waarmee de AI leert van de gegeven data.

Training (fase)

De training is het proces waarin het algoritme leert van data en voorspellingen doet.

Test (fase)

Nadat het model is getraind, kan het worden getest met behulp van testdata.

Gebruik (fase)

Het AI-systeem creëert nu met echte data output. Wanneer de input wordt gevonden die niet kan worden verwerkt tot enige output, zal deze uit zijn routine stappen en wordt er een melding gegeven. Dit is nuttige input voor het testen van nieuwe versies van de oplossing.

Broncode (Algoritme)

Een AI-systeem heeft veel minder regels code dan de klassieke oplossingen. Toch is er broncode die fouten kan bevatten en daarom moeten er op ontwikkelniveau unittesten en andere relevante tests aanwezig zijn.

Infrastructuur

De infrastructuur heeft een sterke link met de niet-functionele specificaties van een product, die moeten worden gecontroleerd en getest.

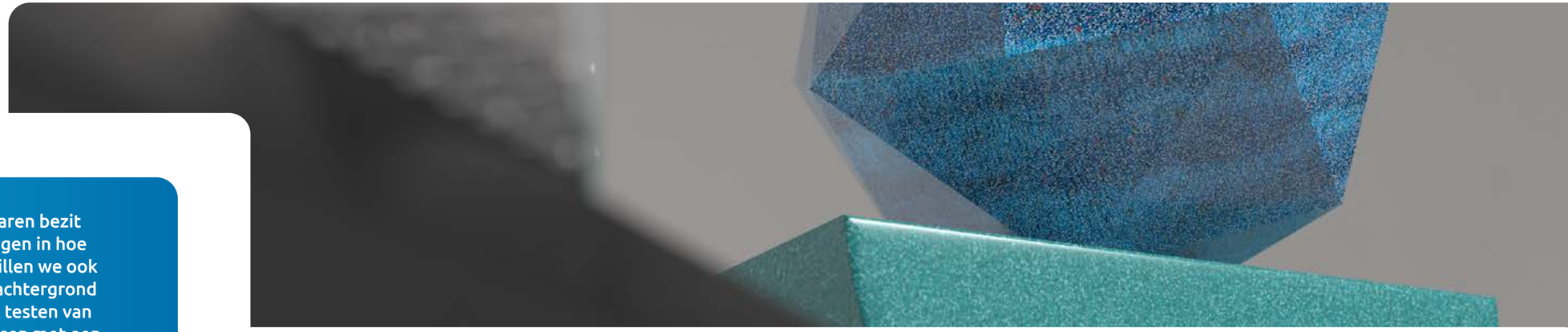
Specificaties

De specificaties op het gebied van AI moeten zorgvuldig worden geïnspecteerd. De technologie is nieuw, dus het is mogelijk dat de verwachtingen onrealistisch zijn of buiten wettelijke of ethische grenzen vallen.

Invoer/uitvoerwaarden

De meest basale testobjecten zijn de invoer- en uitvoerwaarden. Dit is waar de acceptatiecriteria worden geverifieerd. De invoerwaarden in AI-oplossingen zijn cruciaal omdat het onbekend is hoe de data wordt verwerkt.





Leren is een vaardigheid die de mens al duizenden jaren bezit en door de jaren heen hebben we veel inzicht gekregen in hoe mensen leren. Nu we machines creëren die leren, willen we ook testen of de machine goed heeft geleerd. Met een achtergrond in softwaretesten lijkt het misschien logisch om het testen van AI als softwaretesten te benaderen. Maar voor mensen met een onderwijskundige achtergrond is het veel logischer om een lerende machine op dezelfde manier te benaderen als een leerling.

Wat zou dit laatste betekenen voor het testen van AI? Het begint met de leerdoelen. Wat moet het systeem leren? Dit gaat over gewenste competenties, gewenst prestatieniveau en de manier van leren (bijv. begeleid, onbegeleid of door middel van beloningen).

Wanneer het systeem leert, willen we weten hoe goed het leert. Dit is een kwestie van meten. Net als wanneer mensen leren, kunnen we het systeem een examen laten afleggen om zijn capaciteiten te tonen. En natuurlijk doen we voor het eindexamen voorlopige examens om de leercurve te monitoren. Dit alles kan worden gezien als een prestatie-analyse van de intelligente machine.

Deze prestatie-analyse kan worden weergegeven via een dashboard dat de relevante aspecten en de vooruitgang in de tijd toont. Dit kan ook worden gebruikt om het moment te bepalen waarop de intelligente machine voldoende heeft geleerd om zakelijke waarde te bieden.

Voor het leren krijgt het systeem specifieke trainingsdata die de basis vormen voor zijn leerproces. Maar ook worden algoritmen geselecteerd die het beste passen bij het doel van de intelligente machine.

Tijdens het testen voeren de testers specifieke testdata in het systeem in om te verifiëren of de reacties zijn zoals verwacht. En op basis van de analyse van de leervermogens en resultaten worden nieuwe trainingsdata gemaakt om het leerproces verder te verbeteren en worden testdata gemaakt om te worden gebruikt bij een volgend examen.

Examen

Mensen leren allerlei vaardigheden, en sommige moeten worden bewezen door vaardigheid, bijvoorbeeld de vaardigheden om een auto te rijden. In dat geval worden ze geëxamineerd om te beoordelen of ze goed genoeg presteren. Nadat ze het examen hebben gehaald, vertrouwen we erop dat mensen in de toekomst altijd de juiste beslissingen zullen nemen. Het examen controleert of je een set regels correct hebt geleerd en kunt toepassen. Het interessante is dat, bijvoorbeeld bij autorijden, het echte leren pas echt begint na het examen. In de loop van de tijd kom je met autorijden in veel situaties waarin je deze regels moet toepassen. Zelfs na een leven lang autorijden heb je nog niet alle mogelijke situaties meegemaakt waarin je kunt belanden. Geen twee personen hebben dezelfde ervaringen met autorijden. Toch vertrouwen we erop dat alle mensen met een rijbewijs de regels correct toepassen in elke nieuwe situatie.

Dezelfde denkrichting kan worden toegepast op AI-oplossingen, en we kunnen onderscheiden tussen:

- examinering van AI
- periodieke examinering van AI
- continue examinering van AI



Examinering van AI

Net als bij mensen willen we bij een intelligente machine ook weten of deze goed genoeg presteert. Testen is natuurlijk de basisaanpak om de kwaliteit te evalueren. Omdat het niet mogelijk is om alle mogelijkheden te testen, kunnen we de aanpak van examinering gebruiken om te beslissen of de kunstmatige intelligentie kan worden vertrouwd om de taak uit te voeren. Zodra de AI het examen haalt, kan deze in live-operatie worden gebruikt.

Periodieke examinering van AI

Als er een AI-algoritme wordt gebruikt om met data te leren en de uitkomst wordt bevroren, dan verandert de intelligente machine na het examen niet van gedrag. Als de intelligente machine blijft leren, moeten we echter af en toe opnieuw examens afnemen.

Bij een intelligent systeem is een periodiek examen heel goed mogelijk, mits het examen zelf ook door een geautomatiseerd systeem kan worden uitgevoerd. Hierdoor kan er snel een test worden uitgevoerd en is het niet saai om continue dezelfde test uit te moeten voeren. Het is dan ook mogelijk continue variatie toe te voegen (aan een stabiele test basis) om het continu lerende systeem scherp te houden en te zien of alle reacties binnen gestelde examinering grenzen blijft.

Continue examinering van AI

Bij periodieke examinering is een fundamentele vraag hoe vaak dit examen afgenomen moet worden. Dat heeft te maken met de periodiciteit van het gebruik van de intelligente machine, maar de meeste machines zullen continu functioneren. Waarschijnlijk zal het nodig zijn om continue “operationele toetsing” uit te voeren. Vooral bij closed-loop-AI en andere vormen van AI-systemen die continu hun gedrag verbeteren, is het belangrijk om continue monitoring van de resultaten te implementeren om te controleren of de resultaten binnen de toleranties blijven die als “goed gedrag” zijn gedefinieerd.

Is het AI-systeem geslaagd of gezakt voor zijn examen?

“Heeft het AI-systeem examenstress?” is dus niet zo’n vreemde vraag. Misschien moet examinering opgenomen worden in de productontwikkelingscyclus. Een reguliere regressietest geeft niet het vertrouwen dat evoluerende AI-oplossingen zo robuust en “correct” zijn als je zou willen. Een examen is bedoeld om vertrouwen te hebben dat de geëxamineerde persoon (of in dit geval AI-oplossing) de taak nu en in de toekomst kan uitvoeren.

Ook her-examinering is een goed idee. Houd het kennisniveau up-to-date en controleer of nieuwe inzichten en nieuwe situaties nog steeds volgens de examineringstandaard worden afgehandeld. AI’s moeten misschien heel veel examens doen met een veel hogere frequentie dan mensen. Met testautomatisering kunnen we gemakkelijk een mechanisme implementeren van regelmatige AI-examinering met een hoge frequentie (dagelijks of zelfs ieder uur!).

Door veel examens af te nemen kunnen we continue naar het rapport kijken van de AI. Net als op school kan er bepaald worden of de AI geslaagd is of gezakt om door gaan naar een volgende periode of een volgend niveau waar zijn omgeving langzaam maar zeker complexer wordt.

Over de auteur



Tom van de Ven | Portfolio director Capgemini Engineering (NL)
Tom is Portfolio Director voor Capgemini Engineering en High Tech enthousiasteling. Hij spreekt en inspireert voor en over tech en is auteur van de boeken IoTMap en Testing in the digital age.

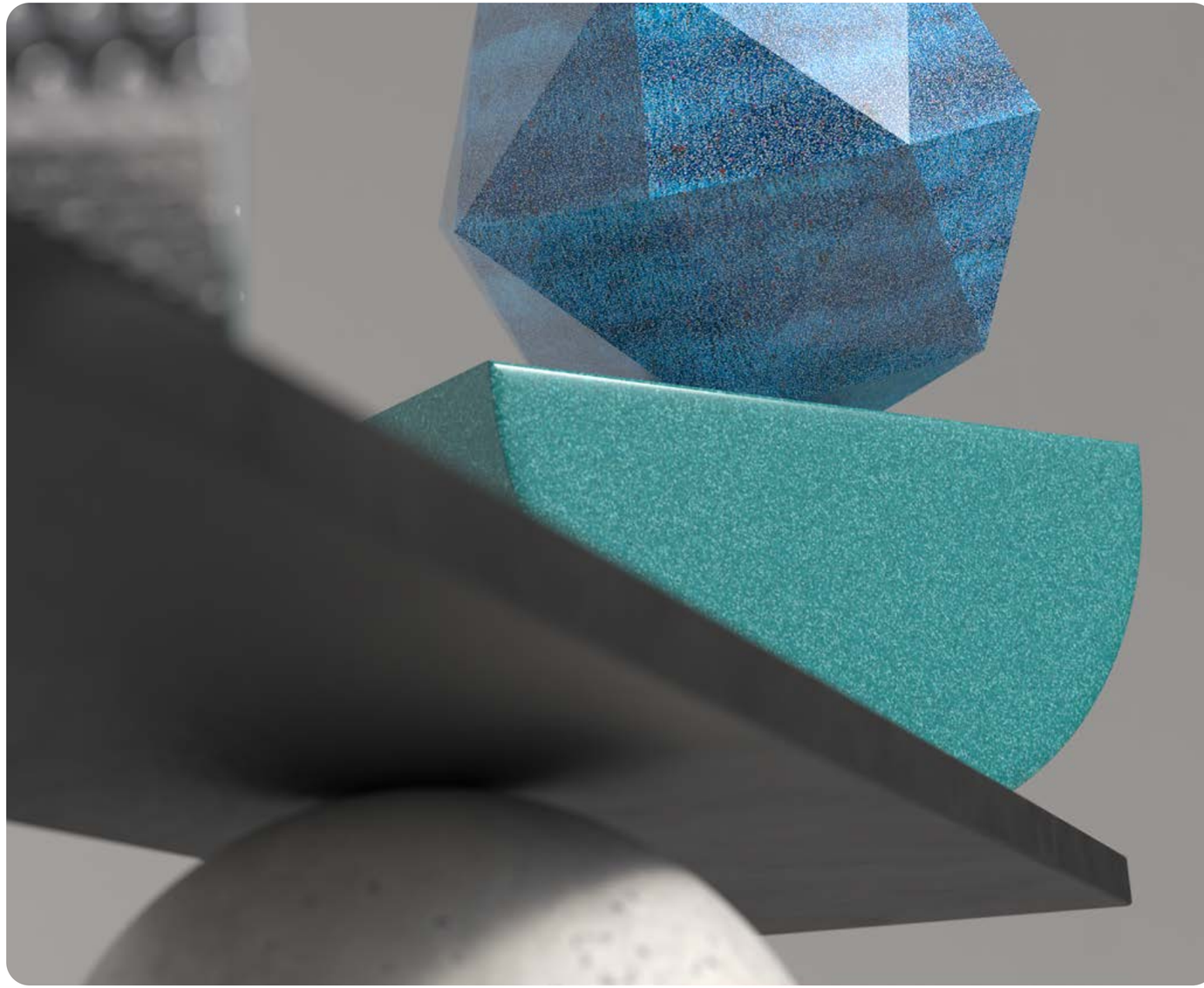
05 *Gen AI:* de waarheid claimt zichzelf niet

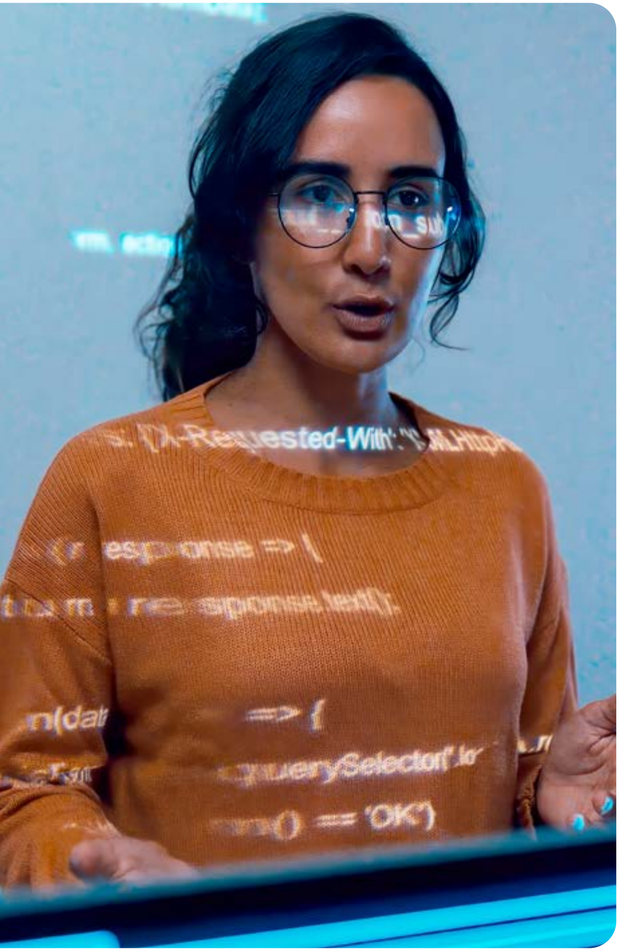
Generative Artificiële Intelligentie (AI) heeft de afgelopen jaren aanzienlijke vooruitgang geboekt en is een krachtig hulpmiddel geworden in verschillende domeinen, van kunst en literatuur tot wetenschap en technologie. Generative AI kan teksten samenvatten, illustraties genereren en onverwachte oplossingen voor bestaande problemen verzinnen. Vooral door dat laatste is het gebruik van generative AI zo aantrekkelijk voor het bedrijfsleven.



Highlights:

- Waarheidsgetrouwe Artificiële Intelligentie (AI) is essentieel voor kritische toepassingen.
- AI-systemen moeten om te leren gevoed worden met ware en actuele informatie.
- Strikte trainings-, validatie- en testprocedures zijn noodzakelijk.
- Domeinexperts zullen het waarheidsgehalte van de AI-antwoorden moeten controleren.
- AI zal nooit altijd de waarheid kunnen spreken omdat de wereld te complex is.





Taalmodellen en Generative AI

Generative AI is in essentie een “taalmodel”. Het leert op basis van statistische analyses van bestaande objecten, zoals tekstfragmenten, de onderliggende structuur te extraheren en op basis daarvan nieuwe taal te genereren. Dat kunnen teksten zijn, maar ook beeld- of muziektaal. Omdat in de tekstfragmenten, waarmee AI leert, ook kennis zit verborgen, wordt Generative AI ook gebruikt als kennissysteem. We kunnen een AI-opdracht geven om antwoorden te verzinnen bij onze vragen. Deze antwoorden zijn gebaseerd op wat de AI geleerd heeft. Deze kennis waarmee programmeurs de AI hebben gevoed, zal de Generative AI dan op ingenieuze manier samenvoegen en reproduceren.

“We denken dat mensen falen, maar technologie niet.”
- Marleen Stikker

Maar generative AI-modellen presenteren vaak onnauwkeurige antwoorden en doen alsof deze antwoorden correct zijn. Dit verschijnsel wordt AI-hallicunaties genoemd. AI-hallucinaties gebeuren wanneer een AI antwoorden op een vraag verzint in plaats van nauwkeurige informatie aan te bieden. AI-hallucinaties kunnen in ernst variëren, van de AI die kleine details in een antwoord verkeerd heeft, tot het verzinnen van de hele reactie. Ze worden vaak veroorzaakt door beperkte informatie in het systeem, vooroordelen in trainingsgegevens en problemen met het algoritme.

Hallucinaties kunnen op meerdere manieren worden bestreden. Aan de invoerkant door correcte, juiste en actuele data te gebruiken. Dus niet het gehele internet als databron, maar zorgvuldig geselecteerde databronnen waarvan het waarheidsgehalte hoog is. Bijvoorbeeld, in de gezondheidszorg wordt dan gebruik gemaakt van peer-reviewed artikelen en handboeken. Aan de uitvoerkant kunnen de antwoorden worden vergeleken met bij voorbeeld de resultaten uit andere kennissystemen, of kan worden nagegaan of er bewijs – onderbouwing - is in de databronnen voor het antwoord.

Het waarheidsprobleem van AI

Dit brengt ons bij de kern van het waarheidsprobleem van AI. Zelfs fabrikanten van AI, zoals OpenAI, waarschuwen ervoor dat de antwoorden die hun AI kan geven, fout kunnen zijn. En gezien de statistische werking van deze AI-algoritmes, zal het wel in de loop van de tijd beter worden, maar nooit helemaal perfect. Wanneer we een sprookje willen schrijven of een fantasie-afbeelding willen maken, is waarheidsgetrouwheid niet zo erg belangrijk. Maar wanneer het gaat om AI-gebaseerde beslissingen met grote gevolgen, is correctheid essentieel bij het gebruik van generative AI.

Wanneer Generative AI een hand tekent met zes vingers, is dat grappig. Maar bij veel domeinen is correctheid van groter belang. Domeinen zoals medische diagnostiek, gezondheidszorg, financiële beslissingen, juridische toepassingen, onderwijs, onderzoek, en media en informatieverspreiding. In al deze domeinen kunnen fouten leiden tot ernstige gevolgen, variërend van gezondheidsrisico’s en financiële verliezen tot juridische complicaties en de verspreiding van desinformatie: fake news.

1. https://waag.org/nl/article/technologie-kan-ook-falen/
2. https://www.palantir.net/blog/practical-ai-hallucination-awareness
3. https://charitydigital.org.uk/topics/generative-ai-content-the-need-for-accuracy-11306

01

Medische diagnostiek en gezondheidszorg

In de gezondheidszorg kan generative AI worden gebruikt voor het analyseren van medische beelden, het stellen van diagnoses en het aanbevelen van behandelingen. In deze context is correctheid van levensbelang. Een foutieve diagnose of een onjuist behandelplan kan ernstige gevolgen hebben voor de gezondheid van de patiënt en kan zelfs levensbedreigend zijn. Daarom moeten AI-systemen in de gezondheidszorg grondig getest en gevalideerd worden om ervoor te zorgen dat ze nauwkeurige en betrouwbare resultaten leveren.

Een concreet voorbeeld is een AI-systeem dat een röntgenfoto van een patiënt analyseert en ten onrechte concludeert dat er geen tekenen van kanker zijn, terwijl er in werkelijkheid wel een tumor aanwezig is. Dit kan ertoe leiden dat de patiënt niet de noodzakelijke behandeling krijgt, wat de prognose van de patiënt aanzienlijk kan verslechteren.

02

Financiële beslissingen

Generative AI wordt steeds vaker ingezet in de financiële sector voor taken zoals het voorspellen van markttrends, het beheren van beleggingsportefeuilles en het detecteren van fraude. In deze sector kunnen fouten leiden tot aanzienlijke financiële gevolgen, zowel voor individuen als voor bedrijven. Correctheid is hier cruciaal om ervoor te zorgen dat investeringen veilig zijn en dat financiële beslissingen gebaseerd zijn op nauwkeurige gegevens en analyses.

Onvoldoende duidelijkheid over de onderliggende gegevens van vooraf getrainde modellen, mogelijke vooroordelen en een gebrek aan inclusiviteit vormen juridische en reputatierisico's voor organisaties, net als "hallucinaties" en de kans op gegevenslekken. (Capgemini)

03

Juridische toepassingen

In de juridische sector kan generative AI worden gebruikt voor het opstellen van juridische documenten, het analyseren van jurisprudentie en het bieden van juridisch advies. Onjuiste informatie of interpretaties kunnen leiden tot juridische complicaties, verkeerde beslissingen en zelfs rechtszaken. Daarom is het essentieel dat AI-systemen in deze context nauwkeurig en betrouwbaar zijn om juridische professionals te ondersteunen bij hun werk.

04

Onderwijs en onderzoek

Generative AI kan een waardevol hulpmiddel zijn in het onderwijs en onderzoek, bijvoorbeeld bij het genereren van onderwijsmateriaal, het samenvatten van wetenschappelijke artikelen en het ondersteunen van onderzoeksprojecten. In deze context is correctheid belangrijk om ervoor te zorgen dat studenten en onderzoekers toegang hebben tot accurate en betrouwbare informatie. Fouten in onderwijsmateriaal kunnen leiden tot misverstanden en verkeerde conclusies, wat de kwaliteit van het onderwijs en onderzoek kan ondermijnen.

05

Media en informatieverspreiding

In de media-industrie wordt generative AI gebruikt voor het schrijven van nieuwsartikelen, het creëren van content en het verspreiden van informatie. Correctheid is hier van groot belang om de integriteit van de informatie te waarborgen en de verspreiding van desinformatie te voorkomen. Onjuiste of misleidende informatie kan leiden tot verlies van vertrouwen in de media, publieke "verwarring", politieke misleiding en zelfs maatschappelijke onrust.

4. https://prod.ucwe.capgemini.com/wp-content/uploads/2023/07/GENERATIVE-AI_-Final-Web-1-1.pdf
5. https://www.capgemini.com/insights/research-library/ai-and-the-ethical-conundrum/

Klanten verwachten transparante en eerlijke AI-interacties, met een duidelijke verantwoording. (Capgemini)



Testen, testen en nog eens testen

Correctheid is fundamenteel bij generative AI, vooral in domeinen waar de gevolgen van fouten ernstig kunnen zijn. Het waarborgen van de nauwkeurigheid en betrouwbaarheid van AI-systemen is essentieel om de voordelen van deze technologie te maximaliseren en de risico's te minimaliseren. Door strikte validatie- en testprocedures te implementeren en voortdurend te streven naar verbetering, kunnen we ervoor zorgen dat Generative AI een positieve en betrouwbare bijdrage levert aan onze samenleving. Hieronder noem ik enkele stappen en methoden die vaak worden gebruikt:

1. Vergelijking met de Gouden Standaard:
   Een veelgebruikte methode is het vergelijken van de output van de AI met een "gouden standaard" of referentie. Dit houdt in dat de gegenereerde resultaten worden vergeleken met een set van vooraf goedgekeurde en correcte antwoorden die door experts zijn opgesteld.
2. Dubbel modelleren:
   Dual modelling in AI-tests verwijst naar het gebruik van twee verschillende modellen die samen gebruikt worden tijdens het testproces. Het ene model is meestal het AI-systeem dat wordt getest en het andere is een referentie- of orakelmodel. Het orakelmodel wordt beschouwd als de bron van de waarheid en wordt gebruikt om de output van het geteste AI-systeem te verifiëren.
3. Menselijke evaluatie: Menselijke evaluatoren, vaak experts in het betreffende domein, beoordelen de output van de AI op nauwkeurigheid, relevantie en coherentie. Dit kan helpen om fouten te identificeren die mogelijk niet door automatische methoden worden opgemerkt.
4. Gebruikersfeedback:
   Feedback van eindgebruikers kan waardevolle inzichten bieden in de prestaties van de AI in de praktijk. Gebruikers kunnen fouten rapporteren en suggesties doen voor verbeteringen, wat kan helpen om de AI verder te verfijnen.
5. Technieken voor het beperken van vooroordelen:
   Technieken zoals data augmentation, eerlijkheidsvereisten en algoritmen voor het detecteren van vooroordelen kunnen helpen bij het verminderen van vooroordelen in Generative AI-modellen. Audits en bias-beoordelingen moeten een doorlopend onderdeel van het testproces zijn.
6. Testen in verschillende scenario's:
   Het is belangrijk om de AI te testen in een breed scala aan scenario's en contexten om ervoor te zorgen dat deze robuust en betrouwbaar is. Dit kan inhouden dat de AI wordt blootgesteld aan verschillende soorten input en situaties om te zien hoe goed deze presteert onder verschillende omstandigheden.

Continuous improvement, het voortdurende proces van het analyseren van prestaties, het identificeren van kansen en het aanbrengen van stapsgewijze wijzigingen – helpt het AI-system fit te houden. Dit zorgt ervoor dat de AI up-to-date blijft en kan omgaan met nieuwe uitdagingen en veranderingen in de omgeving.

De rol van domeinexperts

Door deze validatiemethoden te combineren, kunnen ontwikkelaars en onderzoekers de correctheid van Generative AI-systemen effectief beoordelen en verbeteren. Maar de ontwikkelaars en onderzoekers moeten dan wel weten wat correcte informatie is. Dat betekent dat zij domeinkennis moeten hebben. Domeinexperts zorgen daarnaast ervoor dat de AI-toepassingen aansluiten bij de bedrijfsdoelstellingen en -processen. Zij spelen ook een belangrijke rol bij het ontwikkelen van richtlijnen voor verantwoord gebruik van AI. Hun expertise zorgt ervoor dat AI-toepassingen effectief, ethisch en verantwoord worden ingezet.

De rol van domeinexperts bij het ontwikkelen van Generative AI is daarom cruciaal. Ze brengen gespecialiseerde kennis en ervaring in die essentieel zijn voor het ontwerpen, trainen en implementeren van AI-systemen die betrouwbare en nauwkeurige resultaten opleveren. Domeinexperts helpen bij het definiëren van de juiste datasets, het valideren van de output en het waarborgen van de relevantie van de AI-toepassingen voor specifieke bedrijfstakken of vakgebieden.

Wil je weten of een röntgenfoto correct is geïnterpreteerd door de AI, zal je een medisch specialist of onderzoeker nodig hebben om de AI te controleren. Dat kan je niet laten doen door een willekeurige data-analist of prompt-engineer. We denken vaak dat een AI-systeem werk van experts kan overnemen, maar zolang de correctheid van de uitkomsten van AI niet zijn gegarandeerd, zullen experts juist meer betrokken moeten worden.

Content curation

Generative AI doet verbazingwekkende dingen door al uw bedrijfsinformatie te “lezen” en het antwoord (en) op een of meer vragen te vinden. Maar de inhoud die de AI leest moet nauwkeurig zijn om tot een nauwkeurig antwoord te komen. Gecureerde inhoud is goedgekeurde, nauwkeurige en kwalitatieve inhoud en bevat de juiste antwoorden voor de organisatie.

De belangrijkste taak van een curator in het museum bestaat uit onder andere uit het kiezen van de te tonen objecten - ook wel “betekenisvol kiezen” genoemd - voor een tentoonstelling. “Betekenisvol kiezen” is een activiteit die ook bij AI-systemen nodig is. De zogenaamde content curator bepaalt welke informatie relevant, betrouwbaar, compleet, actueel en accuraat – dus: correct en waar – is. Het is juist deze informatie waarmee we de AI willen trainen. Omdat hiermee de correctheid van de resultaten in hoge mate wordt bepaald.

Hoewel contentcuratoren waardevolle bijdragen kunnen leveren, zijn ze geen vereiste om succes te krijgen bij de contentcuratiestrategie . Factoren als kosten, aanwezigheid van interne expertise en het controleren op persoonlijke vooringenomenheid moeten ook in overweging worden genomen bij de beslissing om al dan niet een contentcurator aan te stellen.

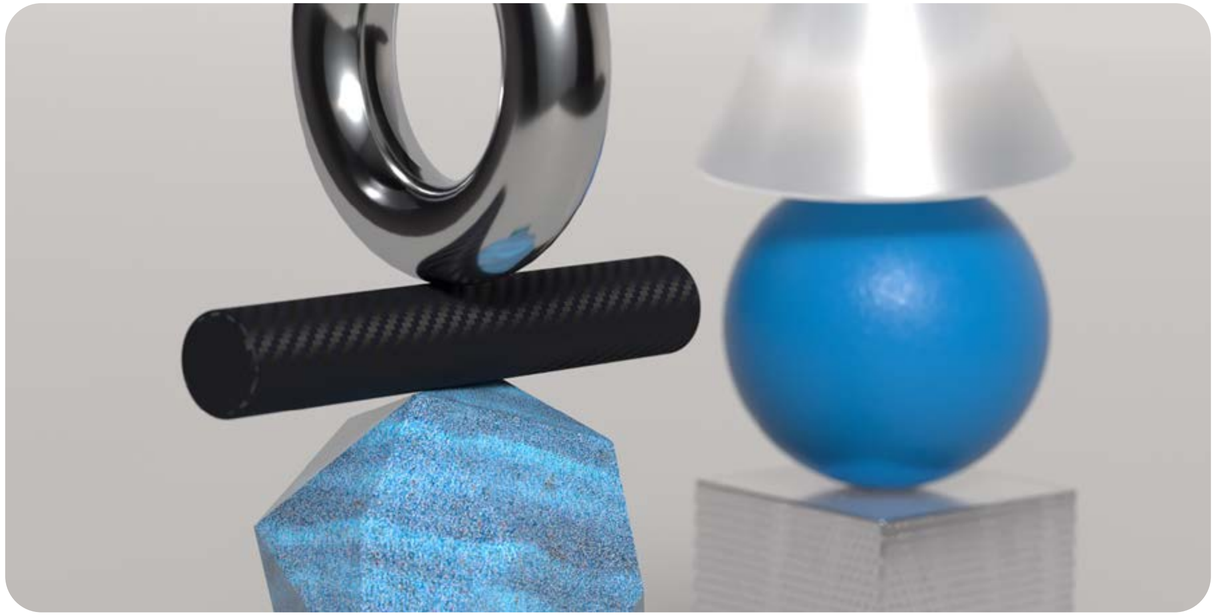
Medewerkers - die al de artikelen en inzichten die u hoopt te gaan ontsluiten - lezen en verwerken, kunnen ook curatiewerkzaamheden doen. Met het juiste proces en de juiste technologieën zijn deze personen in staat om inhoud net zo effectief te beheren als een aparte contentcurator.

Maar ondanks alles blijft content curation een dure en tijdrovende – maar bitter noodzakelijke – aangelegenheid. Alleen met een goede kosten-batenanalyse kan blijken of de investeringen aan de voorkant van de AI kostenvoordelen oplevert aan de achterkant

Altijd correct?

Zal AI ooit perfect worden en foutloos zijn? Het is een interessante vraag. AI-systemen, inclusief Generative AI, zijn ontworpen om taken uit te voeren die normaal menselijke intelligentie vereisen. Hoewel AI aanzienlijk kan verbeteren en steeds geavanceerder wordt, is het onwaarschijnlijk dat AI ooit volledig foutloos zal zijn. Dit komt door verschillende factoren:

1. **Complexiteit van de wereld:** De wereld is extreem complex en veranderlijk, wat betekent dat AI moet kunnen omgaan met een oneindige hoeveelheid scenario’s en variabelen.
2. **Beperkingen van data:** AI leert van data. Als de data onvolledig, vooringenomen of foutief is, kan dit de prestaties van AI beïnvloeden.



3. **Veranderende contexten:** Wat als ‘correct’ wordt beschouwd, kan veranderen afhankelijk van de context, cultuur of tijd, waardoor het moeilijk is voor AI om altijd ‘perfect’ te zijn.
4. **Onvoorspelbaarheid:** Onvoorziene gebeurtenissen of situaties die niet in de trainingsdata zijn opgenomen, kunnen leiden tot fouten. Resultaten uit het verleden bieden geen garantie voor de toekomst.

AI kan worden ontworpen om in veel gevallen uitzonderlijk nauwkeurig te zijn, maar de mogelijkheid van fouten volledig uitsluiten is niet realistisch. De vraag is of dat erg is. Mensen zijn ook niet foutloos. Het gaat te ver hier dieper op in te gaan, maar zolang mensen de uitkomsten van Generative AI “blind” geloven, is er reden tot waakzaamheid. Zeker wanneer AI niet aangeeft wanneer het onzeker is over een uitkomst en gaat bluffen.

Het is belangrijk om AI-systemen te blijven verbeteren en te leren van fouten om hun betrouwbaarheid en effectiviteit te vergroten. Daarom is voortdurende samenwerking tussen AI-ontwikkelaars, domeinexperts en gebruikers essentieel om AI-systemen te optimaliseren en foutmarges te minimaliseren.

Conclusie

Correctheid is een fundamenteel aspect van Generative AI, vooral in domeinen waar de gevolgen van fouten ernstig zijn. Het waarborgen van de nauwkeurigheid en betrouwbaarheid van AI-systemen is essentieel om de voordelen van deze technologie te maximaliseren en de risico's te minimaliseren. Door strikte trainings-, validatie- en testprocedures te implementeren en voortdurend te streven naar verbetering, kunnen we ervoor zorgen dat Generative AI een positieve en betrouwbare bijdrage levert aan onze samenleving.

Voorbeeld: Verbetering van spraakherkenning

Hoe gebruikersfeedback heeft geleid tot verbetering van een AI-systeem bij sspraakherkennings-technologieën, zoals die van virtuele assistenten zoals Siri, Google Assistant en Alexa.

Probleem: In de beginfase hadden veel gebruikers problemen met de nauwkeurigheid van spraakherkenning. De AI-systemen hadden moeite met het begrijpen van verschillende accenten, dialecten en spreekstijlen, wat leidde tot frustratie bij gebruikers.

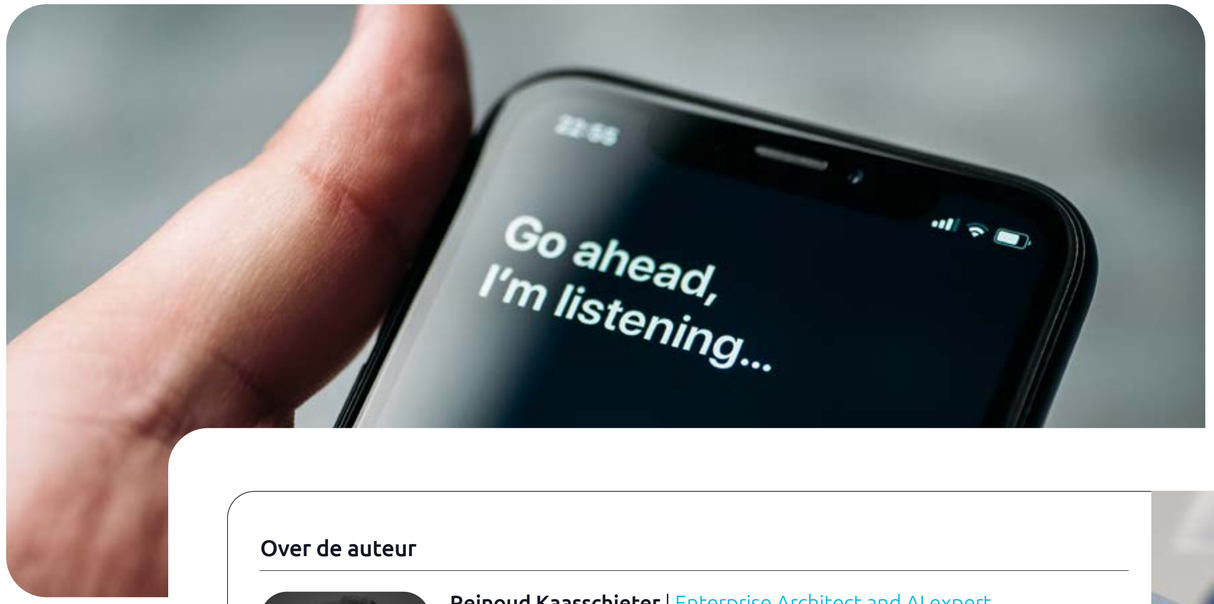
Gebruikersfeedback: Gebruikers begonnen massaal feedback te geven over de problemen die ze ondervonden. Ze rapporteerden specifieke gevallen waarin de AI hun spraak niet correct herkende of verkeerd interpreteerde. Deze feedback werd verzameld via verschillende kanalen, zoals app-beoordelingen, forums, en directe meldingen binnen de apps.

Verbeteringen:

1. Training met diverse data: Op basis van de feedback hebben ontwikkelaars de AI-systemen getraind met een veel meer diverse dataset die verschillende accenten, dialecten en spreekstijlen omvatte. Dit hielp de AI om beter te presteren in real-world scenario's.
2. Aanpassing van algoritmen: De algoritmen werden aangepast om beter om te gaan met variaties in spraak. Dit omvatte verbeteringen in de manier waarop de AI fonetische patronen herkent en verwerkt.
3. Contextuele begrip: Gebruikersfeedback leidde ook tot verbeteringen in het contextuele begrip van de AI. Dit betekent dat de AI beter in staat werd om de context van een gesprek te begrijpen en daardoor nauwkeurigere antwoorden te geven.
4. Continue updates: Op basis van doorlopende feedback blijven de ontwikkelaars de systemen regelmatig bijwerken en verbeteren. Dit zorgt ervoor dat de AI zich blijft aanpassen aan de veranderende behoeften en verwachtingen van gebruikers.

Resultaat

Dankzij deze verbeteringen is de nauwkeurigheid van spraakherkenningstechnologieën aanzienlijk toegenomen. Gebruikers ervaren nu minder fouten en een meer naadloze interactie met hun virtuele assistenten. Dit voorbeeld illustreert hoe waardevolle gebruikersfeedback kan zijn bij het verfijnen en verbeteren van AI-systemen.



Over de auteur

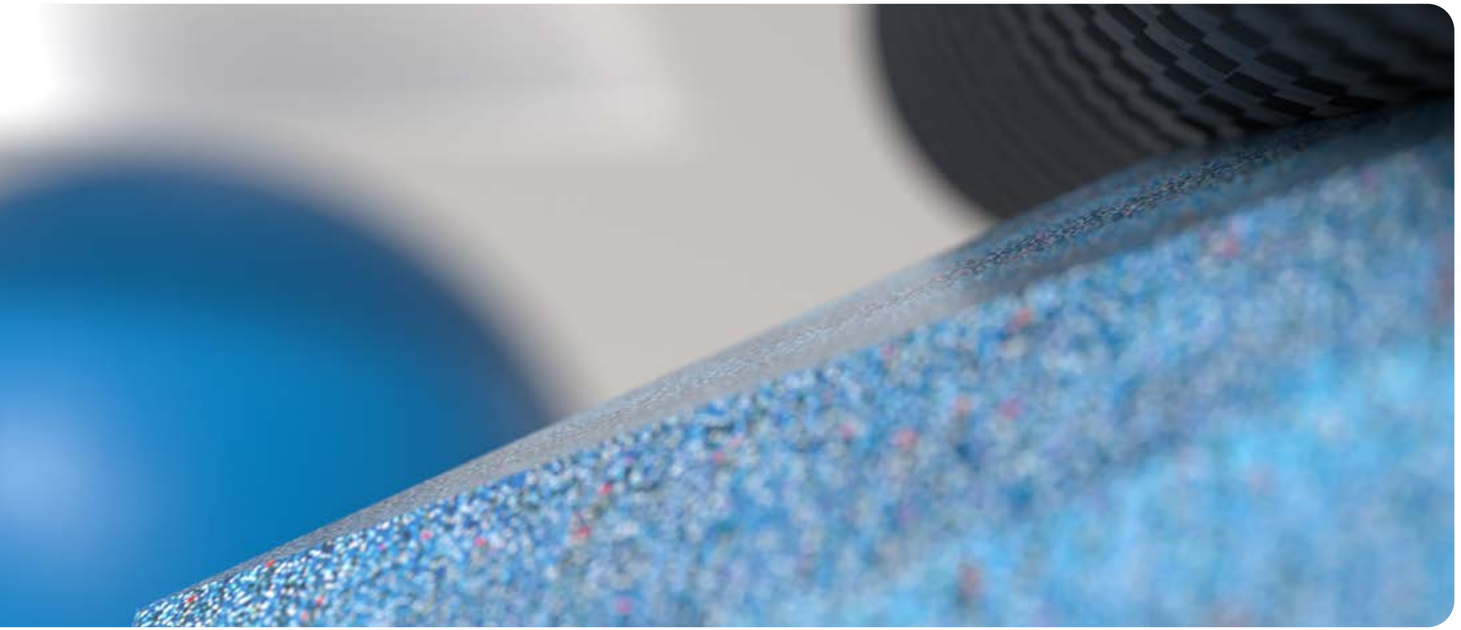


Reinoud Kaasschieter | Enterprise Architect and AI expert
Reinoud is een enterprise architect in de Nederlandse publieke sector en expert op het gebied van kunstmatige intelligentie en ethiek. Hij deelt zijn kennis en inzichten actief door te publiceren en spreken over AI en ethiek, zowel binen Capgemini als daarbuiten.



06 Ready, set, go! Het voorbereiden van klantdata voor *Gen AI in CX*

Welke stappen nu al gezet kunnen worden om een voorsprong te krijgen in de door Gen AI gedreven race naar de CX top



Highlights:

- Gen AI belooft de manier waarop bedrijven klantinteracties vormgeven radicaal te veranderen.
- Een cruciale voorwaarde om die belofte waar te maken is de beschikbaarheid van accurate, gestructureerde en toegankelijke klantdata.
- Bedrijven die nu investeren in hun datakwaliteit, zetten een eerste stap richting een ongeëvenaard concurrentievoordeel in customer experience (CX).
- Een Customer Data Platform (CDP) kan helpen om die voorsprong te pakken.

Hyperpersonalisatie, real-time aanbevelingen en automatisch gegenereerde content zijn geen toekomstmuziek meer, maar worden de nieuwe norm in customer experience (CX). Echter loopt zelfs Gen AI vast als klantdata verspreid, verouderd, of onvolledig is. Bedrijven die hun klantdata nu al op orde brengen, staan op pole position om het potentieel te verwezenlijken.

Met oogkleppen op

Gen AI was het buzzword van 2024 en het belang voor marketing, sales en service wordt al vrij algemeen erkend. We zien met name vrij kleinschalige Proof of Concepts waarbij (Gen) AI wordt toegepast binnen een beperkt vakgebied om de productiviteit van medewerkers te verhogen en kosten te besparen. Gen AI wordt zodoende vaak benaderd ‘met oogkleppen op’: bedrijven foucssen op dat wat direct in het zicht ligt, dus op korte termijn praktisch toegepast kan worden.

Meest gebruikte hedendaagse toepassingen van (Gen) AI binnen CX

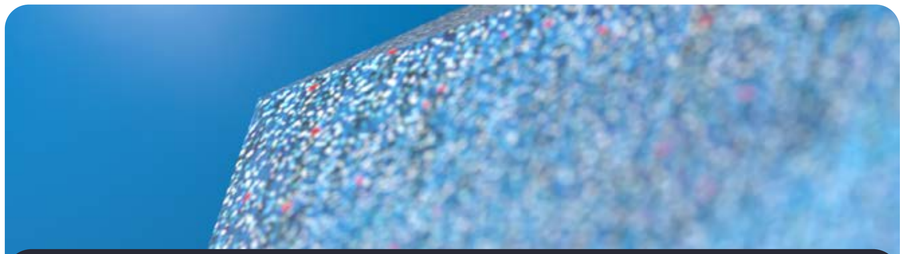
- **Contentcreatie en copywriting:** Van het creeëren van verschillende varianten van productomschrijvingen tot het invullen van de achtergrond van een campagnefoto.
- **Campagneoptimalisatie:** Voorstellen van de meest geschikte momenten op de dag om emails te verzenden, maar ook aanbevelingen om de contactdruk aan te passen.
- **Chatbots:** Met name gebruikt om klantenservicevragen te beantwoorden of om aanbevelingen te doen.
- **Vertaling en transcripties:** Onder andere het vastleggen van klantenservicegesprekken of vertalen van marketingmaterialen voor internationale bedrijven.
- **Sentiment- en data-analyse:** Van het inschatten van de gemoedstoestand van een klant tot het op grote schaal analyseren van data om patronen te herkennen.

De weg verleggen

Deze focus is wellicht te verdedigen vanuit de 70-20-10 ratio voor de innovatiehorizonnen, die voorschrijft om slechts 10% van de aandacht te besteden aan disruptieve innovatie. Toch kan het geen kwaad de oogkleppen af te zetten en eens na te gaan hoe Gen AI gebruiktkan worden om CX disruptief te veranderen. Het wordt dan duidelijk dat de ware kracht van Gen AI ligt in het analyseren van enorme hoeveelheden data en die inzichten vervolgens gebruiken om zelf creatief en adaptatief te zijn. Dat wil zeggen, het genereren van nieuwe content en interacties en die razendsnel aanpassen op basis van experimenten en feedback. Ondanks dat dit op een geaggregeerde manier plaatsvindt, biedt dit enorme mogelijkheden om juist op individueel niveau te personaliseren. Stel je voor dat de toon, boodschap, timing, en kanaal afgestemd worden op de behoeften van een klant, en dat over marketing, sales en service heen. Gen AI is de eerste technologie die een dergelijke datagedreven aanpak met deze schaal en snelheid mogelijk maakt.

Obstakels overwinnen

De waarheid gebiedt te zeggen dat we daar momenteel nog niet zijn. Er zijn nog aanzienlijke technische uitdagingen om de beloofde real-time adaptiviteit te bereiken. Kortom, om zonder vertragingen de enorme hoeveelheden data te kunnen analyseren en daarop te reageren. Daarnaast zijn er begrijpelijkerwijs zorgen over privacy, veiligheid en ethiek, waar bedrijven huiverig zijn voor de strikte regelgeving en de risico's om bias te introduceren in hun modellen. Ook moet er expertise en infrastructuur worden opgebouwd om Gen AI te implementeren en beheren. Maar wellicht de grootste uitdaging, en een die veel over het hoofd gezien wordt, is datakwaliteit en -toegang. Veel organisaties hebben te maken met ongeorganiseerde, onvolledige, of verouderde data. Ze hebben enorme uitdagingen om data te integreren over marketing, sales en service heen waardoor er geen compleet klantbeeld is. Zonder een sterk datafundament kan Gen AI simpelweg niet optimaal presenteren.



Checkbox: Is jouw klantdata klaar voor Gen AI in CX?

- **Relevant:** Gegevens moeten direct bijdragen aan klantinzichten en interactieoptimalisatie
- **Nauwkeurig:** De data moet up-to-date en vrij van fouten en bias zijn
- **Volledig:** Relevante klantkenmerken, gedrag en voorkeuren moeten zo volledig mogelijk worden vastgelegd
- **Toegankelijk:** Data moet makkelijk toegankelijk zijn, zonder silo's of barrières tussen verschillende afdelingen
- **Real-time:** Data moet zo veel mogelijk in real-time worden bijgewerkt in reactie op veranderend klantgedrag
- **Consistent:** Gegevens moeten op een gestandaardiseerde en uniforme manier worden vastgelegd

Een Customer Data Platform als turbocharger

Het goede nieuws is dat de uitdaging van datakwaliteit steeds meer wordt erkend, waardoor CX leveranciers hierop inspringen. O.a. Microsoft, Salesforce en Adobe bieden een CDP (Customer Data Platform) als centraal datafundament binnen hun stack. Een CDP stelt bedrijven in staat om alle klantdata te centraliseren, te verrijken en te gebruiken voor betere interacties over de volledige klantreis. Het bouwt 360-graden klantprofielen op door data uit verschillende bronnen en kanalen te combineren, zoals CRM-systemen, e-commerceplatformen, websites, apps, sociale media en fysieke interacties in winkels. Hierdoor krijgen bedrijven een goed inzicht in wie hun klanten zijn en wat hen beweegt. Vervolgens biedt een CDP uitgebreide integratiemogelijkheden om deze data te kunnen activeren in andere CX tools, bijvoorbeeld in de vorm van gepersonaliseerde emails, productaanbevelingen of real-time push notificaties.

Uit de startblokken

Uiteraard is een CDP geen wondermiddel – ook hier geldt: “Garbage in, garbage out”. Als de gegevens die je invoert van lage kwaliteit, onvolledig of foutief zijn, zullen de resulterende klantprofielen ook onbetrouwbaar of nutteloos zijn. Toch kan een CDP op verschillende manieren helpen om uit de startblokken te schieten voor datakwaliteit:

- Een CDP is bij uitstek geschikt om data over verschillende bronnen heen te verzamelen en consolideren op één centrale plek. Hierdoor verminder je de risico’s op verouderde of dubbele data en zorg je ervoor dat er een enkel systeem beschikbaar is als input voor Gen AI;
- CDP’s normaliseren en standaardiseren klantdata, zodat alle data in een consistent format komt. Gen AI modellen ontvangen hiermee uniforme data en hoeven niet om te gaan met verschillende data formats, wat de betrouwbaarheid en effectiviteit verhoogt;
- Customer Data Platformen zijn gemaakt om data in real-time te ontvangen en verzenden. Hierdoor is er ook actuele data beschikbaar voor Gen AI om snel te reageren op veranderend klantgedrag;
- Een CDP kan enorm helpen bij de uitdaging van identiteitsresolutie (het matchen van meerdere klantprofielen en deze terug herleiden tot dezelfde persoon) en het beheren van klantdata volgens wettelijke voorschriften en privacyregels. Hierdoor wordt duplicatie voorkomen en kan ervoor gezorgd worden dat er alleen geautoriseerde data wordt gebruikt in AI-processen.

En tot slot ook niet onbelangrijk: Een CDP kan ook op korte termijn al waarde opleveren. De use cases beperken zich namelijk niet tot door Gen AI gegenereerde content. Een beter klantinzicht helpt ook om ‘ouderwetse’ human-to-human interacties te verbeteren. Hiermee kan een CDP een mooie brug slaan tussen de innovatiehorizonnen en zorgen voor toekomstbestendigheid én praktische toepasbaarheid op de korte termijn (speciaal voor de besluitvormers met oogkleppen op).

Op volle kracht vooruit

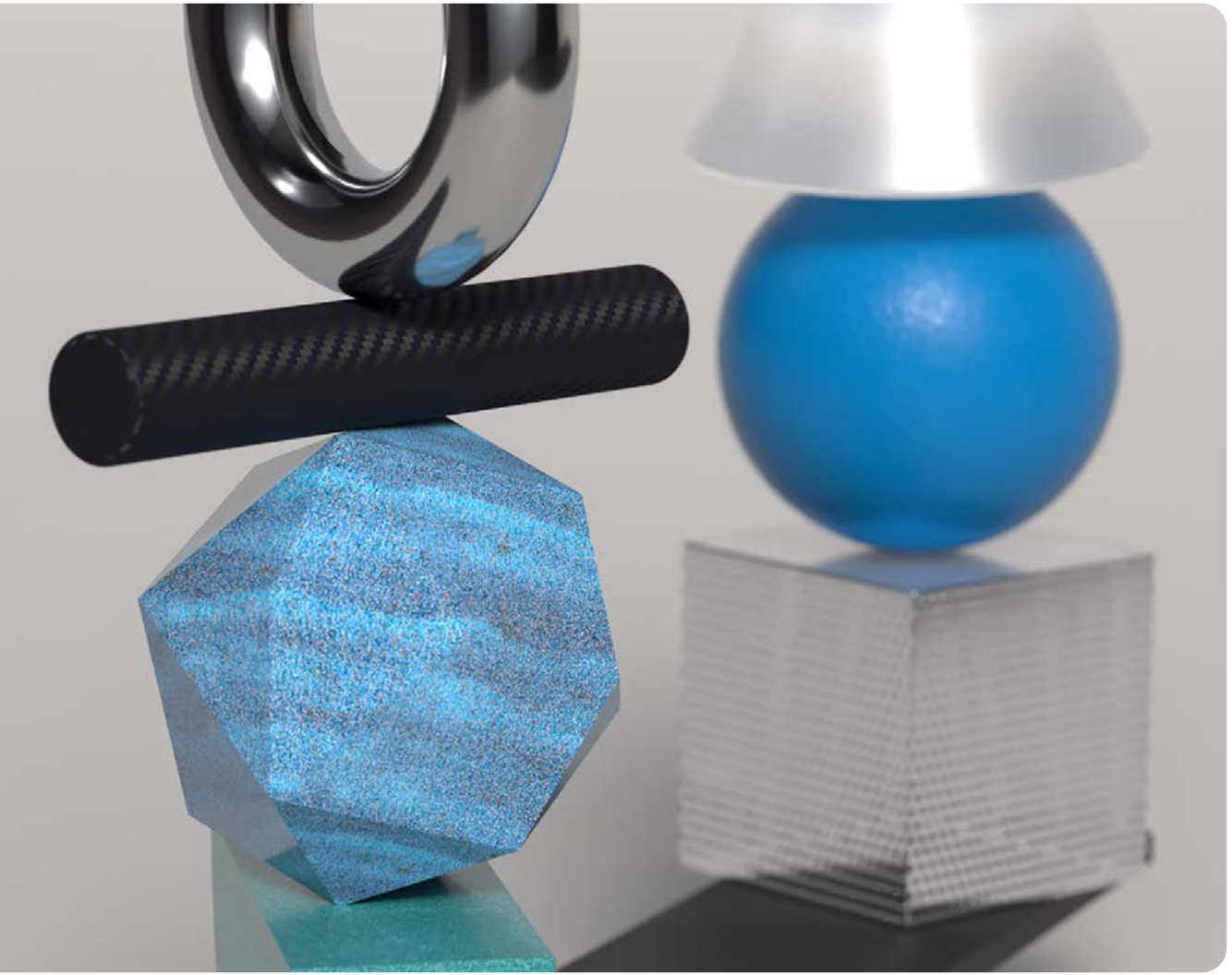
De race naar de top kent altijd obstakels, maar een voorsprong bij de start maakt het verschil. Een sterk klantdatafundament kan die voorsprong bieden voor Gen AI in CX. Bedrijven doen er goed aan om hun klantdata te evalueren op nauwkeurigheid, volledigheid, toegankelijkheid en consistentie. Een CDP kan vervolgens helpen om hierin een verbeterslag te slaan. Zo kunnen bedrijven nu al het noodzakelijke fundament leggen voor Gen AI om gepersonaliseerde klantervaringen te bieden. Tegelijkertijd kan dit ook op de korte termijn al helpen om relevantere interacties te voeren. Dus waarom wachten? Ready, set, go!

Over de auteur



Mandy Steenbergen | Marketing & CX Architect
Mandy is een Marketing & CX Architect binnen de Digital Customer Experience practice. Zij helpt klanten met het ontwerpen en uitvoeren van transformaties om de digitale ervaring te verbeteren.





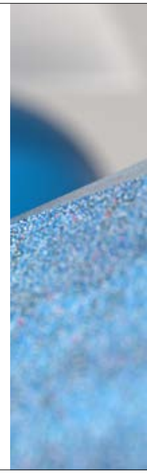
07 Benut de kracht van *Gen AI* voor operationele weerbaarheid en de naleving van regelgeving

Generative AI: bedrijfsvoering versterken met veerkracht en naleving

Bedrijven moeten digitaal innoveren om concurrerend te blijven. Gen AI tools bieden oplossingen om de (digitale) operationele weerbaarheid en naleving van regelgeving te verbeteren. Dit artikel onderzoekt hoe Gen AI bedrijfsvoering kan revolutionaliseren, verstoringen kan voorspellen en aan regelgeving kan voldoen.

Highlights:

- AI-tools en -technieken transformeren bedrijfsprocessen door veerkracht te vergroten en naleving te waarborgen.
- Hoogwaardige data en gestructureerd databeheer zijn cruciaal voor nauwkeurige AI-inzichten.
- Het aanpakken van privacyvraagstukken, bias en transparantie is essentieel voor succesvolle AI-implementaties.
- Gerichte trainingsprogramma's helpen medewerkers effectief gebruik te maken van Generative AI-tools.
- Door uitdagingen op het gebied van data, ethiek en training te overwinnen, kunnen bedrijven een concurrentievoordeel behalen.



Gen AI -tools

Gen AI , als deelgebied van kunstmatige intelligentie, richt zich op het genereren van nieuwe inhoud - tekst, afbeeldingen en code - op basis van bestaande gegevens. Deze tools maken gebruik van machine learning-algoritmen om output te genereren die menselijke creativiteit en probleemoplossend vermogen nabootsen. In de context van bedrijfsvoering kunnen Gen AI tools enorme hoeveelheden gegevens analyseren, patronen herkennen en bruikbare inzichten bieden om besluitvormingsprocessen te verbeteren.

Verbetering van operationele weerbaarheid

Operationele weerbaarheid is het vermogen van een organisatie om verstoringen te anticiperen, zich voor te bereiden, erop te reageren en ervan te herstellen. Deze weerbaarheid is cruciaal voor het handhaven van de bedrijfscontinuïteit en het minimaliseren van de impact van onverwachte gebeurtenissen. In het digitale tijdperk neemt het belang van digitale componenten in operationele weerbaarheid toe. Gen AI -tools verbeteren deze weerbaarheid door voorspellende analyses te gebruiken om zo verstoringen te voorspellen en proactieve maatregelen mogelijk te maken.

Zo gebruikt Alaska Airlines Gen AI om vluchtgegevens, weerspatronen en onderhoudsgegevens te analyseren, waardoor proactief onderhoud en schema-aanpassingen mogelijk zijn. Duke Energy gebruikt Gen AI om infrastructuur te inspecteren, storingen te voorspellen en onderhoud te optimaliseren door weerspatronen en apparatuurprestaties te analyseren.

Voorspellende analyses

Gen AI tools helpen ook bij het analyseren van historische gegevens en het identificeren van trends die mogelijke verstoringen kunnen aangeven. Een voorbeeld hiervan is Walmart, dat Gen AI gebruikt om zijn toeleveringsketen te optimaliseren. Door enorme hoeveelheden gegevens te analyseren, kunnen de AI-systemen van Walmart potentiële verstoringen in de toeleveringsketen voorspellen en alternatieve leveranciers of routes voorstellen. Maar denk ik aan Amazon, dat Gen AI gebruikt om de vraag van klanten nauwkeurig te voorspellen. Door historische verkoopgegevens en markttrends te analyseren, kan Amazon voorspellen welke producten veel gevraagd zullen zijn en de voorraadniveaus dienovereenkomstig aanpassen

Realtime monitoring

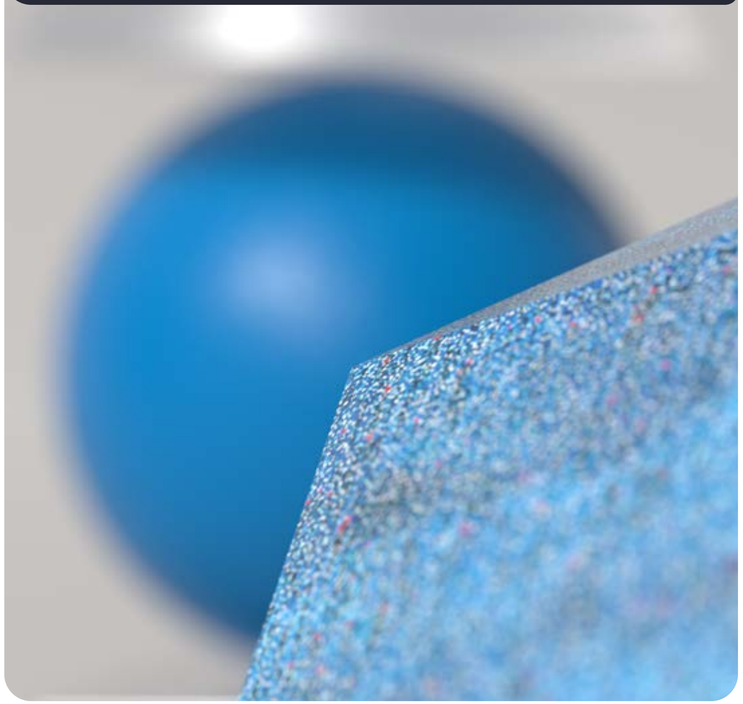
Realtime monitoring is essentieel voor het handhaven van operationele weerbaarheid. Gen AI tools kunnen continu verschillende aspecten van bedrijfsvoering monitoren, zoals productieprocessen, voorraadniveaus en klantinteracties. Door afwijkingen en afwijkingen van normale patronen te detecteren, kunnen signalen van deze tools besluitvormers waarschuwen voor potentiële problemen, waardoor vroegtijdige interventies mogelijk zijn.

In de productie kunnen Gen AI tools productielijnen in realtime monitoren, waarbij metrics zoals machineprestaties, productkwaliteit en productiesnelheid worden gevolgd. Als een afwijking wordt gedetecteerd, zoals een plotselinge daling van de productiesnelheid of een toename van defecte producten, kan het systeem operators waarschuwen om het probleem te onderzoeken en op te lossen, waardoor stilstand wordt geminimaliseerd en de productkwaliteit behouden blijft.

Rolls-Royce gebruikt Gen AI om realtime gegevens van de motoren te monitoren en kan daarmee afwijkingen detecteren en potentiële storingen voorspellen voordat ze optreden. BMW daarentegen heeft AI-gestuurde robots geïmplementeerd in zijn fabriek in Spartanburg om productieprocessen te optimaliseren.

Winkeliers en magazijnen gebruiken Gen AI om continu voorraadniveaus te monitoren. IKEA gebruikt bijvoorbeeld AI-analyse om de vraag naar producten te voorspellen en tijdige aanvulling te garanderen. Dit voorkomt tekorten en zorgt ervoor dat klanten altijd de producten kunnen vinden die ze nodig hebben.

Gen AI tools helpen ook bij het analyseren van historische gegevens en het identificeren van trends die mogelijke verstoringen kunnen aangeven. Een voorbeeld hiervan is Walmart, dat Gen AI gebruikt om zijn toeleveringsketen te optimaliseren.



Geautomatiseerde responsmechanismen

In het geval van een verstoring kunnen Gen AI tools responsmechanismen automatiseren om stilstand te minimaliseren en de bedrijfscontinuïteit te waarborgen. Als een kritisch systeem uitvalt, kunnen deze tools taken automatisch omleiden naar back-upsystemen of vooraf gedefinieerde herstelprocedures starten.

Een recent onderzoek laat zien hoe Siemens Gen AI -tools gebruikt om de operationele weerbaarheid te verbeteren: Het bedrijf gebruikt een geïntegreerde AI-assistent voor veldonderhoudswerkers, dat voorspellende modellen gebruikt om systeemstoringen te voorspellen en reparatie-teams naar de juiste locaties te sturen. Bovendien biedt de AI-assistent on-the-spot kennis over bijvoorbeeld reparatieprocedures met behulp van Gen AI -modellen om stilstand te minimaliseren en de bedrijfscontinuïteit te waarborgen door de noodzaak van handmatige interventie te verminderen.

Zorgen voor naleving van regelgeving

Naleving van regelgeving, compliance, is cruciaal voor bedrijfsvoering, aangezien niet-naleving kan leiden tot zware boetes en reputatieschade. Gen AI tools vereenvoudigen naleving van regelgeving door taken te automatiseren en organisaties op de hoogte te houden van veranderende regelgeving.

Het naleven van nieuwe EU-regelgeving, zoals de EU-richtlijn Directive Corporate Sustainability Due Diligence (CSDDD) in 2024, is hiervan voorbeeld. Deze richtlijn vereist dat bedrijven mensenrechten en milieu-impact in hun operaties en toeleveringsketens aanpakken, waarbij Gen AI een cruciale rol speelt.



EnerSys, een toonaangevende fabrikant van industriële batterijen, gebruikt bijvoorbeeld Gen AI om hun duurzaamheid gegevensverzameling en rapportage te verbeteren. Hun AI-aangedreven platform, ESG Flo, verwerkt gegevens van energierekeningen over 180 locaties wereldwijd, waardoor de nauwkeurigheid en controleerbaarheid van gegevens wordt verbeterd.

De NIS 2-cyberbeveiligingsrichtlijn legt strengere eisen op, waaronder verplichte incidentrapportage en verbeterde beveiligingsmaatregelen. IBM gebruikt zijn Watson AI om bedrijven te helpen voldoen aan cyberbeveiligingsvoorschriften. Dit bereikt het bedrijf door enorme hoeveelheden gegevens te analyseren om daarmee potentiële beveiligingsincidenten te detecteren. Hierdoor kunnen zij ervoor zorgen dat alle verplichte incidentrapporten tijdig worden ingediend.

Een ander voorbeeld is Darktrace, dat AI gebruikt om netwerkverkeer in realtime te monitoren en afwijkingen te detecteren. Deze proactieve benadering stelt bedrijven in staat om potentiële beveiligingsbedreigingen snel te identificeren en erop te reageren, waardoor ze voldoen aan de verbeterde beveiligingsmaatregelen die vereist zijn door NIS 2.

Splunk biedt ook AI-gedreven oplossingen die organisaties helpen hun nalevingsprocessen te automatiseren. De AI-tools van Splunk kunnen continu beveiligingsgegevens monitoren en analyseren, waardoor bedrijven voldoen aan de nieuwste cyberbeveiligingsvoorschriften en incidenten rapporteren zoals vereist.

Deze voorbeelden tonen aan hoe Gen AI -tools al worden gebruikt door toonaangevende bedrijven om nalevingstaken te automatiseren en cyberbeveiligingsmaatregelen te verbeteren.



Regelgeving zoals de Digital Operational Resilience Act (DORA) en MiFID in de EU leggen strikte nalevingsverplichtingen op. Gen AI tools helpen door nalevingsmonitoring te automatiseren, realtime risicoanalyses uit te voeren en naleving van regelgevingskaders te waarborgen.

Zo gebruikt HSBC AI om transacties in realtime te monitoren en verdachte activiteiten zoals witwassen of fraude te detecteren. Evenzo gebruikt JPMorgan Chase Gen AI -tools om nalevingsrapporten te automatiseren en snel onregelmatigheden te identificeren om te voldoen aan regelgeving zoals MiFID II. Het AI-platform van Deutsche Bank volgt regelgevingswijzigingen en past nalevingsstrategieën aan, waardoor de bank voorop blijft lopen en boetes vermijdt

Uitdagingen en strategieën voor implementatie

Hoewel de voordelen van Gen AI tools aanzienlijk zijn, brengt de implementatie ervan uitdagingen met zich mee. Hieronder worden een drietal elementen besproken.

Gegevenskwaliteit en integratie

GenAI tools vereisen hoogwaardige gegevens voor nauwkeurige inzichten, maar het waarborgen van gegevenskwaliteit en integratie is complex. Organisaties moeten robuuste gegevensbeheerpraktijken implementeren en gegevensbeheerstructuren opzetten met duidelijke beleidslijnen en regelmatige audits. Geavanceerde integratietechnieken zoals Extract, Transform, Load (ETL) processen en het gebruik van datalakes vergemakkelijken naadloze gegevensuitwisseling. Het aanpakken van deze uitdagingen zorgt ervoor dat Gen AI -tools effectief werken en waardevolle inzichten leveren.

Ethische overwegingen

Het gebruik van Gen AI -tools roept ethische overwegingen op het gebied van gegevensprivacy, vooringenomenheid en transparantie. Organisaties moeten ethische richtlijnen implementeren en ervoor zorgen dat AI-algoritmen eerlijk, transparant en verantwoordelijk zijn. Regelmatige audits en monitoring kunnen helpen om eventuele vooroordelen in AI-gegenereerde output te identificeren en te verminderen.

Opleiding van personeel

De implementatie van Gen AI -tools vereist een bekwaam personeels bestand. Organisaties moeten investeren in gerichte training voor relevante medewerkers om een effectief gebruik van deze technologieën te

Over de auteurs



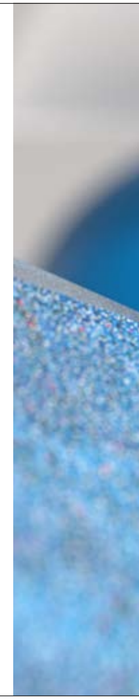
Rahul Rauniyar | Senior Managing Consultant - IT Security Strategy, Risk & Compliance at Capgemini

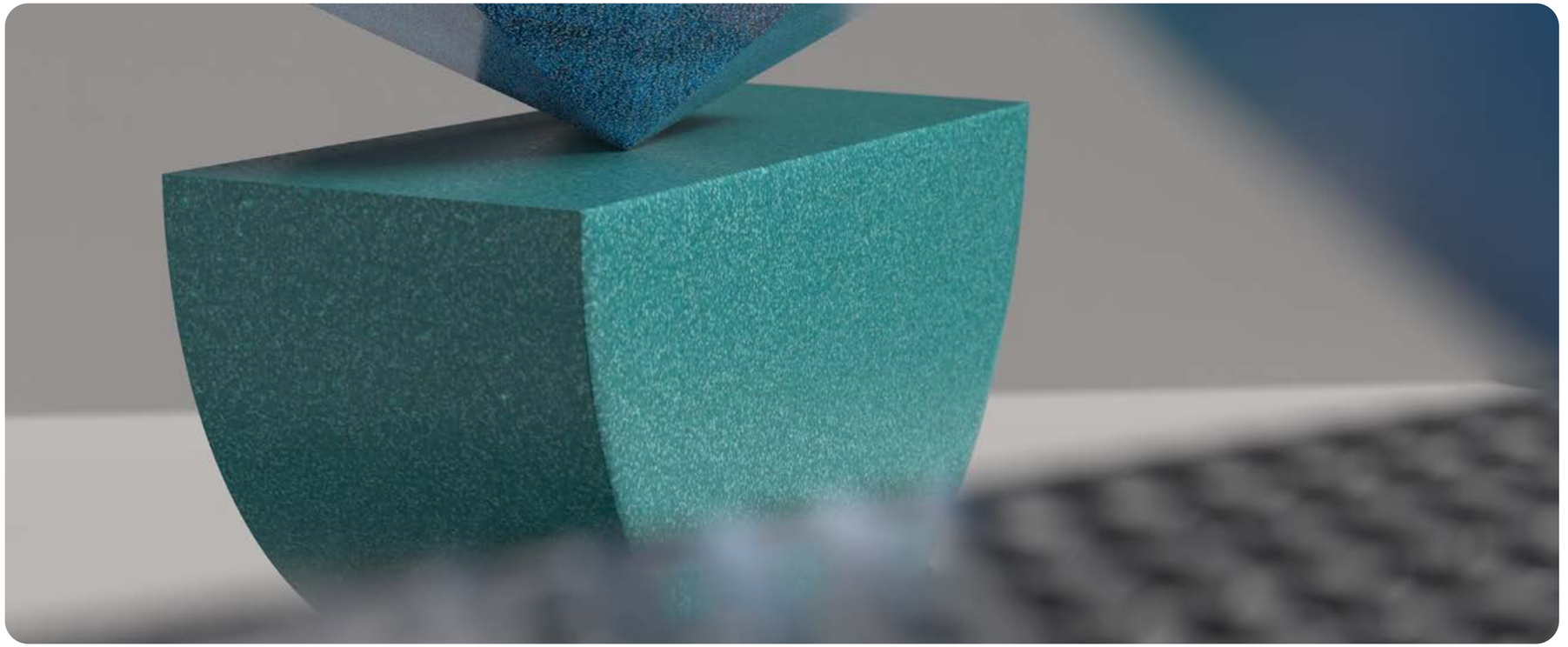
Rahul is een Senior Managing Consultant bij Capgemini NL en leidt het IT security Strategy, Risk & Compliance-team in de financiële sector. Als DORA Offer Lead helpt hij klanten bij het verbeteren van hun digitale veerkracht en compliance. Met Generative AI versnelt hij nalevingsprocessen en versterkt hij beveiliging binnen diverse sectoren.



Marieke van de Putte | Global Portfolio Lead Continuous Compliance | Operational Resilience

Marieke van de Putte is een registercontroller, Cyber Executive en Global en NL Domain Lead voor Continuous Compliance. Ze heeft meer dan 25 jaar ervaring met het adviseren van multinationals op het gebied van security, riskmanagement en compliance. Daarnaast is zij bestuurslid van AIquity in Nederland.





Highlights:

- AI-gedreven robotica transformeert retail met gepersonaliseerde service en efficiëntie.
- RAG combineert AI-modellen met realtime data voor betere productaanbevelingen.
- Multi-agent systemen optimaliseren complexe retailtaken via ‘Chain of Thought’ prompting.
- Praktijkvoorbeeld: Robot als interactieve winkelassistent bij grote Nederlandse retailers.
- Toekomst retail: Synergie tussen geavanceerde AI-oplossingen en menselijke interactie.

08 De Retail Revolutie met *Generative AI* - gedreven Robotica

Van gepersonaliseerde aanbevelingen tot efficiënte automatisering, in de snel veranderende wereld van retail wordt de traditionele winkelervaring steeds meer beïnvloed door technologische innovaties. Eén van de meest spannende ontwikkelingen in dit veld is de integratie van AI-gedreven robotica, die niet alleen de manier waarop retailers met hun klanten communiceren kan transformeren, maar ook een innoverende stap zet voor efficiëntie en gepersonaliseerde service.

Dit artikel onderzoekt de integratie van Generative AI en robotica in de retailsector, met specifieke aandacht voor Retrieval-Augmented Generation (RAG) voor productaanbevelingen, AI en Retail, en het gebruik van multi-agent systemen.

De opmars van AI in Retail

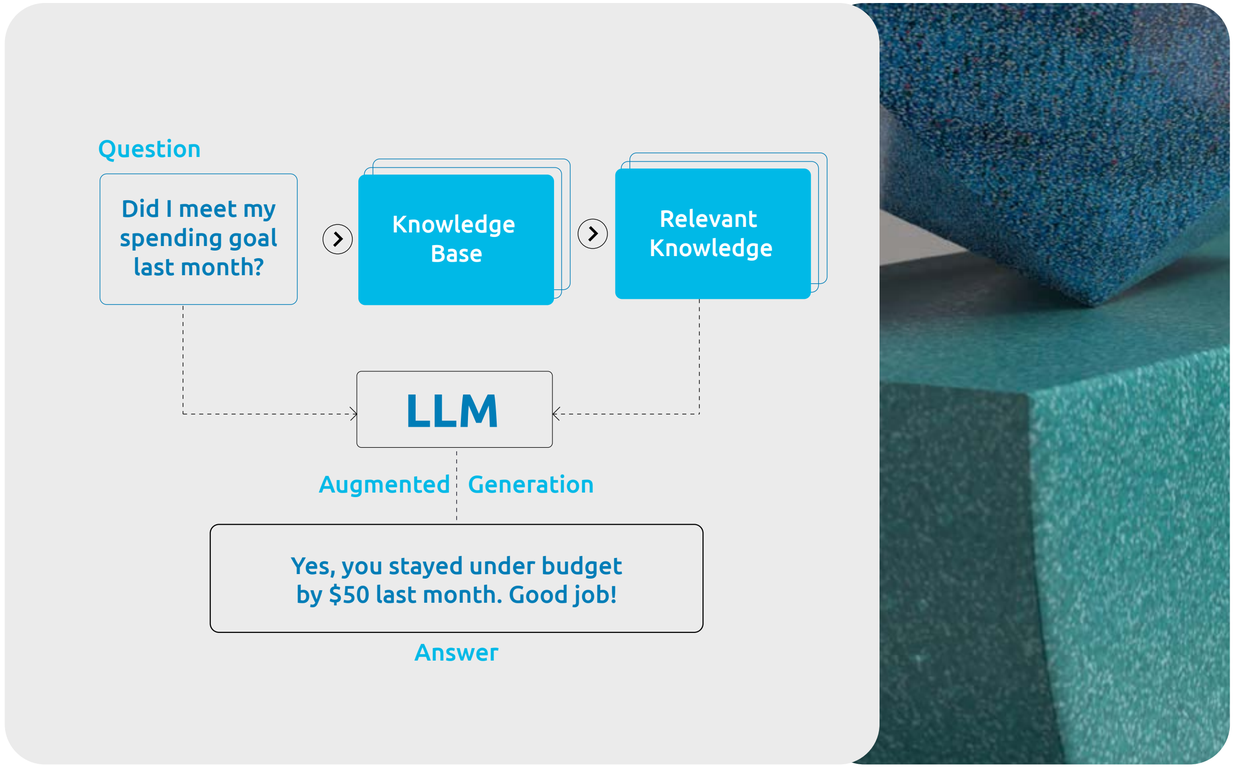
De integratie van AI in de retailsector is geen recent fenomeen. Al in de vroege jaren 2000 begonnen retailers met het gebruik van machine learning om gepersonaliseerde aanbevelingen te doen en voorraadbeheer te optimaliseren. Naarmate de technologie zich ontwikkelde, nam de complexiteit en de kracht van deze toepassingen toe. In de afgelopen jaren heeft de introductie van Generative AI deze trend echter in een stroomversnelling gebracht. Generative AI, die in staat is om creatieve en geavanceerde output te produceren op basis van enorme hoeveelheden data, kan gecombineerd worden met robotica om een revolutie in retail te ontketenen.

Hoewel AI-oplossingen zoals chatbots en voorspellende analyses het nieuwe normaal zijn, is het gebruik van AI-gedreven robotica een relatief nieuwere ontwikkeling. Deze robots zijn niet langer beperkt tot simpele, repetitieve taken; dankzij de integratie van LLM-agents met toegang tot Large Language Modellen zoals Anthropic's Claude 3.5 Sonnet of OpenAI's ChatGPT-4o. Hierbij kunnen robots toegankelijker gebruikt worden als communicatiemedium door de uitgebreide real time contentgeneratie zoals de social robot Nadine. ((Hangyeol Kang, 2024)).

RAG voor Productaanbevelingen

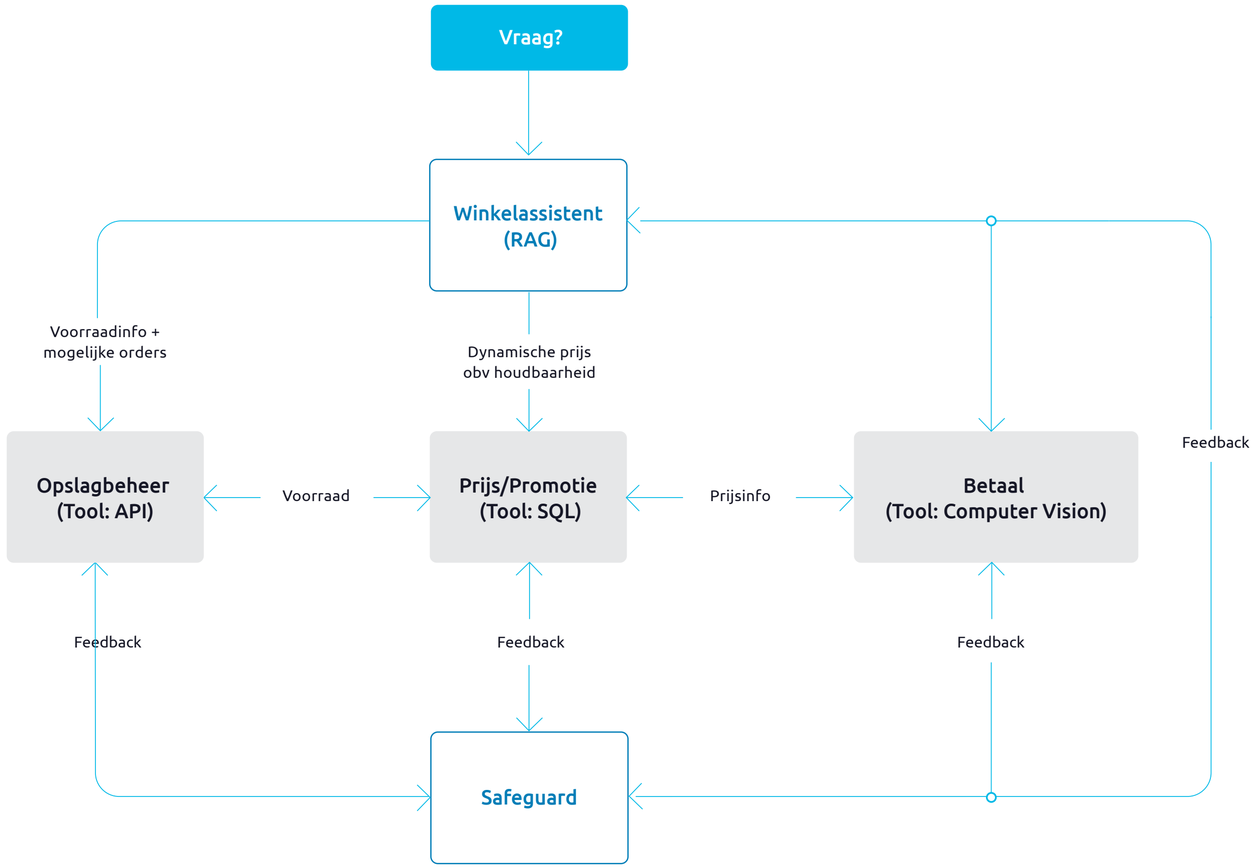
Een van de meest veelbelovende toepassingen van Generative AI in retail is Retrieval-Augmented Generation (RAG). Dit is een techniek waarbij een Gen AI-model relevante informatie ophaalt uit een vector database (retrieval) en deze vervolgens gebruikt om nieuwe content te genereren (generation). In een retailcontext betekent dit dat Large Language-modellen (LLM's) in staat zijn om gepersonaliseerde productaanbevelingen te doen op basis van zowel eerdere interacties van de klant als op basis van realtime beschikbare productinformatie.

Stel je voor dat een klant op zoek is naar een nieuwe winterjas. Een traditioneel AI-aanbevelingssysteem zou zich baseren op de zoekgeschiedenis en eerdere aankopen om soortgelijke producten aan te bevelen. Met RAG kan het systeem echter veel verder gaan. Het kan productinformatie in realtime ophalen, bijvoorbeeld over de nieuwste collectie jassen die net zijn binnengekomen. Vervolgens kan het AI-model een aanbeveling genereren die niet alleen is gebaseerd op wat de klant eerder heeft bekeken, maar ook op wat er momenteel beschikbaar is in de winkel. Dit resulteert in veel relevantere en dus meer gepersonaliseerde aanbevelingen.



(Aquino, 2024) https://qdrant.tech/articles/what-is-rag-in-ai/





Multi-agent Systemen: De sleutel tot geavanceerde retail automatisering?

Naast RAG is er nog een andere technologie die de retailsector fundamenteel kan gaan veranderen: multi-agent systemen (MAS). Een multi-agent systeem bestaat uit meerdere autonome agents met verschillende toolsets die met elkaar communiceren en samenwerken om complexe taken uit te voeren. Deze systemen maken gebruiken van een techniek genaamd ‘Chain of Thought’ prompting, waarbij de agents ‘hardop denken’ en itereren over hun denkproces. Zie het als een multidisciplair team dat complexe taken onderling verdeeld en oplost. Een voorbeeld van een multi-agent systeem binnen de retailsector zou kunnen zijn, een winkelsysteem waarbij de shopping experience en operaties worden geoptimaliseerd. Een winkelassistent agent die contact heeft met de klant (tekst of spraak) voor het vinden van producten, een opslagbeheer agent die kennis heeft over voorraden en plaatst orders wanneer nodig. Een prijs en promotie agent die zich bezighoudt met dynamic pricing en gepersonaliseerde kortingen. Een betaalagent die aan de hand van computer vision, producten kan herkennen en zorgt dat alle producten op de bon staan. En om het process compleet te maken, een feedback agent die het hele proces feedback geeft om te verzekeren dat alles loopt zoals het moet lopen.

Praktische Toepassing: Een Nederlandse Retail Casestudy

Hoe wordt Robotica dan ingezet in Retail? Binnen Capgemini is een demonstrator ontwikkeld voor een use case bij één van Nederlands grootste retailers. In dit project fungeerde een robot als een interactieve winkelassistent met kennis over producten aanwezig bij deze retailer. Gebruikers konden de robot benaderen voor het verkrijgen van cadeau aanbevelingen op basis van budget, de gelegenheid en de voorkeuren van de persoon die het cadeau ontvangt. Daarnaast konden meerdere vragen worden gesteld rondom onderwerpen zoals specifieke productinformatie, winkelroutes en zelfs specifieke promoties.

Er is een tweede use case ontwikkeld voor een andere grote Nederlandse retailer. Hierbij werd de robot ook ingezet als interactieve winkelassistent, maar ook als voedingsexpert. Gebruikers konden vragen stellen rondom allergieën en dieetrestricties/voorkeuren. Zo konden vragen gesteld worden als: “Als ik een notenallergie heb, kan ik <specifiek product> nuttigen?”, “Ik wil een bepaald gerecht maken, wat zou een alternatief product zijn op basis van eerdergenoemde allergie?”

Conclusie

De integratie van Generative AI en robotica in de retailsector markeert een nieuwe fase van technologische innovatie. Door technieken zoals Retrieval-Augmented Generation (RAG) en multi-agent systemen toe te passen, kunnen retailers niet alleen gepersonaliseerde productaanbevelingen doen, maar ook hun voorraadbeheer, prijsstrategie en klantenservice optimaliseren.

Het gebruik van AI-gedreven robotica biedt een unieke kans om niet alleen de winkelervaring te verrijken, maar ook efficiëntie en klanttevredenheid te verhogen. De praktijkvoorbeelden voor Nederlandse retailers tonen aan dat deze technologieën niet langer toekomstmuziek zijn.
De toekomst van retail ligt in de synergie tussen geavanceerde AI-oplossingen en menselijke interactie, waarmee een nieuwe standaard wordt gezet voor zowel gemak als innovatie.

Referenties

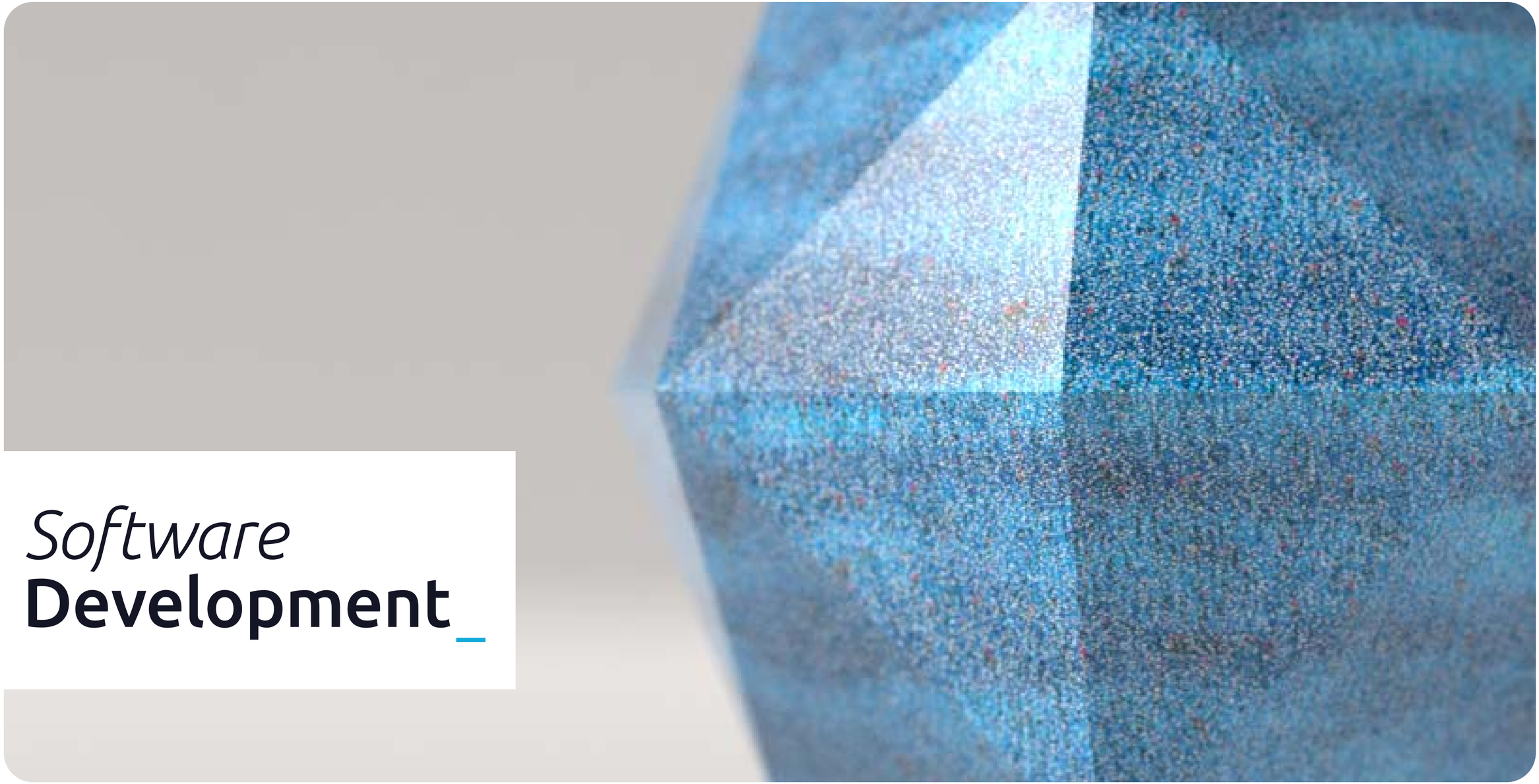
- **Aquino, S. (2024, March 19). What is RAG:** Understanding Retrieval-Augmented Generation. Retrieved from Qdrant: https://qdrant.tech/articles/what-is-rag-in-ai/
- **Hangyeol Kang, M. B.-T. (2024). Nadine:** An LLM-driven Intelligent Social Robot with Affective Capabilities and Human-like Memory.

Over de auteur



Randy van Breugel | Data Scientist, AI-Engineer,
Randy is een Data Scientist binnen Capgemini Engineering’s Digital Manufacturing practice. Hij specialiseert zich in de implementatie van Generative AI-oplossingen voor de retailsector en richt zich daarnaast op complexe data vraagstukken bij klanten in de productie- en farmaceutische industrie.



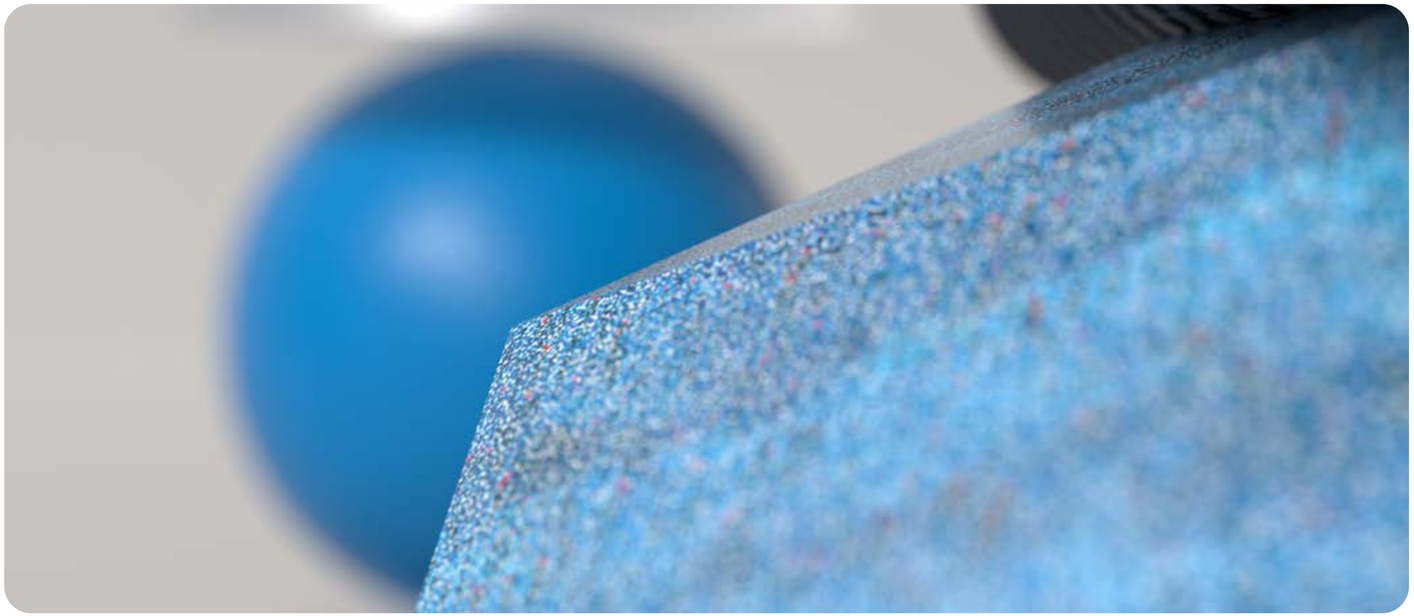


Software **Development**_

09

Het gebruik van *GenAI* om legacy software te verbeteren

Een nieuwe aanpak om legacy software te moderniseren door de kracht van AI te combineren met traditionele tools. Het project van Capgemini Engineering, TNO-ESI en Philips toont hoe Large Language Models de toekomst van softwareonderhoud transformeren.



Highlights:

- Onderhoud van legacy code is een kritieke uitdaging.
- LLMs in combinatie met graph databases bieden een revolutionaire aanpak voor code-analyse en -begrip.
- Ons hybrid systeem maakt natural language querying van complexe codebases mogelijk, wat de efficiëntie drastisch verbetert.
- Deze technologie heeft vooral impact in hoogwaardige industrieën zoals medische apparatuur fabricage.
- De toekomst van software engineering ligt in de synergie tussen AI en traditionele analysetools.

Gen AI ’s beïnvloeden positief de manier waarop we met software omgaan. Door AI te combineren met traditionele hulpmiddelen, transformeren we de aanpak van technical debt en codecomplexiteit in verschillende industrieën, waardoor softwareonderhoud efficiënter en toegankelijker wordt.

De uitdaging van legacy code

In de snelle wereld van technologie maakt legacy software een unieke paradox zichtbaar. Het is de ruggengraat van belangrijke systemen in verschillende industrieën, van gezondheidszorg tot financiën, en belichaamt jaren van opgebouwde kennis en intellectueel eigendom. Toch is het ook een groeiende last, die steeds complexer onderhoud vereist en enorme middelen verbruikt.

De omvang van het probleem

- Op dit moment neemt softwareonderhoud veel van de ontwikkelingscyclus in beslag [1]
- Gartner voorspelt dat tegen 2025 de technical debt verder zal oplopen. Dit zet IT-budgetten onder druk [2], [3]
- In de medische apparatuur productie is legacy software cruciaal voor functionaliteit en veiligheid. Daarentegen is deze moeilijk te onderhouden.

Dit is niet alleen een technisch probleem. Het kan tevens van invloed zijn op belangrijke uitdagingen binnen bedrijven, zoals innovatie en het concurrentievermogen.

Het innovatieve gebruik van LLM’s in RENAISSANCE

Capgemini Engineering, TNO-ESI en Philips zijn pioniers in een revolutionaire aanpak van dit structurele probleem. Door de kwaliteiten van Large Language Models (LLM’s) optimaal te benutten wordt niet alleen de legacy code onderhouden, maar wordt de technologie verder gebracht.

De kracht van AI in codeanalyse

LLM’s zoals GPT-4 en Code-Llama hebben opmerkelijke resultaten getoond in het begrijpen en genereren van code. Het onderzoek van Capgemini Engineering, TNO-ESI en Philips heeft aangetoond dat deze modellen in staat zijn om:

1. Nauwkeurige architectuurdiagrammen te genereren uit ruwe broncode
2. Gedetailleerde UML-diagrammen te creëren, die visuele representaties van complexe systemen bieden
3. Codepatronen in natuurlijke taal te interpreteren en uit te leggen
4. Code te transformeren en documentatie te genereren

De statistische aard van LLM’s kan echter soms tot fouten leiden. Dit neemt een risico met zich mee. Een risico dat in de gezondheidszorg zo veel mogelijk gemeden dient te worden. Binnen deze sector zou een veelvoorkomende fout, zoals een LLM die ‘hallucineert’, of simpelweg gegevens ‘verzint’ dodelijke gevolgen kunnen hebben. Daarom is het van belang om veiligheidsmaatregelen te nemen.

Hybrid intelligence: Het beste van twee werelden

Onze oplossing is een hybride benadering die de sterke punten van traditionele statische analysetools combineert met de intuïtieve kracht van LLM’s. Het gaat als volgt te werk:

01

Code als data:

We zetten complexe codebases (met name C++) om in een AST (abstract syntax tree) door middel van graph databases. Het opstellen van queries voor een graph database is echter een complexe aangelegenheid en uitdagend. Deze aanpak legt de ingewikkelde relaties en afhankelijkheden binnen de code vast en maakt deze bevraagbaar.

02

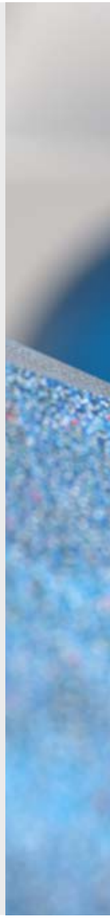
LLM’s als tolken:

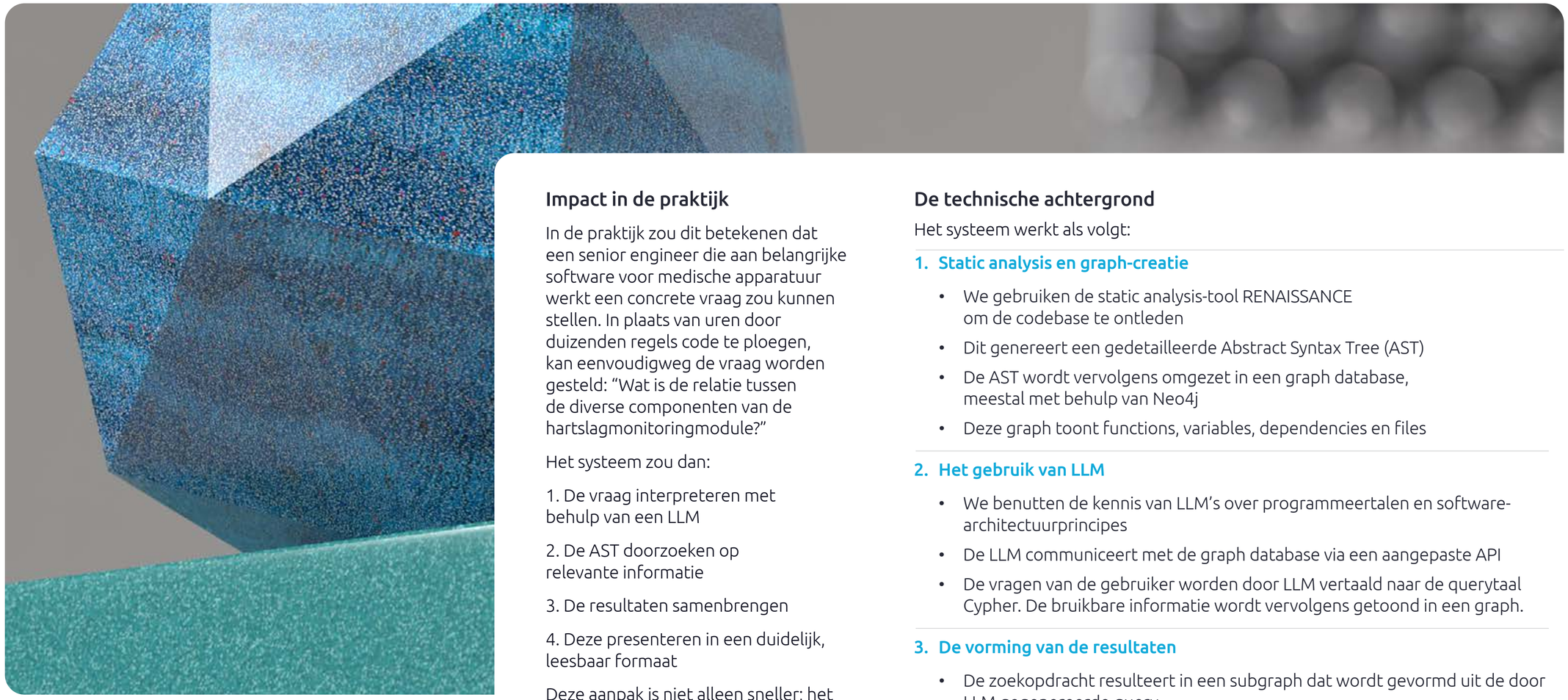
We gebruiken LLM’s als ‘tussenpersoon’ tussen ontwikkelaars en de AST. Ze begrijpen zowel natuurlijke taalvragen als codestructuren, waardoor intuïtieve interactie met complexe systemen mogelijk wordt.

03

Deterministische algoritmen en menselijke ogen voor verificatie:

De gecreëerde kracht van AI wordt gevormd door deze te combineren met RENAISSANCE, gespecialiseerde algoritmen om de resultaten te controleren en een visuele output genereren. Op deze manier kunnen gebruikers snel antwoorden toetsen op feiten.





Impact in de praktijk

In de praktijk zou dit betekenen dat een senior engineer die aan belangrijke software voor medische apparatuur werkt een concrete vraag zou kunnen stellen. In plaats van uren door duizenden regels code te ploegen, kan eenvoudigweg de vraag worden gesteld: “Wat is de relatie tussen de diverse componenten van de hartslagmonitoringmodule?”

Het systeem zou dan:

1. De vraag interpreteren met behulp van een LLM

2. De AST doorzoeken op relevante informatie

3. De resultaten samenbrengen

4. Deze presenteren in een duidelijk, leesbaar formaat

Deze aanpak is niet alleen sneller; het verbetert de effectiviteit van het proces. Door de ingenieurs te voorzien van advies kan deze zich concentreren op innovatie in plaats van verzeild te raken in grote hoeveelheden code-archeologie.

De technische achtergrond

Het systeem werkt als volgt:

1. Static analysis en graph-creatie

- We gebruiken de static analysis-tool RENAISSANCE om de codebase te ontleden
- Dit genereert een gedetailleerde Abstract Syntax Tree (AST)
- De AST wordt vervolgens omgezet in een graph database, meestal met behulp van Neo4j
- Deze graph toont functions, variables, dependencies en files

2. Het gebruik van LLM

- We benutten de kennis van LLM’s over programmeertalen en software-architectuurprincipes
- De LLM communiceert met de graph database via een aangepaste API
- De vragen van de gebruiker worden door LLM vertaald naar de querytaal Cypher. De bruikbare informatie wordt vervolgens getoond in een graph.

3. De vorming van de resultaten

- De zoekopdracht resulteert in een subgraph dat wordt gevormd uit de door LLM gegenereerde query
- Een visualisatie van de subgraph wordt aan de gebruiker gepresenteerd. Daarnaast voorziet het systeem in een antwoord en uitleg.

Overwegingen en toekomstige richtingen

Terwijl we de grenzen van AI in software-engineering verleggen, is het van belang tevens stil te staan bij de implicaties die hiermee gepaard gaan.

- Baanverdringing:
  Het systeem maakt codeonderhoud efficiënter. Het gaat hier echter geenszins om het vervangen van ingenieurs door middel van deze technologie. In plaats daarvan ligt de focus op het vergroten van hun capaciteiten. Bovendien neemt deze technologie hen werk uit handen zodat zij zich kunnen concentreren op meer creatieve taken.
- Gegevensbeveiliging:
  Bij de omgang met bedrijfseigen code, in het bijzonder in gevoelige industrieën, is gegevensbescherming van groot belang. Ons systeem is ontworpen met robuuste beveiligingsmaatregelen om de vertrouwelijkheid van code te waarborgen.
- AI-bias:
  Capgemini Engineering, TNO-ESI en Philips onderzoeken actief manieren om vooroordelen die LLM's kunnen introduceren tijdens het proces te detecteren en te beperken.

Voor de toekomst onderzoeken de drie partners:

- Het veelvuldiger gebruiken van LLM's in de softwareontwikkelingscyclus
- Het uitbreiden van RENAISSANCE om zonder problemen meerdere programmeertalen naast elkaar te gebruiken
- Het ontwikkelen van door AI ondersteunende code-refactoring tools die automatisch verbeteringen kunnen voorstellen en implementeren
- Geautomatiseerd testen om ons systeem continu te ontwikkelen en onze prestaties te verbeteren.

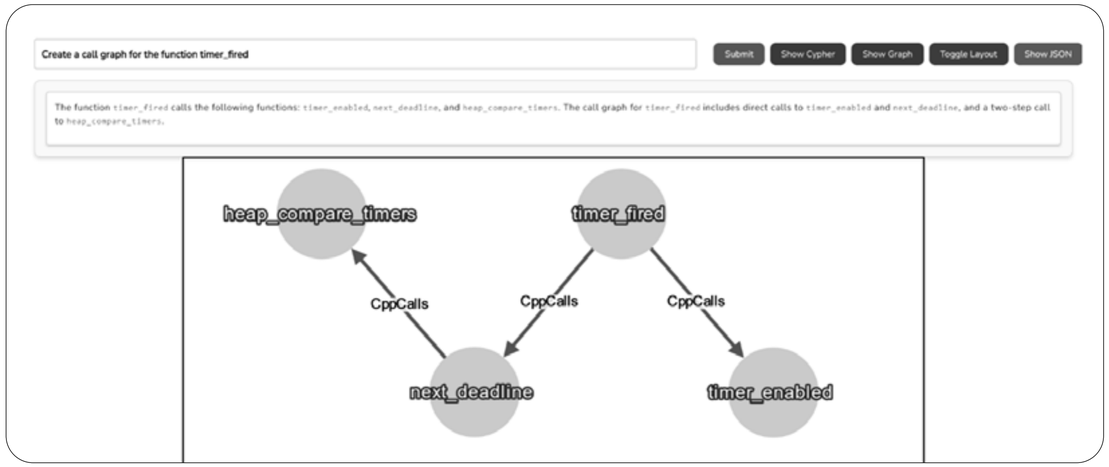
De weg vooruit

De fusie van LLM's met traditionele software-engineering tools lijkt een kleine verbetering. Daarentegen is het een grote paradigmaverschuiving. Terwijl RENAISSANCE zal worden verfijnd en de mogelijkheden van deze technologie wordt uitgebreid, zullen de grenzen van softwareonderhoud en -ontwikkeling worden verlegd.

Deze aanpak heeft de potentie om...

- De kosten en tijd die gepaard gaan met het onderhoud van legacy systemen drastisch te verminderen
- De levensduur van belangrijke softwaresystemen te verlengen
- De kloof tussen oude codebases en moderne innovaties te overbruggen
- Het begrip van code algemeen bekender te maken, waardoor complexe systemen toegankelijk worden voor meer professionals

Concluderend toont de ontwikkeling van RENAISSANCE door Capgemini Engineering, TNO-ESI en Philips dat we op de drempel staan van een nieuw tijdperk in software-engineering, met uitgebreide mogelijkheden en kansen. De partners onderhouden niet alleen code, maar ontketenen het volledige potentieel van decennia van softwareontwikkeling; stap voor stap, één query tegelijk.



Over de auteur



Joe Reynolds | Data Scientist

Als data scientist is Joe gespecialiseerd in het toepassen van machine learning-technieken om bestaande problemen op vernieuwende manieren op te lossen. Zijn achtergrond in natuurkunde en zijn levenslange interesse in computing hebben hem een passie gegeven voor het effectief inzetten van nieuwe technologieën.



10 Versla het software legacy beest met *R.E.B.O.R.N.*

Hou je code up-to-date terwijl je teams zich concentreren op wat echt belangrijk is: het ontwikkelen van geweldige software!

De uitdagingen van software ontwikkelteams

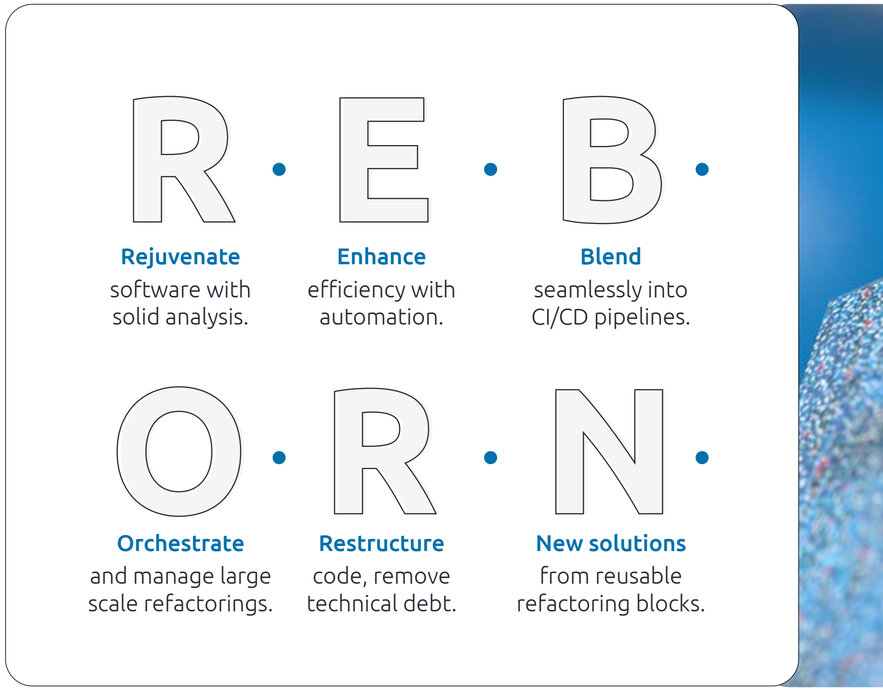
Stel je een software ontwikkelteam voor dat volledig in beslag wordt genomen door het leveren van functies en het oplossen van kritieke bugs. Zonder tijd of focus om hun software bij te werken naar de nieuwste architectuur-, kwaliteits- en technologiestandaarden. Het structureel uitstellen van dergelijke software-evolutietaken zal echter onvermijdelijk hun software in legacy veranderen. Legacy software is berucht inefficiënt om uit te breiden met nieuwe functionaliteit, waardoor de innovatiesnelheid van een bedrijf wordt vertraagd en legacy software direct een bedrijfsrisico wordt.

De R.E.B.O.R.N. oplossing

R.E.B.O.R.N. is een oplossing die software-evolutie holistisch faciliteert gedurende de ontwikkelingscyclus. Onderdeel van R.E.B.O.R.N. zijn de serviceteams die naadloos integreren in klantorganisaties en zorgen voor software-analyse en refactoring om te zorgen dat software aan moderne standaarden blijft voldoen. Automatisering is cruciaal voor efficiënte analyse en refactoring van grote codebases, waarvoor tools nodig zijn die zijn afgestemd op specifieke technologieën, klantomgevingen en evolutie vraagstukken. Met behulp van automatisering kunnen de R.E.B.O.R.N. serviceteams analyse en refactoring op grote schaal uitvoeren, iets wat anders praktisch onmogelijk zou zijn als dit handmatig gedaan zou worden. Tegelijkertijd ontlast dit de ontwikkelteams, waardoor zij zich kunnen concentreren op hun kerntaken.

De R.E.B.O.R.N. Engine

De R.E.B.O.R.N. Engine biedt een framework voor het ontwikkelen en integreren van deze aangepaste automatiseringstools. Het bevat krachtige bibliotheken en frameworks zoals TNO-ESI's Renaissance, wat de analyse en refactoring van C/C++ code versnelt. Renaissance haalt kennis uit broncode en publiceert deze in een graph database, zodat deze op een efficiënte manier kan worden gequeryd en gevisualiseerd.. Bovendien biedt Renaissance een API voor schaalbare, geautomatiseerde refactoring. Het is fundamenteel gebaseerd op bekende en bewezen compilertechnologieën (bijv. direct opvragen/herschrijven van de Abstract Syntax Tree) om codepatronen met precisie en determinisme te extraheren, te vinden en te vervangen.



Software ontwikkelaars versnellen met automatisering

De R.E.B.O.R.N. Engine, geïntegreerd in de CI/CD-omgeving van de ontwikkelaar, ondersteunt software ontwikkelaars door repetitieve, tijdrovende en foutgevoelige taken te automatiseren.

Analyse taken

Analysetaken worden versneld door de geautomatiseerde reverse-engineering van de geïmplementeerde softwarearchitectuur. Hoe is de software gestructureerd? Wat zijn de componenten en hun interface-afhankelijkheden? Waar in de code bevindt zich technical debt of worden verouderde technologieën gebruikt? Dergelijke vragen worden beantwoord door de benodigde structurele en statische kennis uit beschikbare code- en architectuurartefacten te extraheren en deze kennis in een graph database te publiceren. Uiteindelijk kan deze kennis eenvoudig worden opgevraagd en adequaat worden gevisualiseerd met rapporten en visualisaties, zodat de softwarecode efficiënt kan worden geanalyseerd.

Code refactoring taken

Refactoring taken zijn in feite het vinden en vervangen van code patronen in software broncode. Bijvoorbeeld, het oplossen van code smells, het migreren van code naar nieuwe API's, het bijwerken van code om verbeterde inter-process communicatie frameworks te gebruiken, het migreren van testcases naar nieuwe test frameworks en het verhogen van de codekwaliteit. Taken die normaal gesproken handmatig worden uitgevoerd, maar nu met grotere efficiëntie en kwaliteit.

Reële impact

Architecten en ontwikkelaars van onze klanten in de hightech-industrie geven aan dat ze het prettig vinden om te worden ontzorgd van deze analyse- en refactoring taken. Ze waarderen de waardevolle tijd die ze daarmee besparen, die ze nu kunnen besteden aan waardevoller werk met meer impact.

Vanuit het perspectief van het management van de klant zien ze de last en dreiging van legacy software verdwijnen met meer focus, toewijding en snelheid. En zo verlengen ze de levenscyclus van de software waarin ze over de decennia zwaar hebben geïnvesteerd. Deze software wordt nu behouden als hun erfgoed, in plaats van een bedrijfsrisico te worden.

Conclusie

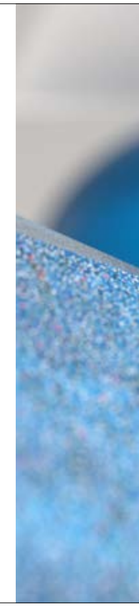
R.E.B.O.R.N.-teams en de R.E.B.O.R.N. Engine vertegenwoordigen een aanzienlijke vooruitgang op het gebied van software legacy en modernisering van software. Door software ontwikkelaars uit te rusten met krachtige automatiseringstools, kunnen we het legacy software beest verslaan en de weg effenen voor een meer innovatieve en efficiënte toekomst.

Over de auteur



Niels Brouwers | Solution Architect MDE & Automated Software Rejuvenation

Niels heeft een master in technische informatica en is Solution Architect bij Capgemini. Hij heeft een sterke passie voor het helpen van organisaties bij het versnellen van hun softwareontwikkeling. Dit wordt bereikt door het abstractieniveau te verhogen en het automatiseringsniveau te intensiveren bij het ontwikkelen van software. Enerzijds door het hanteren van een model gedreven software engineering aanpak, anderzijds met het het extraheren van modellen uit code bases voor de analyse en refactoring van grote hoeveelheden software broncode.



11 *OpenRewrite:* Automatiseren van Code Refactoring voor Moderne Applicaties

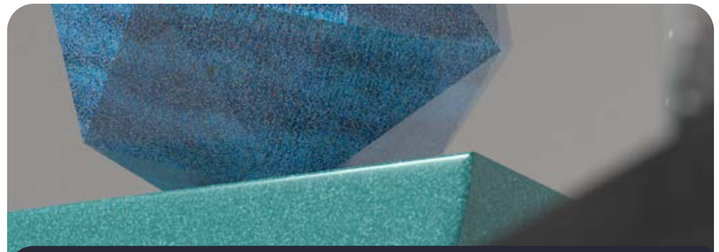
Transformeer uw Java CodeBase met krachtige, declaratieve en veilige refactoring-recipes

In onze softwaregedreven wereld is automatisering een integraal onderdeel van de meeste oplossingen. Toch gaat niet alles vanzelf. Software-ingenieurs en architecten staan nog steeds voor uitdagingen bij het bijwerken van afhankelijkheden, het toepassen van beveiligingspatches en het upgraden van frameworks. Veel bedrijven worstelen met legacy code; daardoor is technical debt een constante bron van zorg. Het bijwerken van afhankelijkheden in meerdere build-bestanden en het omgaan met grote framework-upgrades vereist vaak dat ontwikkelaars handmatig class files, eigenschappen en build-bestanden zoals pom.xml of build.gradle beoordelen en aanpassen.

Wat is OpenRewrite?

OpenRewrite is een geautomatiseerd refactoring-ecosysteem dat specifiek is gemaakt om de gevolgen van grootschalige broncode-refactoring aan te pakken. Het biedt een breed scala aan recipes die automatisch Java-code en verschillende build-configuratiebestanden, waaronder pom.xml en build.gradle, kunnen aanpassen. Als open-source initiatief faciliteert het semantische analyse en code-refactoring, waardoor gebruikers in hun vrije tijd als opensource-contributor aan het OpenRewrite-project kunnen deelnemen en bijdragen. Deze tool stelt individuele ontwikkelaars in staat om code binnen een enkele repository te zoeken en te corrigeren.

Bovendien integreert OpenRewrite naadloos met build-tools en kan het worden opgenomen in diverse workflows, variërend van continue integration setups tot uitgebreide refactoring over meerdere microservices en repositories.



Hoe werkt OpenRewrite?

OpenRewrite genereert een geavanceerde versie van de Abstract Syntax Tree (AST): de Lossless Semantic Tree (LST). Deze LST biedt diepgaande structurele informatie, terwijl de oorspronkelijke formatting van de code behouden blijft. Daardoor kunnen ontwikkelaars:

- Code wijzigen terwijl consistente opmaak behouden blijft
- Metadata (bijv. build-markeringen) aan tree-elementen toevoegen voor dieper inzicht
- Recipes toepassen voor de refactoring van code in zowel bron- als build-bestanden

OpenRewrite in Actie

De integratie van OpenRewrite is eenvoudig. Ontwikkelaars kunnen het in hun build-proces opnemen met behulp van Maven- of Gradle-plugins. Zodra de plugin is geconfigureerd, kunnen OpenRewrite-recipes direct op het project worden toegepast.

Toepassingen in de Praktijk

OpenRewrite kan worden gebruikt voor verschillende bedrijfsapplicaties zoals:

- **Upgraden van Spring Boot:** Naadloos migreren tussen grote versies van Spring Boot
- **Junit-versies wisselen:** Testcases automatisch bijwerken van Junit4 naar Junit5
- **Java-versies upgraden:** Code migreren om nieuwere Java-versies te gebruiken (bijv. JDK 1.8 naar 11/17/21)
- **Code Smells oplossen:** Automatisch oplossingen toepassen voor problemen die zijn geïdentificeerd door statische code-analysetools
- **Log4j naar SL4J-migratie:** Logframeworks bijwerken voor verbeterde prestaties en beveiliging

De ontwikkeling van recipes begon met de Java-taal en is in de loop van de tijd uitgebreid naar verschillende technologieën, zoals Kotlin, Quarkus, Micronaut, Scala, SQL, Node.js, Maven, Feature Flags, Jenkins, Gradle, Groovy en nog veel meer.

Voordelen

- **Geautomatiseerde Refactoring:** Automation-recipes voor wat normaal handmatig, foutgevoelig werk zou zijn. Daardoor kunnen ontwikkelaars zich concentreren op het bouwen van nieuwe functies
- **Schaalbaar:** OpenRewrite kan worden toegepast op meerdere repositories en grootschalige projecten
- **Consistentie:** Zorgt voor uniforme toepassing van codestandaarden, opmaak en migratie over de hele codebase
- **Aanpasbare Recipes:** Ontwikkelaars kunnen hun eigen recipes maken en aanpassen aan specifieke behoeften het Moderne platform is beschikbaar voor de refactoring van grote codebases en verschillende soorten projecten.

Uitdagingen

- **Recipe stability:** Aangezien OpenRewrite in ontwikkeling is, zijn sommige recepten mogelijk niet altijd stabiel. Het up-to-date houden van de tool en recepten is essentieel voor een soepele refactoring.

Over de auteur



Mahendra Rao Bandaru | Senior Architect

Technische Oplossing & Applicatie Architect met 17,5 jaar uitgebreide ervaring in Architectuur, Oplossingen, Gespecialiseerd in Java, Spring Framework, Spring Boot, Web Services, Kotlin, Microservices, Angular, React en AWS Cloud. De afgelopen 5+ jaar bij Capgemini ben ik betrokken geweest bij meerdere opdrachten voor klanten, met focus op delivery en groei. In mijn vrije tijd werk ik als opensource-contributor aan spring-projecten zoals spring boot migrator, spring rewrite commons, OpenRewrite Projects en als Contributor Editor bij foojay.io (OpenJDK-community).



12 Foutloze controllersoftware binnen handbereik dankzij Synthesis-Based Engineering (SBE) en Poka Yoke

Cyber-fysieke systemen, zoals tunnels, industriële printers en wafer scanners, vereisen betrouwbare controllersoftware om veilig en correct te functioneren. Het ontwikkelen van dergelijke software wordt echter steeds complexer, temeer door toenemende prestatie-eisen en een groeiend aantal systeemfuncties. Ontwikkelaars moeten alle mogelijke situaties in kaart brengen en ervoor zorgen dat de controller daarin de juiste keuzes maakt. De huidige aanpak leidt echter tot late foutdetectie, wat resulteert in vertragingen en hoge kosten.

SBE is een moderne methode waarbij ontwikkelaars zich kunnen richten op wat het systeem moet doen (eisen), in plaats van hoe het dat moet doen (implementatie). Met een druk op de knop wordt vanuit de eisen middels synthese automatisch een controllermodel uitgerekend. Synthese is een wiskundig algoritme dat alle situaties doorrekent en garandeert dat deze allemaal correct afgehandeld worden. Dit controllermodel kan effectief worden gesimuleerd en geanalyseerd, wat zorgt voor snelle feedback cycli. Vanuit dit controllermodel wordt mede automatisch een correcte software-implementatie gegenereerd. SBE heeft als doel om met relatief beperkte inspanning en tegen lagere kosten, controllers van hogere kwaliteit te ontwikkelen.

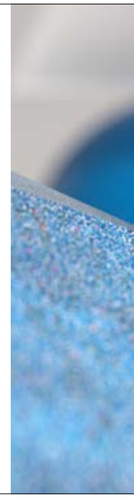
Hoewel SBE terrein wint in de industrie, brengt de adoptie ervan ook uitdagingen met zich mee, omdat het nieuw is en afwijkt van bestaande ontwikkelingstechnieken. Daarom werken we bij Capgemini in het Poka Yoke project samen met de Embedding Systems Innovation groep van TNO. Poka Yoke is een Japanse term voor het ontwerpen van producten om menselijke fouten te voorkomen. In dit project wordt gewerkt aan het verbinden van SBE met meer vertrouwde specificatietalen zoals UML en SysML. Met de combinatie van intuïtieve specificatie talen, simulatie en SBE ligt foutloze controllersoftware binnen handbereik.

Over de auteur



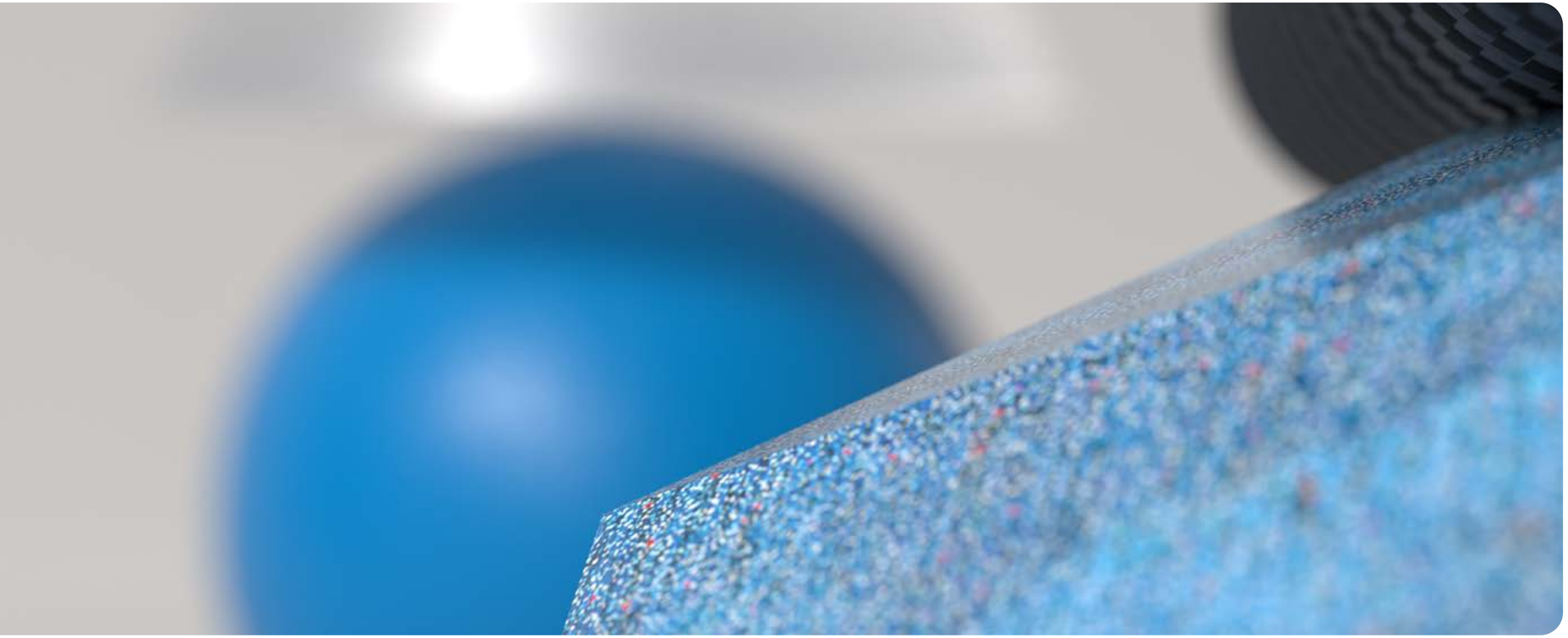
Yuri Blankenstein | Lead Digital Engineer

Yuri was een ervaren lead digital engineer bij Capgemini. Hij werkte als scientific programmer die academische onderzoeksresultaten vertaalt naar praktische automatiseringsoplossingen. Met een expertise in procesoptimalisatie hielp hij organisaties efficiënter te werken door complexe wetenschappelijke inzichten om te zetten in schaalbare technologische innovaties. Tijdens zijn tijd bij Capgemini was Yuri gedreven door de overtuiging dat technologie en wetenschap samen de sleutel zijn tot continue verbetering en slimmer werken.





Sustainability_



Highlights:

- De afhankelijkheid van batterijen voor de werking van IoT-apparaten brengt aanzienlijke nadelen met zich mee.
- De kern van batterijloze IoT-technologie ligt in energieoogsttechnieken
- Een belangrijke voorwaarde van batterijloze IoT-apparaten is dat ze extreem energiezuinig zijn.
- De grootste uitdaging ligt echter in de acceptatie van batterijloze IoT door de afnemers
- Integratie met AI en Machine Learning zal batterijloze IoT een steeds belangrijkere rol toebedelen.

13 Batterijloos *Internet of Things*

Spaar het milieu en maak onderhoud overbodig

Het Internet- of-Things (IoT) heeft de afgelopen jaren een enorme vlucht genomen. Met miljarden ver-bonden apparaten wereldwijd, verandert het de manier waarop we communiceren, werken, en leven. Echter, de afhankelijkheid van batterijen voor de werking van IoT-apparaten brengt aanzienlijke nadelen met zich mee. In dit artikel bespreken we de opkomst van batterijloze IoT , met specifieke aandacht voor twee toepassingen die laten zien hoe deze technologie de IoT-wereld kan transformeren.

Batterijen vormen een enorme milieubelasting. Ze hebben een beperkte levensduur, moeten regelmatig vervangen of opgeladen worden, en weggegooid. Bovendien kan de noodzaak van frequente batterij-vervanging of het opladen de kosten en complexiteit van IoT-implementaties verhogen, vooral in groot-schalige toepassingen.

De oplossing voor deze problemen komt in de vorm van batterijloze IoT-apparaten. Door energie te oog-sten uit de omgeving, zoals van zonne-energie, thermische energie of vibraties, kunnen deze apparaten volledig autonoom werken, dus zonder noodzaak van batterijen.



De Technologie Achter Batterijloze IoT

De kern van batterijloze IoT-technologie ligt in energieoogsttechnieken. Energieoogsten is het opvangen van kleine hoeveelheden energie uit de omgeving. Er zijn verschillende technologieën voor ener-gieoogsten, waaronder:

Zonne-energie

Zonnecellen kunnen licht omzetten in elektriciteit. Hoewel dit niet nieuw is, zijn er recentelijk aanzienlijke verbeteringen geweest in de efficiëntie van zonnecellen (vooral in die voor gebruik binnenshuis), wat ze aantrekkelijker maakt voor IoT-toepassingen. Het aantal variaties in uitvoering (glas, folie, transparantie) en materialen (silicium, organisch) is ook flink toegenomen.

Thermo-elektrische energie

Deze technologie maakt gebruik van temperatuurverschillen om elektri-citeit op te wekken via het Seebeck effect. Dit kan gebruikt worden in industriële omgevingen waar temperatuurverschillen aanwezig zijn, zoals verwarmingsinstallaties.

Kinetische energie

Bewegingsenergie kan worden omgezet in elektriciteit door middel van piëzo-elektrische materialen of magneetjes die binnen een spoel bewegen. Deze aanpak is geschikt voor IoT-apparaten die in machines of voertuigen worden geplaatst. Lichtschakelaars op basis van dit idee zijn al jaren in gebruik.

Radiofrequentie (RF) energie

Uit omgevingsradiosignalen, zoals die afkomstig zijn van Wi-Fi-routers of mobiele netwerken, kan energie worden geoogst. Ook kunnen IoT-apparaten met speciale zenders worden aangestraald om zo energie toegestuurd te krijgen.

Plant energie

Er kan ook energie onttrokken worden uit de chemische processen die in de grond plaatsvinden (plant-e uit Wageningen).

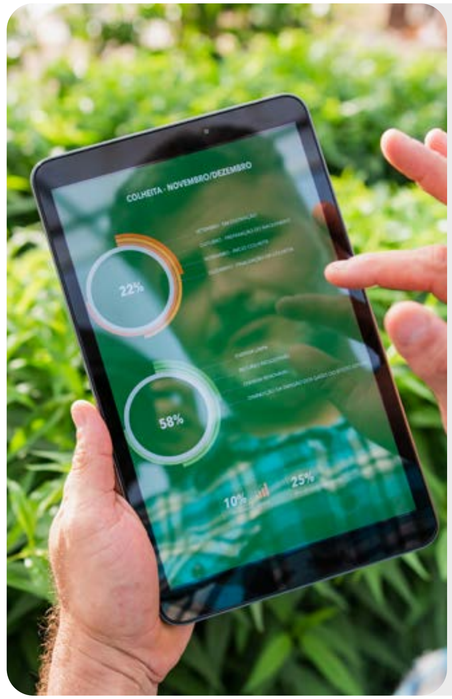
De geoogste energie kan worden opgeslagen in supercondensatoren of in hybride supercondensatoren (Small Lithium Batteries, Solid State Lithium batterijen en Lithium Keramische Batterjjen die, in tegen-stelling tot gewone Lithium-Ion batterijen, tienduizenden op- en ontlaadcycli mee kunnen en niet kunnen exploderen of ontbranden).

Zuinige IoT-apparaten

Een belangrijke voorwaarde van batterijloze IoT-apparaten is dat ze extreem energiezuinig zijn. Omdat de opgewekte energie vaak heel beperkt is, moeten deze apparaten ontworpen worden op maximale energie-efficiëntie. Dit betekent dat ze kort actief zijn en lange periodes van slaap of ultra-laag energie-verbruik kennen tussen actieve periodes. De gebruikte microcontrollers en sensoren zijn dan ook geop-timaliseerd voor ultra-laag energieverbruik, vaak door gebruik te maken van state-of-the-art technolo-gieën zoals sub-threshold IC-ontwerp, superzuinige radio's en energiezuinige draadloze communicatie-protocollen zoals Bluetooth Low Energy. De ontwikkelingen op het gebied van energiezuinigheid gaan steeds verder. Op het gebied van kunstmatige Intelligentie (AI) op IoT-apparaten (Edge-AI) worden de chips steeds krachtiger en zuiniger, zodat ook deze technologie inzetbaar is op batterijloze IoT-apparaten. Het gaat hier dan niet om het trainen van de modellen (daar gaat de meeste energie in zitten) maar om het draaien van de getrainde modellen op binnenkomende gegevens. Dat is namelijk veel ener-gie-zuiniger, zie o.a. de chips van efficient.computer en Innatera die met minder dan 1 mW genoegen nemen.

Toepassingen van Batterijloze IoT

Batterijloze IoT biedt tal van nieuwe mogelijkheden voor toepassingen die voorheen ondenkbaar waren. Hier zijn twee voorbeelden van hoe deze technologie in de praktijk kan worden toegepast.



Toepassing 1: Slimme Landbouw

Een van de veelbelovende toepassingen van batterijloze IoT is in de slimme landbouw. Landbouwers gebruiken steeds vaker IoT-apparaten om gegevens te verzamelen over bodemvocht, temperatuur, en an-dere cruciale factoren die invloed hebben op de oogstopbrengst. Traditionele IoT-apparaten vereisen echter regelmatige batterijvervanging, wat arbeidsintensief en dus kostbaar is.

Batterijloze IoT-apparaten kunnen dit probleem oplossen door gebruik te maken van zonne-energie of plant energie. Zo kunnen sensoren die bodemvocht meten worden uitgerust met kleine zonnecellen of plantenergie, wat voldoende stroom genereert om de sensoren te voeden en verzamelde gegevens te verzenden.

De voordelen hiervan zijn duidelijk. Landbouwers kunnen continu gegevens verzamelen zonder zich zorgen te hoeven maken over onderhoud. Dit kan leiden tot een nauwkeuriger beheer van water en an-dere hulpbronnen, wat niet alleen de opbrengst verhoogt, maar ook de duurzaamheid van de landbouw verbetert.



Toepassing 2: Slimme Steden en Infrastructuur

Slimme steden zijn een ander gebied waar batterijloze IoT-oplossingen aanzienlijke voordelen kunnen bieden. In stedelijke omgevingen worden IoT-apparaten gebruikt voor een breed scala aan toepassingen, van het monitoren van luchtkwaliteit en verkeersstromen tot het beheren van straatverlichting en afval-verwijdering.

Een specifiek voorbeeld is het gebruik van batterijloze sensoren voor het in de gaten houden van struc-turele integriteit van bruggen en gebouwen. Deze sensoren kunnen bijvoorbeeld vibratie-energie ge-bruiken die wordt gegenereerd door verkeer of wind om de integriteit van structuren in real time te controleren. Dit kan steden in staat stellen om problemen te identificeren voordat ze uitgroeien tot ge-vaarlijke situaties, wat uiteindelijk levens kan redden en kostbare reparaties kan voorkomen.

Een ander interessant voorbeeld is het gebruik van batterijloze sensoren voor het beheren van straatver-lichting. Deze sensoren kunnen zonne-energie gebruiken om de verlichting te regelen op basis van de omgevingslichtomstandigheden, wat helpt om energie te besparen en de efficiëntie van het stadsbe-heer te verhogen.



De Voordelen van Batterijloze IoT

De overgang naar batterijloze IoT-apparaten biedt verschillende belangrijke voordelen:

Onderhoudsvrijheid
Omdat er geen batterijen nodig zijn vervalt de noodzaak van regelmatige vervan-ging of opladen ervan. Omdat ook bijwerken van de software geen fysieke nabijheid vereist (via de radio-verbinding plaatsvindt), vervallen de operationele kosten hiervoor.

Duurzaamheid
Batterijen (denk aan tientallen miljarden afgedankte batterijen per jaar!) vormen een enorme bron van chemisch afval, dat giftig en explosiegevaarlijk is. Recyclen is niet voor 100% mogelijk. Het gebruik van batterijloze apparaten vermindert de milieu-impact van IoT-apparaten aanzienlijk. Als bewerkingen op de data op de edge-node plaats vinden, wordt er minder energie in het datacentrum gebruikt. Dat reduceert de milieubelasting door het datacentrum.

Bedrijfszekerheid
Batterijloze apparaten kunnen in veel gevallen meer bedrijfszeker zijn omdat ze niet afhankelijk zijn van een energiebron die kan uitvallen. Dit is vooral belangrijk in toepassingen waar conti-nue werking cruciaal is.

Schaalbaarheid
Omdat de noodzaak om batterijen te vervangen wegvalt, kunnen grootschalige IoT-implementaties eenvoudiger en goedkoper worden uitgevoerd.

Uitdagingen en Toekomstige Ontwikkelingen

Hoewel batterijloze IoT veelbelovend is, zijn er nog steeds uitdagingen die moeten worden overwon-nen. Een van de grootste uitdagingen is de beperkte hoeveelheid energie die kan worden geoogst uit de omgeving. Dit beperkt de complexiteit van de taken die deze apparaten kunnen uitvoeren en kan bete kenen dat ze minder frequent gegevens kunnen verzamelen of verzenden.

Om te kunnen omgaan met een tekort aan energie (bijvoorbeeld ‘s nachts in geval van zonnecel als ener-giebron) kunnen technieken als intermittent computing en energiebewuste scheduling veel problemen voorkomen, Een vergaderzaaldisplay kan bijvoorbeeld ‘s avonds en ‘s nachts het display uitzetten, en functies alleen uitvoeren als daarvoor de energie beschikbaar is.

Er is ook de kwestie van standaardisatie. Hoewel er verschillende protocollen en technologieën bestaan voor batterijloze IoT, is er nog geen breed geaccepteerde standaard die de interoperabiliteit tussen ver-schillende apparaten en systemen kan waarborgen.

Om te laten zien wat er mogelijk is, is ook een prototype batterijloos vergaderkamerdisplay gebouwd, waarmee veel ervaring is opgedaan.

De grootste uitdaging ligt echter in de acceptatie van batterijloze IoT door de afnemers. Deze kijken vooral

naar de initiële kostprijs en niet naar de Total Cost of Ownership over de leeftijd van het apparaat. De kosten van onderhoud en duurzaamheid spelen geen of maar een heel kleine rol bij de besluitvorming over de keuze tussen IoT met batterij en batterijloze IoT. Het verschil in fabricagekosten tussen beide zal, bij brede acceptatie van deze technologie, steeds kleiner worden, waardoor het argument van kostprijs steeds meer aan kracht zal verliezen.

AI en Machine Learning Integratie

De integratie van AI en Machine Learning speelt ook een belangrijke rol in de toekomst van batterijloze IoT. Door Machine Learning-algoritmen te gebruiken om patronen in de verzamelde gegevens te her-kennen en AI te gebruiken om het energiebeheer te optimaliseren kunnen apparaten nog efficiënter worden, wat leidt tot slimmere en meer autonome IoT-systemen. De toekomst van batterijloze IoT hangt af van verdere verbeteringen in energie-oogsttechnologieën en energiezuinige elektronica. Daarnaast zijn de ontwikkeling van nieuwe radiotechnologieën, communicatieprotocollen en netwerkarchitecturen die specifiek zijn ontworpen voor batterijloze apparaten cruciaal voor de bredere acceptatie en implementatie van deze technologie. Omdat de aandacht van de chipmakers hier steeds meer op gericht is wordt het ook steeds goedkoper. (Ambiq Micro, Onio, Silicon Labs, Nordic Semiconductor, OnSemi, STMi-crolectronics).

Conclusie

Batterijloze IoT is een spannende stap vooruit in de evolutie van het Internet-of-Things. Door gebruik te maken van energieoogsttechnieken kunnen IoT-apparaten werken zonder de nadelen en beperkingen van batterijen, wat hun bruikbaarheid en duurzaamheid aanzienlijk vergroot. Integratie met AI en Machine Learning zal batterijloze IoT een steeds belangrijkere rol toebedelen. Hoewel er nog uitdagingen zijn die moeten worden aangepakt, biedt batterijloze IoT een goede oplossing voor de schaalbare en duur-zame uitrol van IoT in verschillende sectoren. Capgemini gaat deze uitdagingen graag aan, en biedt ook hulp aan in de vorm van een tweedaagse Ultra Low Power hands-on workshop die een overzicht geeft van de belangrijkste aspecten van het ontwerp van Ultra Low Power systemen en praktische voorbeelden die laten zien hoe het energiegebruik van ingebedde systemen te minimaliseren.



TL;DR

Batterijloze IoT-apparaten halen energie uit de omgeving, waardoor ze zonder batterijen kunnen wer-ken. Dit maakt onderhoud overbodig, is duurzaam en verhoogt de bedrijfszekerheid. Slimme landbouw en slimme steden zijn enkele toepassingen. De toekomst van IoT zal aanzienlijk verbeteren door toepassing van deze technologie.

Over de auteur



Herman Roebbers | Embedded Systems and Ultra Low Power expert

Herman Roebbers is de Ultra Low Power expert van Capgemini Engineering Nederland. Hij is consultant voor ingebedde systemen en werkte aan first-of-a-kind systemen en NWO-onderzoeksprojecten.

Ook geeft hij gastcolleges en schrijft hij artikelen. Verder geeft hij de workshop “Ultra Low Power for the IoT”.





14 Brandstofcellen in mobiliteit: Kansen en uitdagingen

Hoewel brandstofcellen sinds het einde van de 19e eeuw bestaan, heeft de recente drang naar een energieovergang als oplossing voor klimaatuitdagingen hen opnieuw in de schijnwerpers gezet. De vraag die veel mensen zich stellen is: wat is het probleem dat we proberen op te lossen, en waarom doen we zoveel moeite om brandstofcellen te onderzoeken?

Brandstofcellen bieden een duurzame oplossing voor mobiliteit door waterstof en zuurstof om te zetten in elektriciteit, met als enige bijproduct water. Vergeleken met verbrandingsmotoren, die slechts een efficiëntie tot ongeveer 30% kunnen bereiken, kunnen brandstofcellen tot wel 75% efficiënt zijn dankzij hun directe omzetting van chemische energie in elektriciteit.

Vergeleken met batterij-elektrische voertuigen (EV's) bieden voertuigen met brandstofcellen enkele duidelijke voordelen in specifieke toepassingen. Terwijl EV's efficiënt zijn voor kortere afstanden en dagelijks gebruik, hebben brandstofcelvoertuigen een hoger bereik en kunnen ze snel worden bijgetankt, wat ze ideaal maakt voor langeafstandsritten en zwaar transport zoals vrachtwagens. Deze voordelen maken brandstofcelvoertuigen een uitstekende groene keuze voor toepassingen waar lange laadtijden of beperkte actieradius van batterijen een beperking vormen.

De toekomst van brandstofcellen richt zich op het ontwikkelen van kosteneffectievere oplossingen en onafhankelijkheid van zeldzame aardmaterialen, zoals platinum en nikkel.

Kortom, brandstofcellen vormen een haalbare middenweg tussen e-brandstoffen en batterijen en zijn vooral geschikt voor nichetoepassingen waar de energiebehoefte hoog is en de infrastructuur beperkt. Het is essentieel om zowel de kansen als de uitdagingen van brandstofcellen duidelijk te communiceren, zodat hun bijdrage aan duurzame mobiliteit en hun obstakels voor bredere implementatie helder worden voor een breed publiek.

Over deauteur



Nipun Kaushik
Electro Mechanical Engineer

Nipun is lid van de waterstofcompetentie binnen het centre of Excellence en Sustainability bij Capgemini Engineering. Zijn achtergrond ligt in Automotive Engineering en Energietransitie. Het artikel is het resultaat van Nipun's onderzoekswerk aan de Technische Universiteit Eindhoven, gepresenteerd op de Hydrogen Technology Expo Europe 2024 in Hamburg.

15

Overvloed aan duurzaamheids software in een popcornmarkt

Hoe kies je de juiste software ter ondersteuning voor je duurzaamheidsbeleid, emissie verlaging en compliance, afgestemd op de behoeften van alle stakeholders.

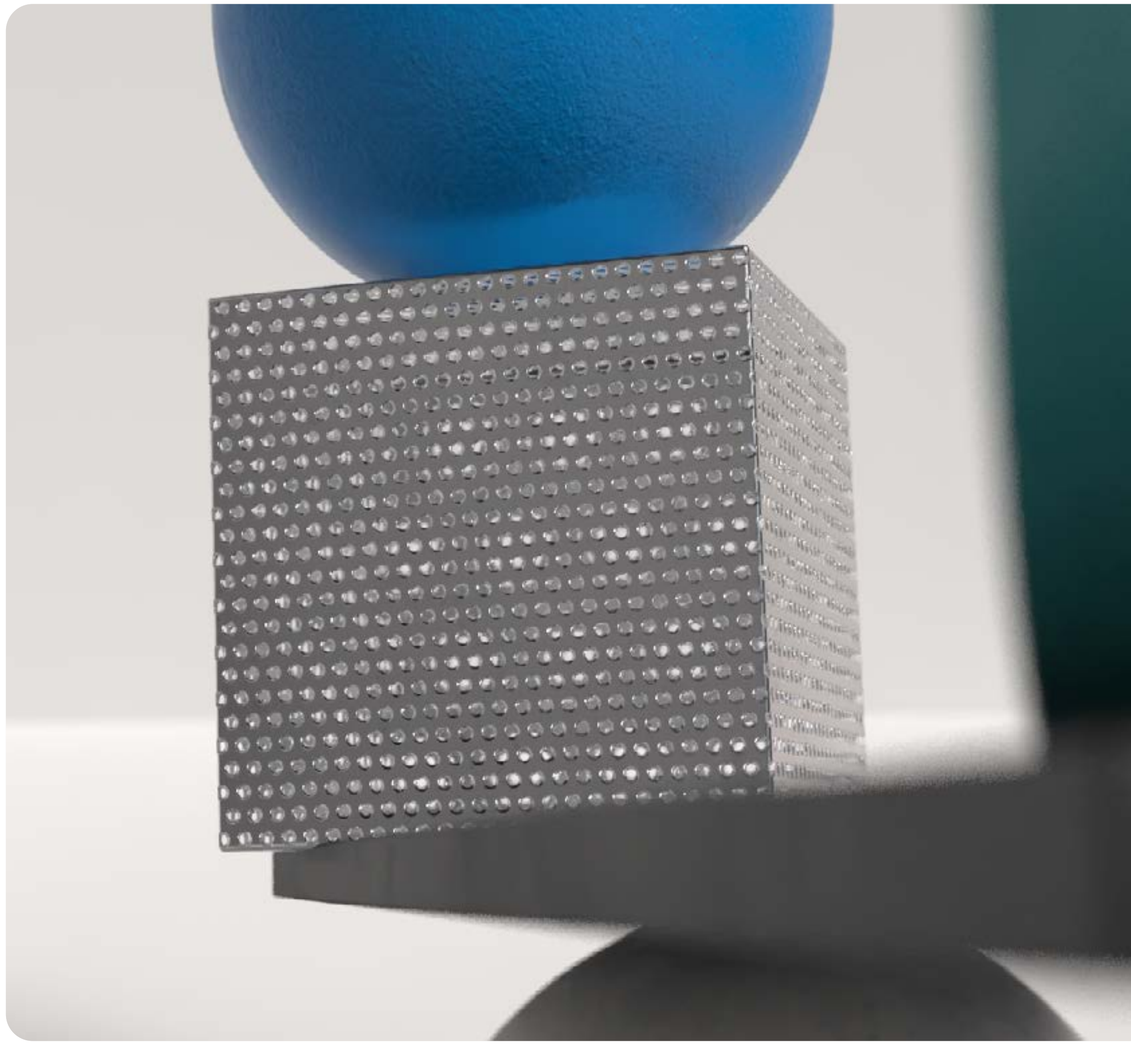
Het aanbod van nieuwe ESG*-software lijkt wel op popcorn. Er is zo veel ter beschikking gekomen als reactie op de toenemende druk op bedrijven om transparant te zijn over hun duurzaamheidsinspanningen. Met dit overweldigende aanbod aan oplossingen wordt de keuze dus steeds complexer.

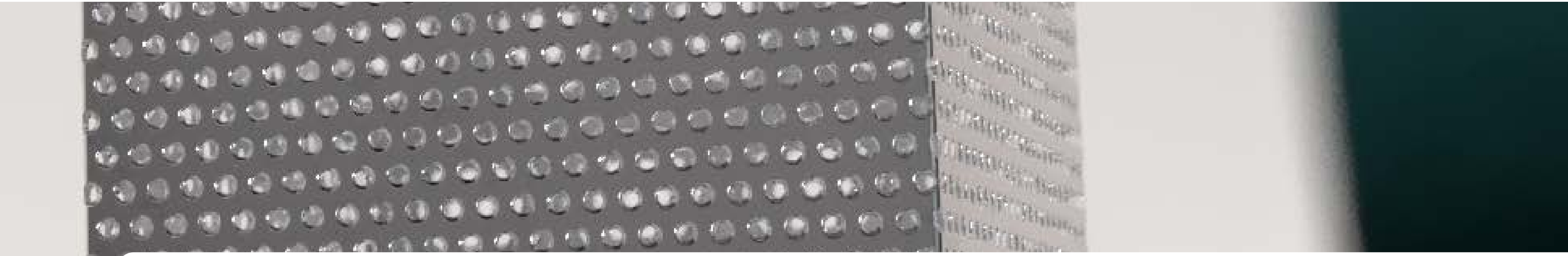
Hoe maak je keuze voor ESG software die, passend binnen de architectuur, output levert voor verplichtingenrapportages (CSRD), faciliteert in emissie reductie doelstellingen en eventueel het ecodesign van producten ondersteunt?

Het begint met het in overweging nemen van de behoeften van verschillende stakeholders. IT-teams kijken naar de bestaande IT-architectuur en de mogelijkheden voor datacollectie. Duurzaamheidsmanagers willen inzicht in de emissies over de hele waardeketen en gebruiksvriendelijke workflows. Voor finance of inkoop is geautomatiseerde, betrouwbare rapportage essentieel voor compliance en planning.

Kwaliteit van data is hierbij van groot belang. ESG-data kunnen gemakkelijk uit 40+ bronnen komen, waardoor effectief datamanagement en datagovernance, maar zeker ook een adequate applicatie-architectuur noodzakelijk zijn om betrouwbare informatie te waarborgen. Het ontbreken hiervan kan leiden tot een arbeidsintensief proces, compliance issues en vertragingen in verplichtingenrapportages.

Het aanbod van nieuwe ESG*-software lijkt wel op popcorn. Er is zo veel ter beschikking gekomen als reactie op de toenemende druk op bedrijven om transparant te zijn over hun duurzaamheidsinspanningen. Met dit overweldigende aanbod aan oplossingen wordt de keuze dus steeds complexer.





Hoe maak je keuze voor ESG software die, passend binnen de architectuur, output levert voor verplichtingenrapportages (CSRD), faciliteert in emissie reductie doelstellingen en eventueel het ecodesign van producten ondersteunt?

Het begint met het in overweging nemen van de behoeften van verschillende stakeholders. IT-teams kijken naar de bestaande IT-architectuur en de mogelijkheden voor datacollectie. Duurzaamheidsmanagers willen inzicht in de emissies over de hele waardeketen en gebruiksvriendelijke workflows. Voor finance of inkoop is geautomatiseerde, betrouwbare rapportage essentieel voor compliance en planning.

Kwaliteit van data is hierbij van groot belang. ESG-data kunnen gemakkelijk uit 40+ bronnen komen, waardoor effectief datamanagement en datagovernance, maar zeker ook een adequate applicatie-architectuur noodzakelijk zijn om betrouwbare informatie te waarborgen. Het ontbreken hiervan kan leiden tot een arbeidsintensief proces, compliance issues en vertragingen in verplichtingenrapportages.

Bij het maken van de juiste softwarekeuze zijn de volgende inzichten belangrijk:

- Inventarisatie van rapportage wensen en verplichtingen. Deze zijn vaak afhankelijk van de sustainability strategie, de marktbehoefte en sectorspecifieke eisen.
- Kenmerken van de huidige operatie of bedrijfsvoering; (markt) kenmerken, de supply chain, de geografische spreiding.
- De eisen voor geavanceerde rapportage en visualisering op bedrijf, proces en/of productniveau gerelateerd aan de (toekomstige) decarbonisatie inspanningen.
- De IT-architectuur vereisten, datamanagement en datagovernance.
- Implementatie en beheer budget.

Door deze inzichten zorgvuldig in kaart te brengen en te vertalen naar requirements, kun je een weloverwogen selectie maken van ESG-software die past bij jouw organisatie.

***ESG-gegevens verwijzen naar informatie over de prestaties van een bedrijf op drie belangrijke gebieden:**

Milieu, Sociaal en Governance. Deze gegevens worden gebruikt door investeerders, analisten, klanten en andere belanghebbenden om de duurzaamheid en sociale ethische impact van een organisatie te beoordelen

Over de auteur

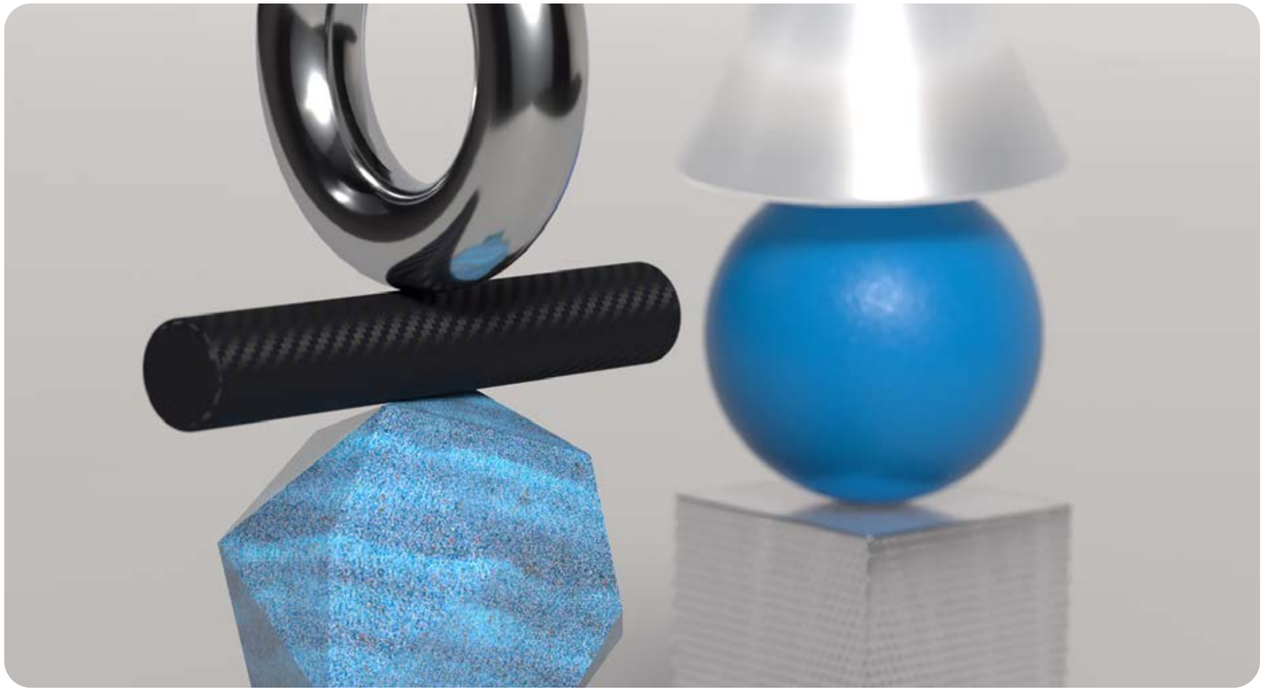


Josieke Moens | Director, Lead Sustainable Business

Josieke Moens is Lead Sustainable Business bij Capgemini, en sinds 2006 bezig met het verduurzamen van bedrijven. Ze heeft expertise ontwikkeld in het implementeren van ESG transities. Haar passie ligt in het helpen van organisaties om te groeien en tegelijkertijd een positieve bijdrage te leveren aan de maatschappij en het milieu. Josieke's kracht ligt in het ontwikkelen van impactvolle duurzaamheidsmodellen met oog op de organisatorische verandering en innovatieve strategie, zodat de noodzakelijke decabonisatie versneld wordt.

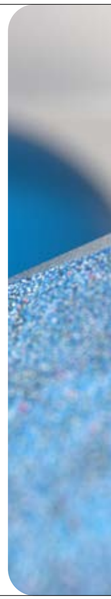
16 Repareer het niet-repareerbare, met additieve manufacturing

Hoe kan AM een waardevolle “(P)Layer” van duurzaamheid zijn?



Highlights:

- Additive manufacturing (AM), ook wel 3D-printen genoemd, is niet een “one size fits it all” technologie. Inzicht in de voordelen en beperkingen ervan is essentieel om het in te zetten als duurzame oplossing.
- Het doel van AM is niet om traditionele productiemethoden te vervangen, maar om deze aan te vullen en de productiemogelijkheden uit te breiden.
- Met AM kunnen specifieke beschadigingen worden gerepareerd die conventionele methoden niet aankunnen, zoals scheuren in metalen onderdelen, complexe geometrieën, en slijtage aan precisiecomponenten.
- AM-processen bewaken honderden procesparameters. Het begrijpen en modelleren van hun invloed leidt tot een betere levenscyclusanalyse (LCA), voorspelbaarheid en processtabiliteit.

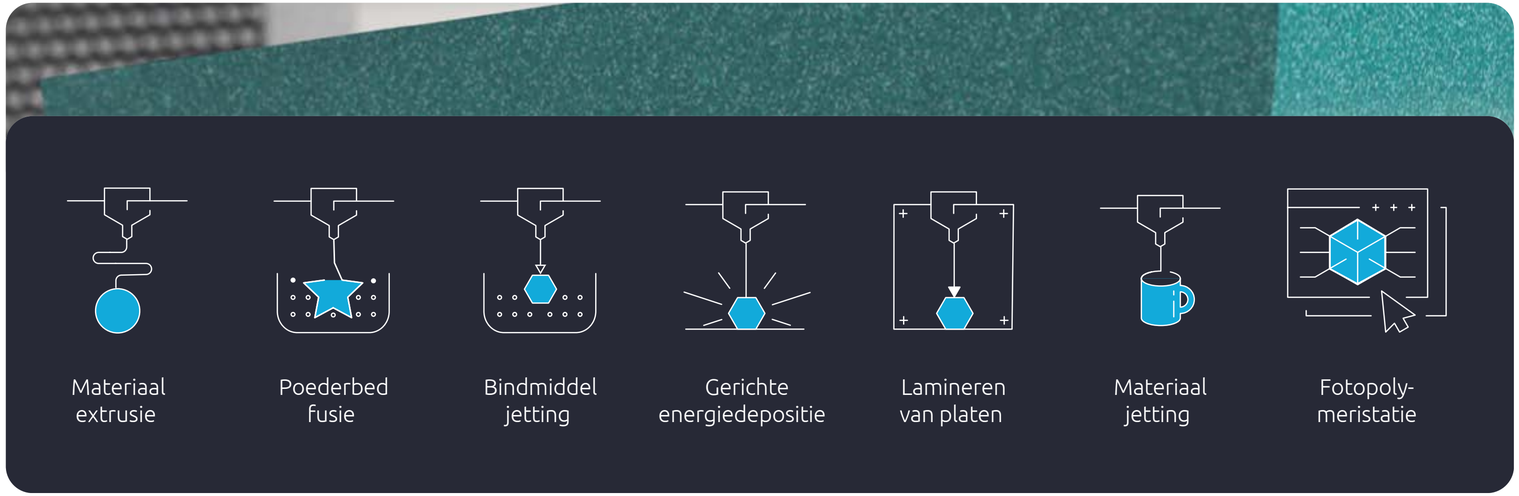


Additive manufacturing (AM), ook wel 3D-printen genoemd, is een printprocess waarbij materialen laag voor laag worden toegevoegd om van een digitaal ontwerp een driedimensionaal object te maken. De stap in de wereld van AM is een complexe beslissing, waarbij het wordt behandeld als een disruptieve technologie die ontwerp-, operationele en strategische veranderingen in een organisatie introduceert. AM kan beschouwd worden als een krachtig hulpmiddel om uitdagingen in ontwikkeling en productie op te lossen en is de manier om duurzame innovatie en evolutie te bereiken.

AM was een echte game-changer technologie in de jaren 90 en is sindsdien bezig een revolutie teweeg te brengen in de manier waarop materialen worden geselecteerd, complexe onderdelen produceert en zelfs assemblages ontwikkeld. Terwijl AM zich in het beginstadium richtte op toepassingen met plastic materialen en rapid prototyping, is het tegenwoordig beschikbaar voor een veelheid aan metalen materialen, biomaterialen, keramiek, beton. Printers worden in productielijnen geïntegreerd, bij fabrikanten uit de ruimtevaart, werktuigbouw, automobiel-, marine- en defensie-industrie.

Niet alleen het materiaalportfolio is groter geworden voor AM, maar ook de manier waarop verschillende materiaallagen op elkaar worden gebouwd. Complexe AM-onderdelen worden gebouwd in poederbedden en kunnen worden gevormd door technieken zoals directe energieafzetting, materiaalsintering, smelten, jetting of fotopolymerisatie, om er maar een paar te noemen. AM-apparatuur kan grondstoffen gebruiken in de vorm van poeder, draad of staven, die worden gesmolten door lasers, elektrisch of plasmaboog, elektronenbundels of kinetische energie. De combinatie van verschillende energiebronnen, bewegingstechnieken en materiaalvormen bepaalt de specifieke AM-technologie.

ISO/ASTM 52900, de eerste standaard om AM-technologieën te definiëren deelt AM in 7 belangrijke technologieën in, waaruit de afgelopen jaren meer dan 20 andere subtypes zijn ontwikkeld.



De 7 soorten AM - infographic van Hybrid Manufacturing Technologies (https://hybridmanutech.com/resources/)

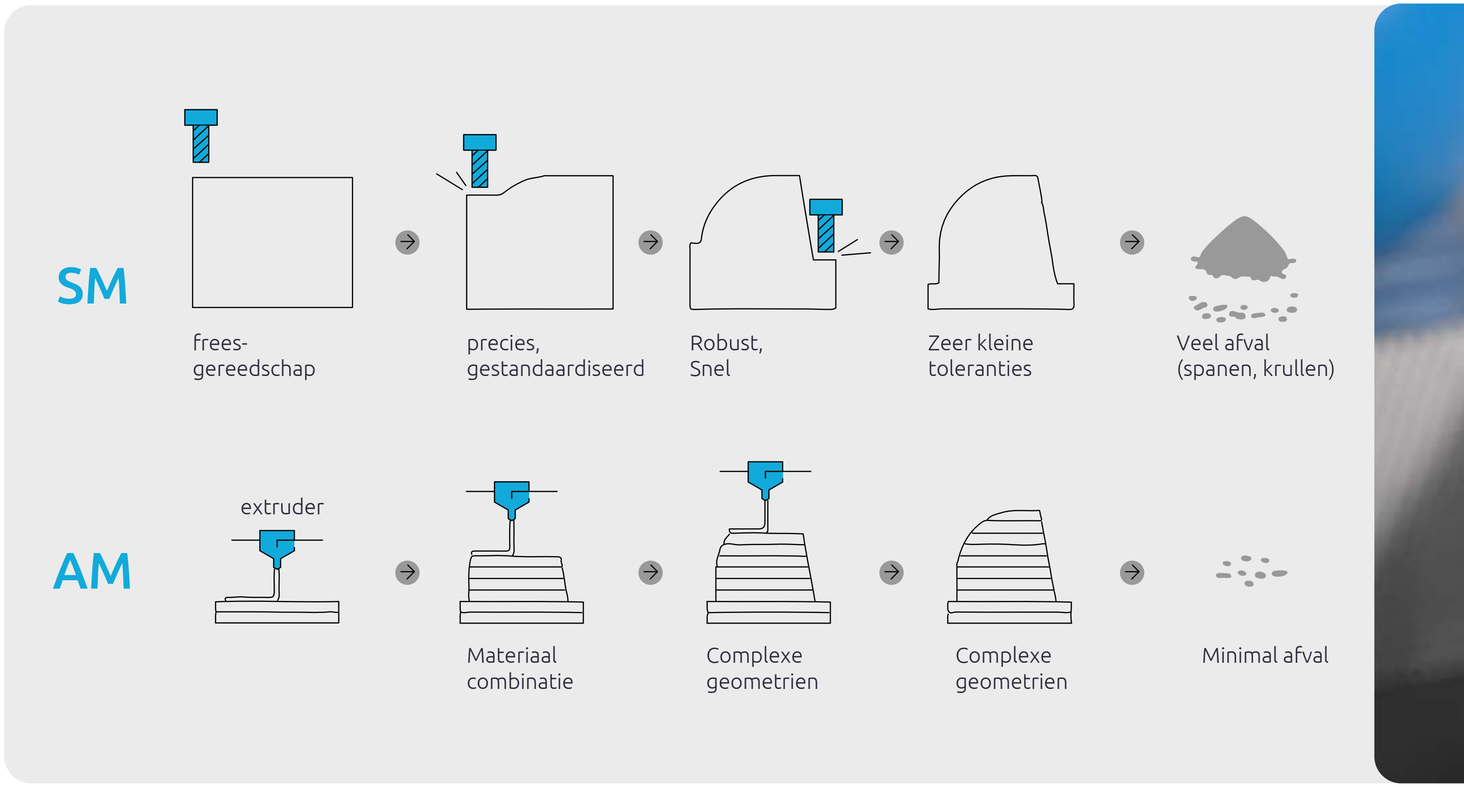
Inzicht in de voordelen en beperkingen van deze technologieën is de sleutel tot de beslissing om in de wereld van AM te stappen en het slim en efficiënt in te zetten om urgente uitdagingen op te lossen. AM is dus niet een “one size fits it all” technologie.

Duurzaamheid van AM

De noodzaak om te verschuiven naar circulariteit, afval te verminderen en meer duidelijkheid te krijgen in LCA, drijft innovatie. AM krijgt veel aandacht binnen dit onderwerp, omdat het complexe, lichtgewicht onderdelen kan produceren, die niet mogelijk zijn met conventionele productiemethoden. Het omzetten van een conventioneel design naar een AM-design en het printen van het onderdeel in plaats van het machinaal bewerken of gieten, leidt echter niet direct tot een “groene oplossing”. Het doel van AM is niet om traditionele productiemethoden te vervangen, maar om waarde toe te voegen als aanvulling daarop. Strikte vergelijkingen tussen productiemethoden zijn als appels met peren vergelijken, want duurzaamheid in AM komt op veel meer manieren tot uiting.

Beginnend bij grondstoffen voor AM, wordt er veel moeite gedaan om poeders en filamenten te verkrijgen uit recycling en hergebruik. Aangezien materiaalkwaliteit essentieel is voor een succesvol AM-proces, wordt de traceerbaarheid transparant gemaakt, waardoor de hele LCA van het AM-proces verbetert.

De materiaalefficiëntie van AM, gekwantificeerd door lage buy-to-fly ratio’s, betekent niet alleen dat AM-onderdelen een geoptimaliseerde vorm en topologie hebben, maar ook dat ze minder nabewerkingen nodig hebben (bv. machinale bewerking), waardoor minder materiaalafval ontstaat. Door een groot en complex onderdeel zoals een blisk of turbine te printen, wordt slechts ongeveer 5% afval gegenereerd tijdens de nabewerking, vergeleken met 60-80% bij traditionele bewerking. Verminderde bewerkingsinspanningen resulteren in energiezuinige bewerkingen.



Vergelijking: subtractief (conventioneel) vs. additief produceren

AM maakt de **combinatie van verschillende materialen** in één enkele print mogelijk. Deze eigenschap opent de deur naar geoptimaliseerd materiaalgebruik, doordat laagwaardige basismaterialen kunnen worden gecombineerd of gecoat met hoogwaardige legeringen om de prestaties van onderdelen te verbeteren of om hun levensduur te verlengen. Onderdelen kunnen worden geüpgraded door materiaal en waarde toe te voegen **precies waar dat nodig is.**

Duurzame AM operaties

Het strategisch inzetten van AM in operaties, zoals gedecentraliseerde productie van reserveonderdelen en reparatie, verbetert de beschikbaarheid van items met een lange levertijd aanzienlijk. Het vermindert risico's in toeleveringsketens en leidt tot circulariteit.

Gezien de voordelen en beperkingen van de verschillende AM-technologieën, is Directe Energie Depositie (DED) het meest geschikt voor reparatietoepassingen. DED werkt met een relatief eenvoudige kinematica: een energiebron (laser of elektronenbundel) wordt geactiveerd en een toevoersysteem (spuitmond/extruder) levert het poeder of de draad aan het doeloppervlak (bij reparatie, aan het beschadigde oppervlak). De geactiveerde energiebron creëert een smeltbad op het doeloppervlak en waarin de metaallaag wordt afgezet en een sterke metallurgische verbinding ontstaat. Flexibiliteit is een belangrijk voordeel van de DED spuitmonden: ze kunnen worden geïntegreerd in meerassige CNC machines of op robotarmen worden geplaatst om lokale, meerassige depositie uit te voeren.

De AM reparatie is relevant voor metalen onderdelen die worden beschouwd als niet repareerbaar, slijtage gevoelige onderdelen, onderdelen met oppervlaktekrassen, deuken of onderdelen die een upgrade nodig hebben.

Eerst wordt een haalbaarheidscontrole uitgevoerd, waarbij de materiaalcompatibiliteit, de hoofdoorzaak en de vorm van het defect, de schadefrequentie en de vraag naar onderdelen worden geanalyseerd. Voor algemeen gebruikte materialen zoals Titanium, roestvrij staal en Inconel zijn de procesparameters goed gedefinieerd. Echter voor moeilijk te lassen metalen zoals Aluminium en speciaal gereedschapsstaal, is de ontwikkeling van procesparameters nodig om reparatie op grotere schaal te implementeren.

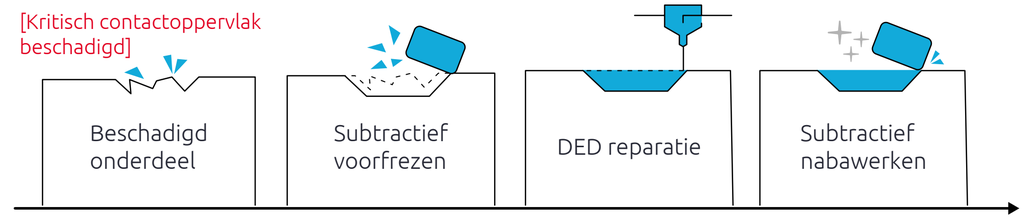
Zodra de haalbaarheid is geanalyseerd, wordt er een business case gebouwd om R&D, toepassing en management op één lijn te krijgen over de reparatiestrategie, de kostenverdeling en de resultaten. Als de business case positief wordt beoordeeld, kan de reparatiestroom van start gaan.

In onze technologiedemonstrator werd een vaak beschadigd onderdeel van aluminium onderzocht voor reparatie, in samenwerking met het Fraunhofer ILT in Aken, Duitsland. Aluminium is niet alleen uitdagend om te printen, maar de reparatie concurreert ook met de stabiele en productieve hogesnelheidsbewerking evenals met de vervanging van deze onderdelen.

In het eerste deel van het project werden de parameters verfijnd. Het vermogen van de laserstraal en de poedermassastroom werden aangepast om een poreus vrije depositie van aluminiumlagen te verkrijgen. De hardheid van de afgezette lagen werd ook geanalyseerd en was vergelijkbaar met de hardheid van het basismateriaal. Na de parameterkeuze werd het beschadigde onderdeel voorbereid voor reparatie: de beschadigde gebieden werden gereinigd en voorgefreesd voor betere toegang tot het gereedschap.

De AM reparatie is relevant voor metalen onderdelen die worden beschouwd als niet repareerbaar, slijtage gevoelige onderdelen, onderdelen met oppervlaktekrassen, deuken of onderdelen die een upgrade nodig hebben.





De reparatieworkflow

Deze stap werd gevolgd door het aanbrengen van aluminiumlagen om de geometrie opnieuw op te bouwen en een laatste freesbewerking om de vereiste maattoleranties te verkrijgen.

Dit demonstratiemodel benadrukt het potentieel van DED in het overwinnen van materiaal- en kostenuitdagingen, maar heeft de reparatiekwalificatie niet in scope. Om deze methode op grotere schaal vrij te geven, is het nodig om testen uit te voeren om de mechanische eigenschappen van de geprinte gebieden te analyseren, zoals treksterkte, vermoeidheid. Niet-destructieve testen, zoals CT-scans, zijn nodig om de microstructuur en de kwaliteit van de afzetting te evalueren.

Toekomstperspectieven

AM-reparaties naar een hoger niveau tillen kan door automatisering en het gebruik van de kracht van AI. De ontwikkeling van evaluatietools kan de besluitvorming, parameterkeuze en algehele reparatietijd versnellen.

Modellen die getraind zijn om in-situ depositiefouten te detecteren, kunnen gebruikt worden om parameters tijdens de depositie te corrigeren en aan te passen, waardoor de kwaliteit van het proces aanzienlijk toeneemt.[1]

Standaardisatie van AM is een **gezamenlijke inspanning,** universiteiten en onderzoeksinstituten werken nauw samen met de industrie. Capgemini ziet AM als een **strategische troef** en heeft zich aangesloten bij lokale en internationale **AM-consortia** om implementatiestrategieën voor de industrie en standaarden voor AM-reparaties te definiëren.

1. Pandiyan, V., Cui, D., Richter, R.A. et al. Real-time monitoring and quality assurance for laser-based directed energy deposition: integrating co-axial imaging and self-supervised deep learning framework. J Intell Manuf (2023). https://doi.org/10.1007/s10845-023-02279-x

   Deshpande, S., Venugopal, V., Kumar, M. et al. Deep learning-based image segmentation for defect detection in additive manufacturing: an overview. Int J Adv Manuf Technol 134, 2081–2105 (2024). https://doi.org/10.1007/s00170-024-14191-6

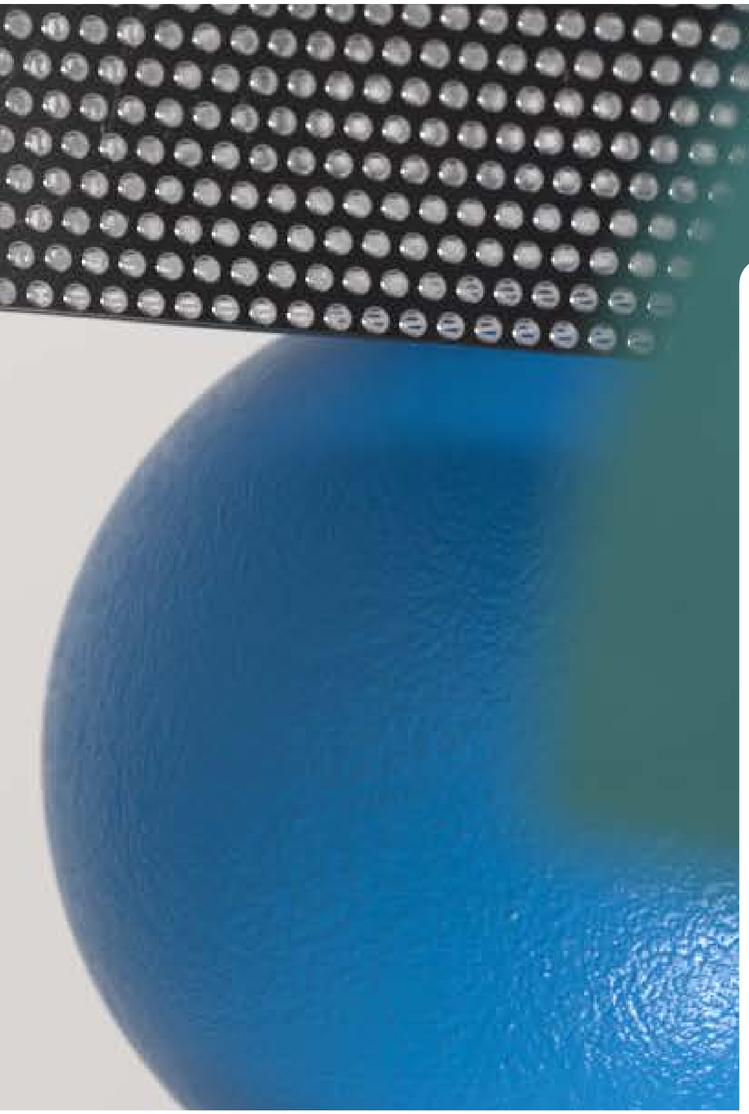
Over de auteur



Adrienn Matis | Technical Lead Additive Manufacturing, GBL

Adrienn is Technical Lead voor AM in de Global Business Line. Ze is verantwoordelijk voor de competentie ontwikkeling en uitvoering van een technische strategie voor AM, gecoördineerd door de CTO voor Manufacturing & Operations. Adrienn heeft een mechanical engineering achtergrond en werkt momenteel ook aan haar PhD op het gebied van AM.





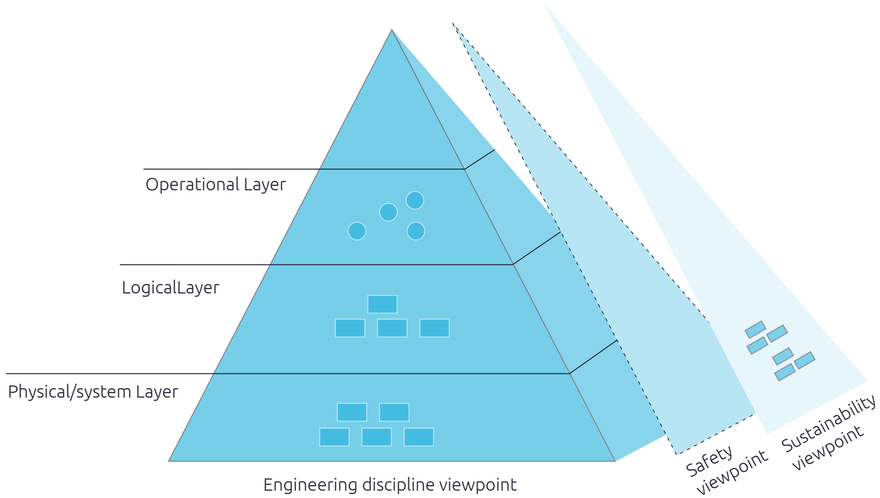
17 Design for *Sustainability*

Meer dan 70% van de ecologische footprint van een bedrijf is afkomstig van zijn producten. Dit is inclusief de toeleveringsketen, de productie en de gebruiksfase. Het is van groot belang voor de wereld om dit te reduceren. Duurzaamheid wordt een nieuwe dimensie, die in evenwicht moet worden gebracht met traditionele KPI's: prestaties, kosten en tijd.

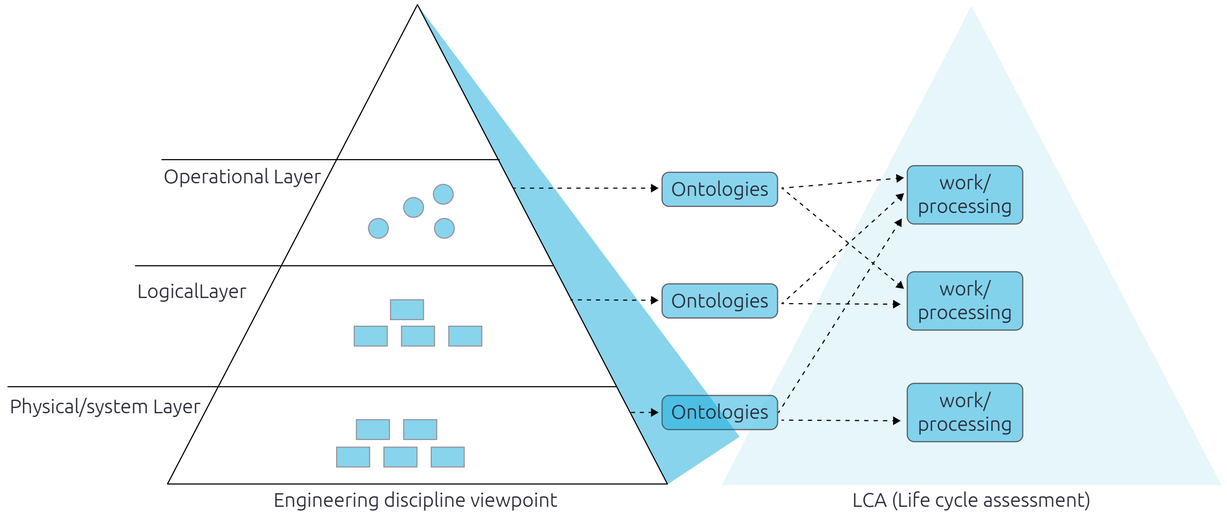
Als duurzaamheid wordt meegenomen in het ontwerpproces, dan brengt dit ook nieuwe uitdagingen met zich mee, te weten:

- Wat zijn precies de duurzaamheid doelstellingen?
- Hoe meet je de duurzaamheid doelstellingen?
- Hoe bouw je betere oplossingen met duurzaamheid als belangrijke drijfveer?
- Hoe ga je om met de steeds grotere hoeveelheid aan data en tooling?

Design For Sustainability omvat specifieke methoden, tools en vaardigheden om duurzaamheidsvraagstukken vanaf de eerste ontwerpfasen mee te kunnen nemen.



Engineering viewpoints



Sustainabilty backbone

Systems Engineering speelt hierbij belangrijke rol. Als bijvoorbeeld duurzaamheidsaspecten meegenomen worden bij de ontwikkeling van een drone, dan moet het complete systeem (holistisch) bekeken worden. Als besloten wordt om de aandrijving te vervangen door een brandstofcel, dan heeft dit impact op de range, gewicht en veiligheid van het complete systeem. Als er andere ‘groene’ materialen gebruikt worden, dan heeft dit impact op het gewicht en bereik.

Ontwerp beslissingen moeten gevalideerd kunnen worden met LCA (Life Cycle Analyses) en circulariteit assessments. Daarvoor is een digitale backbone ontwikkeld welke een uitbreiding is op PLM systemen. Met deze backbone zijn we in staat om ontwerpbeslissingen te meten en valideren vanaf de conceptuele ontwerpfasen. Deze backbone legt de verbinding tussen ontwerp en assessments door het delen van de gemeenschappelijke kenmerken van modellen (via ontologieën).

Kortom: D4S is een gestructureerde, holistische aanpak, gebaseerd op systems engineering om projecten zich te laten aanpassen op sustainability doelen. Ontwerp beslissingen moeten gevalideerd kunnen worden met LCA (Life Cycle Analyses) en circulariteit assessments. Daarvoor is een digitale backbone ontwikkeld welke een uitbreiding is op PLM systemen. Met deze backbone zijn we in staat om ontwerpbeslissingen te meten en valideren vanaf de conceptuele ontwerpfasen. Deze backbone legt de verbinding tussen ontwerp en assessments door het delen van de gemeenschappelijke kenmerken van modellen (via ontologieën).

Over de auteur



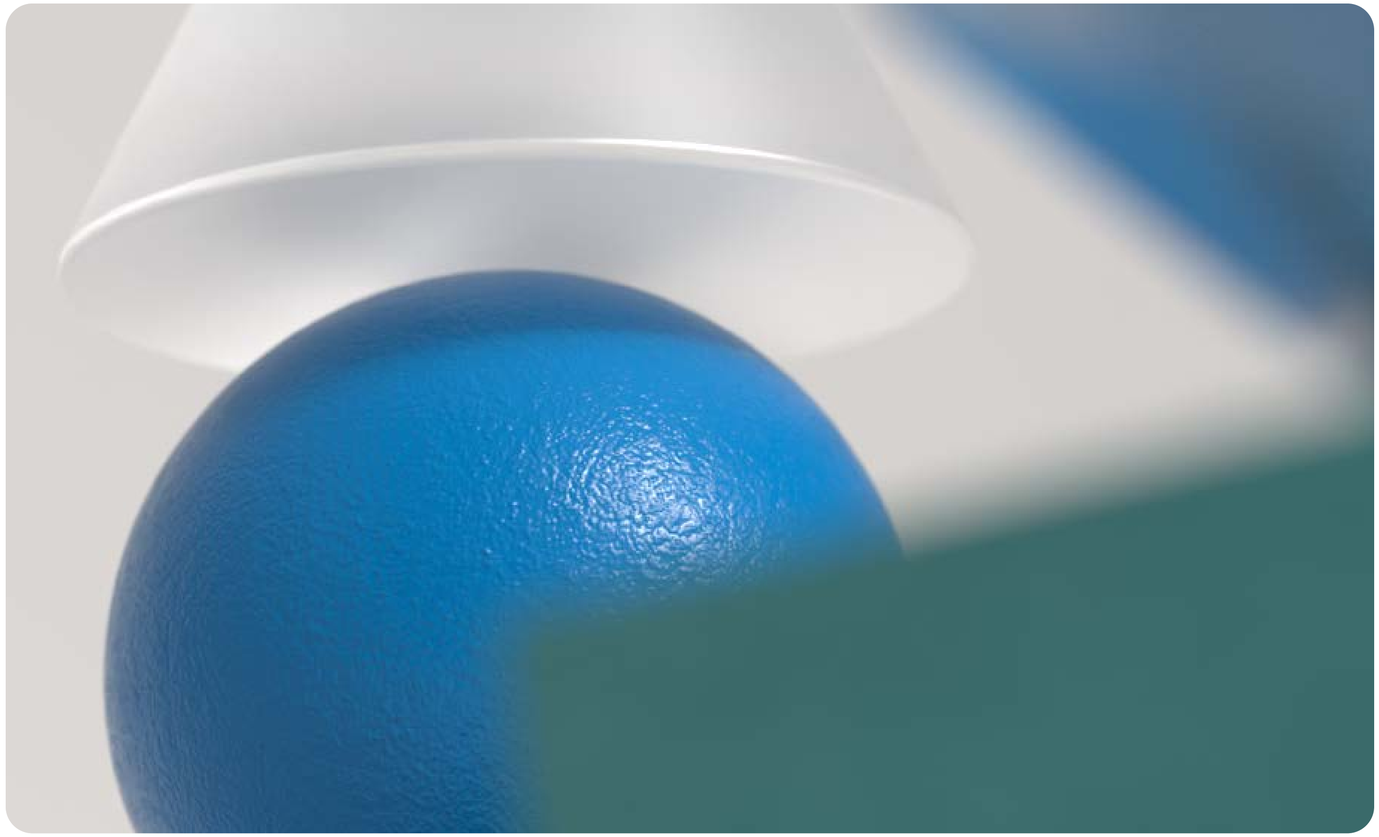
Lambert Verhagen | Technology Leader Complex Systems Engineering

Lambert werkt sinds de jaren 90 aan complexe systemen in uiteenlopende domeinen zoals lithografie, maritiem en tunnels. Multidisciplinair en systeem architectuur zijn terugkerende aspecten in zijn werk. Verbinden van zowel techniek als mensen is wat hem drijft. Van jongs af is hij begaan met duurzaamheid. Met D4S komt dit allemaal samen.





Business *Agility*_



Highlights:

- Verandering: de nieuwe constante.
- Generative AI: een nieuwe, potentieel disruptieve kracht voor de organisatie en de architect.
- Automatiseren, assisteren en augmenteren: hoe Generative AI het werk van de architect ondersteunt.
- Succesvol toepassen van Generative AI: alleen met een robuuste en schaalbare organisatie, processen en infrastructuur.
- Primus inter pares en hofnar: de rol van de Augmented architect in een omgeving waarin iedereen Generative AI gebruikt.



18 De “Augmented” Architect

Primus inter pares en hofnar

We leven in een complexe, roerige wereld waarin steeds snellere verandering norm is geworden. De grote uitdaging voor organisaties hierin is om de complexiteit van hun digitale ecosystemen zowel beheersbaar als tegelijkertijd flexibel te houden.

De architect speelt daarin een sleutelrol. Maar is hij of zij daarvoor voldoende toegerust?

Verandering: de nieuwe constante

De wereld om ons heel verandert steeds sneller en deze veranderingen hebben een steeds grotere impact op onze samenleving. Verandering is de nieuwe constante en voorbereid zijn op disruptieve krachten is vitaal om te overleven en te bloeien.

Om hier het hoofd aan te kunnen bieden moet een organisatie in staat zijn mee te bewegen met veranderingen, maar ook zichzelf opnieuw uit te vinden. Daarvoor moet een organisatie goed geïnformeerd zijn, snel en doeltreffend beslissingen kunnen nemen en met de juiste kennis en kunde vaart kunnen maken in de gekozen richting.

Architectuur wordt steeds belangrijker

Veranderen kan niet zonder een doeltreffende architectuur; de blauwdruk van zowel de organisatie, de voortbrengingsprocessen en de producten en diensten, moeten voortdurend effectief toepasbaar zijn.

De architect speelt hierbij een sleutelrol, die in de afgelopen jaren al sterk is veranderd.

Speelde de architect vroeger voornamelijk een (vooraf) voorschrijvende en (achteraf) controlerende rol in programma's en projecten, nu vaart de architect, als "servant leader" vaak mee met de Agile teams in het voortbrengingsproces. Beslissingen worden grotendeels binnen de teams belegd en architectuur wordt daarbij net op tijd, en net genoeg (Just-in-time, Just-enough) opgeleverd.

Dit vereist van de architect een enorm absorptievermogen van signalen, informatie en vaardigheden.

De kracht van Generative AI

Aan dat speelveld is een nieuwe, potentieel disruptieve kracht, aan toegevoegd: Generative AI. Deze nieuwe technologie maakt het mogelijk om grote hoeveelheden kennis binnen te halen en toe te passen. Dat geeft, zeker voor kleinere organisaties, de mogelijkheid om grotere concurrenten versneld in te halen.

De architect die Generative AI niet gebruikt, zal op termijn worden vervangen door de architect die deze technologie wel gebruikt.



Automatiseren en assisteren

Generative AI-toepassingen zijn uitstekend in staat om grote hoeveelheden documentatie te verwerken, aan te vullen, samen te vatten en om te zetten naar verschillende formaten en stijlen. Denk aan het samenvatten van gespreksverslagen, gepersonaliseerd rapporteren aan verschillende stakeholders, genereren van overzichten en illustraties.

Dit is zeer waardevol, omdat de architect de essentie van grote hoeveelheden informatie snel moet kunnen vatten en hierover doeltreffend met een palet aan stakeholders moet kunnen discussiëren.

In de praktijk bieden architectuur tools (denk aan BizzDesign, Sparx, Orbus, Ardoq) nu al de mogelijkheid om de architect te assisteren met het toepassen van architectuurtalen, maken van modellen, documenteren en doen van eenvoudige analyses.

Ook in de adviserende rol van de architect kan AI ondersteunen. Denk hierbij aan het helpen opstellen van impact analyses, risicoanalyses, opstellen van basisadviezen, uitvoeren van reviews en zelfs het genereren van scenario's, ontwerpen en programmacode.

Dit levert de architect veel tijdswinst op, maar om een "augmented architect" te worden is meer nodig...

Pas wanneer de architect door de inzet van AI meer kan doen dan hij/zij nu kan, spreken we van augmenting. Denk bijvoorbeeld aan het verkrijgen van nieuwe inzichten in complexe situaties die helpen bij het nemen van beslissingen, of het (semi) automatisch nemen van beslissingen.



Stel je eens voor...

Een persoonlijke AI assistent die de architect helpt de pijnpunten van een complex systeem in kaart te brengen, zelf verbeteringen voorstelt en hiervoor alternatieve scenario's beschrijft. In combinatie met andere technologieën biedt dit veel mogelijkheden.

Zo kan een architect samen met een CEO als stakeholder in een virtual reality omgeving verschillende scenario's doorlopen. Doordat actoren, business processen en IT-oplossingen worden gevisualiseerd komen user stories en de implementaties ervan tot leven. Ter plekke kunnen details van de voorgestelde architectuur worden opgevraagd, kunnen uitdagingen inzichtelijk worden gemaakt, en kunnen scenario's worden opgesteld en aangepast zodat de CEO optimaal inzicht krijgt in de reis naar de toekomst van zijn organisatie. Met de architect als reisleider.

Toekomstscenario's worden hierbij gegenereerd op basis van realtime analyse van architecturen/ontwerpen, implementaties en relevante invloeden. Modellen en analyseresultaten worden omgezet in programmatuur waarmee een VR-platform een virtuele omgeving creëert waarin sprekende voorbeelden worden gevisualiseerd. Een doeltreffende samenwerking tussen AI en andere technologieën.

En als we nog verder in de toekomst kijken, kunnen AI-platformen worden geïntegreerd met sensoren waarmee realtime data kan worden verkregen om scenario's nog meer waarde te laten krijgen.

Context is cruciaal

Voor een kwalitatief goede ondersteuning door AI is kennis van context (het probleem- en oplossingsdomein) cruciaal. In de basis hebben AI-systemen deze context niet, maar architecten hebben dit wel. Zij vangen immers uit een breed scala van bronnen informatie op die relevant kan zijn bij hun werk en ze nemen kennis en situationele ervaring mee vanuit het verleden.

Een einde aan de Babylonische spraakverwarring?

Generative AI kan kennis uit tal van verschillende informatiebronnen bijeenbrengen, hierop bevraagd worden, en informatie ook weer in verschillende formaten publiceren. Dat biedt de architect de mogelijkheid om architecturen, ontwerpen, scenario's en zelfs code bijeen te brengen en te gebruiken als "golden source" voor beschrijvingen en analyses. Daarmee worden de conceptuele "kloven" tussen de verschillende abstractieniveaus gedicht.

En ook hiervoor geldt weer dat dit bereikt kan worden door het slim combineren van technologieën. Denk bijvoorbeeld aan het gebruik van Graph databases voor het bijhouden van architectuurmodellen. Dergelijke databases zijn, meer nog dan AI-taalmodellen, geschikt om complexe netwerken van concepten en relaties (wat architectuurmodellen in essentie zijn) op te slaan. Bovendien kunnen ze veel doeltreffender worden bevraagd en geven, in tegenstelling tot taalmodellen, precieze antwoorden. In combinatie met Generative AI kunnen deze architectuurmodellen in natuurlijke taal worden bevraagd, en worden betrouwbare antwoorden gegeven.

Een virtueel team voor complexe problemen

Voor complexe en omvangrijke problemen waarin belangrijke architectuur besluiten genomen moeten worden is vaak specialistische kennis, een speciale oplossingsstrategie, en menselijke interventie nodig.

Dit kan alleen bereikt worden door het slim combineren van technologieen. Hier komen AI Agents in beeld: processen die in staat zijn om autonoom, of in samenwerking met andere Agents, taken uit te voeren door (beperkt) beslissingen te nemen op basis van context, invoer en vooraf gedefinieerde doelen. Denk bijvoorbeeld aan een Agent die, naar aanleiding van wijzigingen die aan een architectuurmodel worden uitgevoerd, deze wijzigingen controleert en verbeteringen voorstelt en voorlegt aan een architectenteam. Of aan een Agent die, gebaseerd op aandachtsgebieden en de communicatievoorkeur van een stakeholder, automatisch een voor hem/haar een actuele en toegankelijke beschrijving van de architectuur genereert.

Met AI Agents heeft de architect als het ware een team van virtuele medewerkers beschikbaar, dat voortdurend nieuwe inzichten verzamelt, kwaliteit toetst, scenario's in kaart brengt en dit doeltreffend met de architect communiceert.

Ecosystemen van Agents kunnen een grote omvang bereiken, en daarbij zelfs het gedrag van hele systemen simuleren (Agent-Based Modeling). Daarmee wordt het mogelijk om scenario's uit te werken ter ondersteuning van besluitvorming.

AI evolueert snel

Het AI-landschap evolueert snel. Van grote modellen (Large Language Models) naar ecosystemen van kleinere, doeltreffender en doelmatiger modellen (Small Language Models), gecombineerd met AI-agents.

De Augmented architect: Primus inter pares en hofnar

De architect is echter niet de enige die gebruik zal maken van deze nieuwe technologien. Iedereen kan "augmented" worden.

Ook softwareontwikkelaars, ontwerpers, managers en business stakeholders zullen hun architectuur kennis en skills kunnen uitbreiden.

Wat heeft deze "brede augmention" voor gevolgen voor de rol van de architect? Zal de rol van de architect vervallen tot een self-service functie binnen de organisatie?

De veranderende rol van de architect

Een aantal vragen hoeven "augmented" collega's niet meer bij de architect neer te leggen, maar om doeltreffend gebruik te maken van AI zullen toch betrouwbare en actuele kennis en inzichten beschikbaar moeten worden gesteld. Daarnaast zijn bij het gebruik van Generative AI het doeltreffend formuleren van vraagstellingen ("prompts" genoemd), en het juist interpreteren van antwoorden cruciaal. Het in goede banen leiden van de conversaties tussen AI en de gebruikers, is met name iets waarbij de architect een sleutelrol zal blijven spelen.

Om doeltreffend te blijven functioneren zal de architect zich daarom moeten blijven verdiepen in de relevante business domeinen, in AI en de bijbehorende voortbrengingsprocessen van zowel applicaties als kennis(modellen).

Gevraagd: Hofnar met kennis en kunde

En hoe indrukwekkend Generative AI ook is, het is geen wondermiddel. De architect blijft verantwoordelijk voor de kwaliteit van de architectuur. De kritische houding, zoals die van een hofnar, blijft noodzakelijk. De architect zal de mogelijkheden en beperkingen van AI moeten kennen, en de organisatie met de beide benen op de grond moeten kunnen adviseren.

Dat betekent ook het openlijk in twijfel durven trekken van plannen, het tegenspreken van sterke meningen en een eigen visie kunnen verkondigen. Zoals een hofnar dat vroeger deed aan het hof.

Over de auteur



Hans van Rijs | Managing Consultant, Architect

Hans is een Solution & Enterprise Architect, met een achtergrond in software engineering & project management. Daarnaast deelt hij graag zijn kennis als trainer bij de Capgemini Academy. Is hij lid van The Open Groups O-AA standard Working Group, onderdeel van Capgemini's Architecture Community of Practice en. als coach van een team van getalenteerde collega's



19 AI unleashed: Op weg naar ultieme innovatiekracht

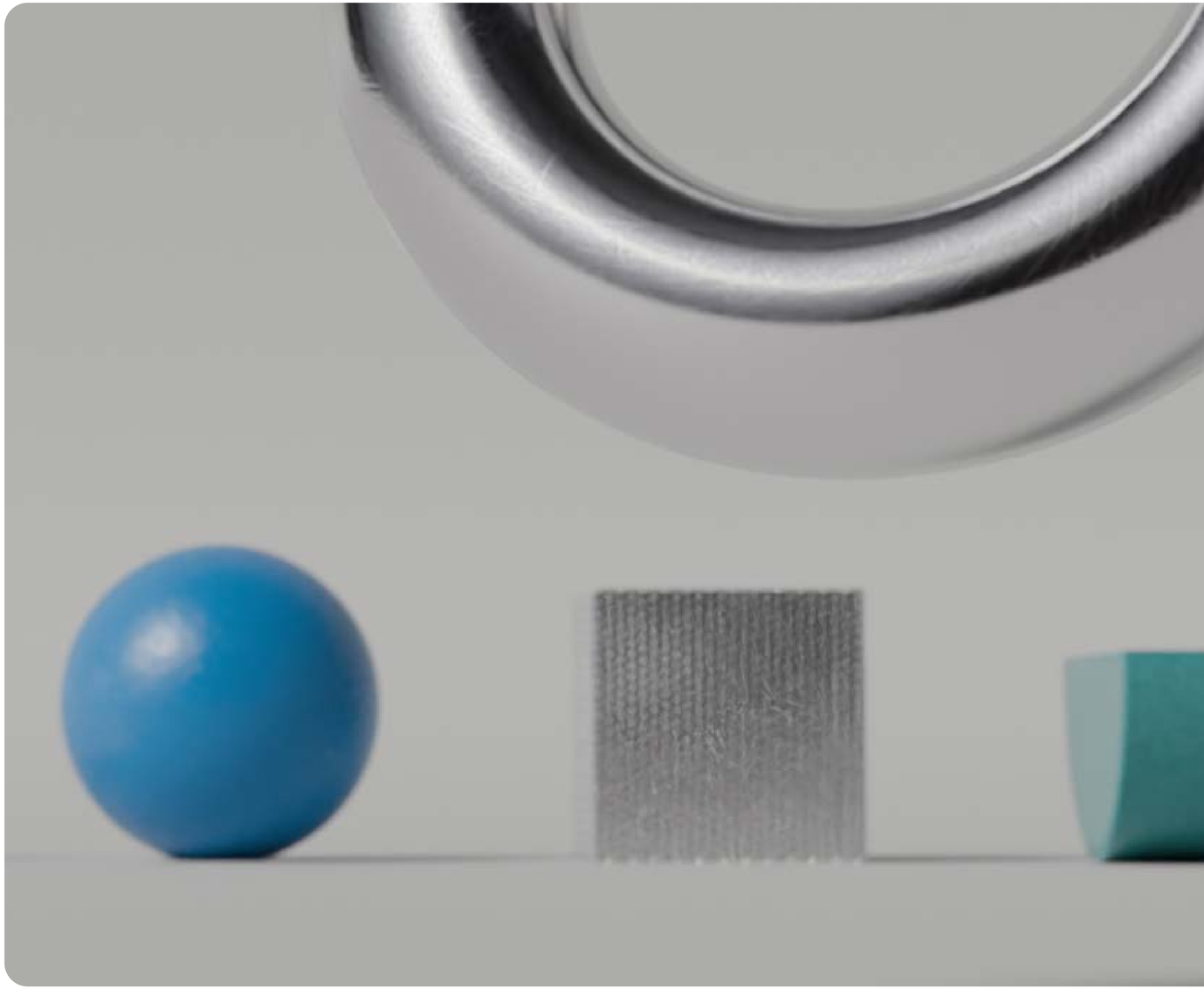
AI als motor voor extreme innovatie

Tesla doet het: het doorvoeren van 27 hard- en softwarewijzigingen per dag in productie. SpaceXdoet het: 10.000 hardware wijzigingen aan de grootste raket in 69 dagen. Klarna doet het: 2,3 miljoen klantenservice gesprekken in 35 talen per maand.



Highlights:

- Autonomie wordt de drijvende kracht achter innovatie.
- Met een Digital Self Management system krijgt iedereen up-to-date, actionable informatie naar een app gepusht.
- Combineer AI met CRM en ERP systemen.
- Werk met micro-doelstellingen.
- Meer data maken een AI-systeem niet altijd slimmer.



In de afgelopen 20 jaar zijn veel bedrijven overgestapt van traditioneel waterval projectmanagement naar moderne methoden die hun wendbaarheid en innovatiekracht aanzienlijk vergroot hebben. Veel organisaties hebben agile frameworks zoals Scrum, SAFe, LeSS en "Spotify" omarmd en zijn er redelijk volwassen in geworden.

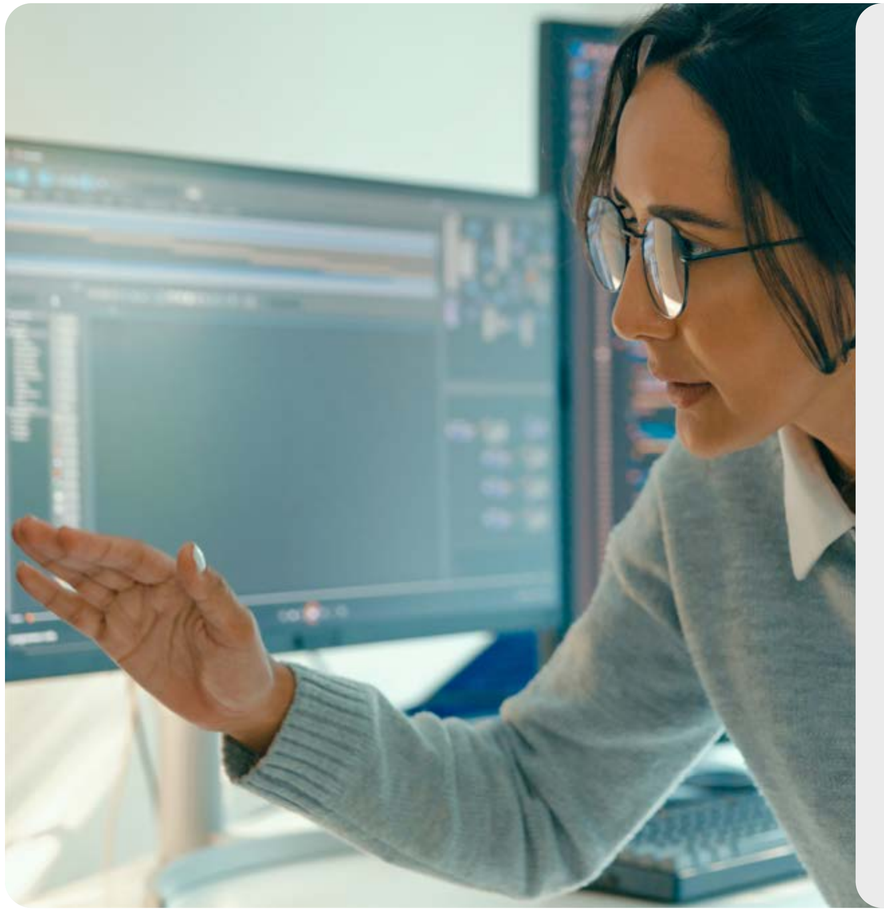
Maar soms vraagt de nieuwsgierige manager: “Wat komt er na "Agile”? Hoe kunnen we nog sneller innoveren? Kan AI ons daarbij helpen?”

Wendbaarheid is van cruciaal belang. Als je niet snel kunt innoveren of als je niet snel kunt reageren op veranderingen, dan ziet de toekomst er donker uit.

In dit artikel besteden we aandacht aan de volgende factoren die onze innovatiekracht vergroten: autonomie, AI, modulariteit en korte feedback loops.

Meer autonomie = Meer innovatiekracht

Als medewerkers zelf mogen beslissen over wat zij ontwikkelen en hoe ze dat doen, bespaart dat meetings, wachttijd en frustratie. Vanuit de high-level doelstellingen kunnen medewerkers vaak heel goed bedenken wat ervoor nodig is om die te bereiken. Zou het mogelijk zijn dat we medewerkers bijvoorbeeld zelf de procurement kunnen laten doen, als er iets aangekocht moet worden voor hun project, zoals bijvoorbeeld een robotarm? Zolang alle informatie zoals kosten, baten, effectiviteit, efficiency maar direct voorhanden zijn, moet iedereen slimme beslissingen kunnen nemen.



De Tesla case study 1:

Tesla heeft jarenlang gewerkt aan een “Digital Self-Management (DSM)” systeem. Elke medewerker krijgt apps op zijn of haar telefoon die live KPI’s tonen over kosten, baten, doelstellingen, problemen, knelpunten, testresultaten, kwaliteit, leverancierssprestaties etc. Als werknemer kom je ‘s ochtends in de fabriek, je checkt je apps, je kijkt waar de knelpunten zitten en JIJ bepaalt waar je die dag behulpzaam kunt zijn en waar je waarde kunt toevoegen.

Tesla heeft hierbij zwaar ingezet op AI voor het extraheren van data uit al deze systemen en het genereren van “actionable, up-to-date insights”.

Eigenlijk is het DSM systeem de moderne variant van de “Obeya room”. Het Japanse woord ‘obeya’ betekent ‘grote kamer’. Vaak is dat een fysieke ruimte, alhoewel virtuele ruimtes ook steeds meer het licht zien. Het is een plaats waar organisatie informatie visueel samengebracht is op borden of muren, zodat teams en management hier over kunnen discussiëren en plannen kunnen maken.

Het Digital Self Management system pusht echter continue de meest actuele informatie naar ieders telefoon. Zoals David Marquet al zei: “Move the authority to where the information is” of in dit geval “Move the information in order to empower”.

Intelligentie uit CRM en ERP systemen

Om innovatiekracht te vergroten is het dus nodig de data uit CRM, ERP en andere systemen slimmer te gebruiken, zodat medewerkers daadkrachtiger worden, sneller kunnen innoveren en betere beslissingen kunnen nemen.

Het is nu al mogelijk om met Co-pilot real-time advies te krijgen van de AI van Microsoft over ontwikkelingen in de wereld die de supply chain de komende tijd zullen uitdagen. Denk aan een melding aan medewerkers van een logistiek bedrijf over een nieuwsartikel, dat grote werkzaamheden aan de aangrenzende snelweg aankondigt en het effect hiervan op de planning. Ook binnen het klantenservice domein gebruiken bedrijven AI steeds vaker. CRM-systemen komen steeds dichter bij het punt waarbij AI alle taken van de klantenserviceagent overneemt. Klarna houdt miljoenen AI-klantenservicegesprekken in 23 landen in 35 talen per maand. Het bedrijf Bland AI gaat verder en heeft een AI-tool in de markt die in gesprek gaat met klanten, als eerstelijns klantenservice of sales-binnendienst. Voor de duidelijkheid, de AI-tool praat.... niet de medewerker.

Steeds vaker worden systemen geïntegreerd met AI of zijn zodanig gebouwd dat ze compatibel zijn met externe AI-software. Zo kunnen Microsoft Dynamics-gebruikers gebruik maken van Co-Pilot en hebben Salesforce-gebruikers Einstein Copilot tot hun beschikking.

De eerste stap die bedrijven zullen moeten maken is zich verdiepen in de licentiestructuur die de leveranciers aanbieden. In de meeste gevallen zullen additionele licenties moeten worden aangeschaft om gebruik te kunnen maken van de AI-oplossing.

Het tweede aandachtspunt heeft alles te maken met data. AI werkt met bedrijfsdata. Organisaties moeten zich de volgende vragen stellen: Zijn er systemen die geïntegreerd moeten zijn? Welke API's moeten geconfigureerd worden zodat de AI toegang heeft? Zijn er Sharepoints die geconfigureerd moeten worden? Zijn onze data geharmoniseerd? Antwoord op deze vragen zullen leiden tot gaps in de datamanagementstrategie, geoptimaliseerd voor AI.

Bij de volgende stap kunnen we verder kijken naar bijvoorbeeld implicaties voor RBAC (Role Based Access Control). AI zal toegang moeten krijgen tot data die hoort bij de gebruiker die op dat moment gebruik maakt van AI. Het is uiteraard niet de bedoeling dat een Inkoop medewerker toegang krijgt tot gevoelige HR-informatie van mede-collega's door aan AI te vragen: "Wat is het jaarsalaris van medewerker X?". AI moet gekoppeld worden aan de toegangsrechten van de gebruiker. De vraag is hoe dit ingesteld kan worden in specifiek situaties en in specifieke CRM- of ERP-systemen. Een stapje verder is om te denken aan dynamische RBAC, waarbij locatie, tijdstip en datatype, bepalend zijn voor AI-toegang, of gescheiden omgevingen voor bepaalde gevoelige informatie.

+[1] Turn the Ship Around!: A True Story of Turning Followers into Leaders Hardcover – May 16, 2013 by L. David Marquet (Author)

Case Study: De sales-assistent in de farmaceutische industrie

Recente ontwikkelingen maken het mogelijk dat AI-modellen open vragen kunnen beantwoorden.

Bij een biofarmaceutisch bedrijf vraagt een verkoper aan een op AI gebaseerd systeem wat de nieuwste ziekte- en voorschrijftrends voor diabetes in zijn regio zijn. Het systeem vertaalt deze vraag van natuurlijke taal naar computercode, die vervolgens wordt uitgevoerd om de nieuwste gegevens te analyseren. De verkregen inzichten worden gedeeld met de verkoper: "De diabetesprevalentie is met 6% gestegen, maar het behandelpercentage is met 12% gedaald. Het systeem geeft ook automatisch aan welke medicijnen het meest worden voorgeschreven en welke therapieën worden aangeraden. Aan de hand wetenschappelijke artikelen en nieuwsberichten signaleert het systeem direct een kans voor de verkoper: "Er is een groeiende belangstelling onder zorgverleners om te leren over alternatieve benaderingen". De verkoper kan op dit moment doorvragen, om meer te weten over deze nieuwsartikelen, of om meer te weten over de kansen in de markt. De verkoper is klaar voor actie en vraagt: "Welke prospects dien ik te benaderen?" Het systeem reageert met suggesties, zoals welke zorgverleners de verkoper persoonlijk moet bezoeken en welke de verkoper via e-mail moet contacteren. De verkoper combineert deze inzichten met eigen ervaringen en is dan beter voorbereid op relevante, inzichtelijke en productieve klantinteracties.

Korte feedbackloops

Het Digital Self Management systeem, zoals Tesla dat gebruikt, levert continue nieuwe “actionable insights” op en dat vraagt om snelle innovatiecycli. En snelle innovatie bereik je door in korte tijd een proof of concept neer te zetten, snel feedback op te halen, om vervolgens, als de eerste impressie goed is, hierop door te itereren.

Vaak zien we dat organisaties hier een doelstelling aan hangen: “Verbeter de XYZ in het volgende kwartaal met 10%”. Maar waarom zouden we niet met micro-doelstellingen kunnen werken? “Reduceer het gewicht van de auto met 10 gram in een iteratie van 3 uur”. Als het gehaald wordt heeft het team wat te vieren en is de dopamine rush de grootste motivator.

De Tesla case study 2

Tesla’s focus ligt op innovatiesnelheid en om snel te kunnen innoveren, zijn korte feedbackloops vitaal. Ze investeren fors in testautomatisering en hoe meer er geautomatiseerd kan worden, hoe beter. Hardware, robots, software, interfaces en apps, engineers zijn continu bezig met het updaten van geautomatiseerde tests. Zelfs de inspectie van de lak van Tesla’s is geautomatiseerd. Het naleven van de regelgeving is geautomatiseerd. Testrijden is geautomatiseerd. Auto’s testen in feite zichzelf. En AI analyseert de testresultaten.

“Ka-Chow, I’m okay!” (uit de film “Cars”)

Modulariteit

We kennen het allemaal wel: Wachten tot een ander team klaar is. Dan pas kunnen wij verder en dat is niet goed voor innovatie. Hoe meer een systeem uit onafhankelijke modules bestaat, hoe beter dat is en hoe minder wachttijd we doorlopen. Modules met duidelijke interfaces. Een team moet makkelijk een wijziging aan kunnen brengen in bijvoorbeeld de koplampen van een auto, zonder zich bezig te hoeven houden met afhankelijkheden met andere teams of met goedkeuringsprocedures. Ze hoeven geen vergaderingen te organiseren om het met andere teams eens te worden en ze hebben geen complexe, dure planningssessies nodig. Ze beslissen zelf omdat ze geloven dat dit het voertuig beter zal maken. De ondersteuning wordt geboden door AI-assistenten.

Maar hoe maken we een slimme AI-assistent? Het trainen met meer data leveren niet noodzakelijkerwijs een slimmer AI-systeem op. De foutgevoeligheid van AI’s kan significant worden verlaagd door de toegankelijkheid van data te limiteren. Voor de ontwikkeling van een AI-assistent kunnen we deze veel beter op een specialistische dataset trainen. Hoe beter en nauwkeuriger de antwoorden van de AI-assistent hoe meer verantwoordelijkheden we aan een dergelijk assistent kunnen geven. Hierdoor creëren we in feite een AI-conciërge. Een AI-conciërge geeft ook advies op basis van situatiefactoren en eerdere genomen keuzes. De volgende stap is een AI-agent. De AI-agent voert zelfstandig acties uit, zonder menselijke inbreng. Dit zijn bijvoorbeeld AI-marketing agenten, AI-warehouse planner agent of een AI-klantenservice agent, zoals die van Klarna.

“The future of AI will be on less data, not more” .

De innovatiemachine van de toekomst zal niet de agile frameworks van vandaag gebruiken, maar zal werken met dynamische, autonome teams, die in zeer korte cycli verbeteringen door kunnen voeren; cycli van uren in plaats van dagen of weken. De motor van deze machine heet AI.

[2] by H. James Wilson, Paul R. Daugherty, and Chase Davenport

January 14, 2019, Harvard Business Review, https://hbr.org/2019/01/the-future-of-ai-will-be-about-less-data-not-more

Over de auteurs



ZAHIR WANGA | Group Portfolio Manager

Zahir is Group Portfolio Manager bij Capgemini en houdt zich dagelijks bezig met de oplossingen die Capgemini te bieden heeft. Zahir heeft een brede oriëntatie op het gebied van business en IT en blijft constant up-to-date met de nieuwste ontwikkelingen.



RIK PENNARTZ | Applications Unleashed Chief editor, agile coach & Scrum Master

Rik is agile coach en gespecialiseerd in het vormen van high performing teams in innovatieve omgevingen. Rik is ook de initiatiefnemer van het Applications Unleashed rapport.





20 Waardengedreven architectuur

Jonge consumenten, waaronder Gen Z en millennials, voelen zich sterk aangetrokken tot merken en bedrijven die streven naar hogere doelen.[1] Belangrijke waarden zoals integriteit, innovatie en duurzaamheid moeten centraal staan in de bedrijfsvoering. Ik ben het eens met Simon Sinek, de bekende managementgoeroe, dat bedrijven niet alleen bestaan om geld te verdienen.[2]

Elke organisatie heeft haar eigen waarden. Dit zijn de fundamentele principes die een organisatie en haar medewerkers naleven in hun gedrag, werk en besluitvorming.[3]

Het doel van Enterprise Architectuur is om de organisatie in staat te stellen haar missie, bedrijfsstrategie en doelen zo effectief mogelijk te bereiken.[4] Daarom moeten kernwaarden ook in de Enterprise Architectuur worden geïntegreerd. Ik stel de volgende stappen voor om dit te doen:

1. **Identificeer de kernwaarden,**
2. **Definieer architectuurprincipes die deze waarden weerspiegelen,**
3. **Veranker waarden in governance,**
4. **Integreer je waarden in de ontwerpen,**
5. **Maak je waarden eigen en communiceer daarover,**
6. **Stem doelen en meetcriteria af op de waarden,**
7. **Bevorder ethische besluitvorming en**
8. **Verbeter continu. Door 'Design for Values'[5] toe te passen, waarbij producten en diensten worden gecreëerd die maatschappelijke en morele waarden realiseren, wordt dit proces aanzienlijk vergemakkelijkt.**

Het opnemen van waarden in de Enterprise Architectuur zal leiden tot betere economische waardeketens en tot meer tevredenheid voor alle belanghebbenden, zowel intern als extern. Enterprise Architectuur moet niet alleen economische waarde creëren, maar ook kernwaardeketens realiseren die het leven van klanten, medewerkers en alle andere betrokkenen ten goede komen.

1. Anna Schaverien: Consumers Do Care About Retailers' Ethics And Brand Purpose, Accenture Research Finds; 10 Dec 2021; https://www.forbes.com/sites/annaschaverien/2018/12/12/consumers-do-care-about-retailers-ethics-and-brand-purpose-accenture-research-finds/
2. Simon Sinek: Businesses do not exist to make money; 11 Feb 2021; https://x.com/simonsinek/status/1359944838226001925
3. Beth Mason: How to make values grow and flourish as a meaningful part of the business; 18 Aug 2021; https://www.employment-studies.co.uk/news/how-make-values-grow-and-flourish-meaningful-part-business
4. The Open Group: The Purpose of Enterprise Architecture; 2022; https://pubs.opengroup.org/togaf-standard/adm-practitioners/adm-practitioners_3.html
5. TU Delft: Design for Values; 17 Jul 2024; https://www.delftdesignforvalues.nl

Over de auteur



Reinoud Kaasschieter | Enterprise Architect and AI expert

Reinoud is een enterprise architect in de Nederlandse publieke sector en expert op het gebied van kunstmatige intelligentie en ethiek. Hij deelt zijn kennis en inzichten actief door te publiceren en spreken over AI en ethiek, zowel binen Capgemini als daarbuiten.

Colofon

Het Capgemini Applications Unleashed Team 2025

Editor in Chief: Rik Pennartz

Design: Puja Sengupta, Johanna Achterberg, Arundhati Bhattacharya

Branding: Els van der Velden, Pamela Datta

Project en Marketing Management: Thomas de Klerk, Rebecca Scott

Thought leaders:

Henk Vermeulen
Donald Hessing
Arin Roy
Douwe van de Ruit
Sjoukje Zaal
Tom van de Ven
Reinoud Kaasschieter
Mandy Steenbergen
Marieke van de Putte
Rahul Rauniyar
Randy van Breugel
Mahendra Rao Bandaru
Niels Brouwers
Joe Reynolds
Yuri Blankenstein
Lambert Verhagen
Josieke Moens
Nipun Kaushik
Herman Roebbers
Adrienn Matis
Hans van Rijs
Rik Pennartz
Zahir Wanga

Met speciale dank aan:

Ron Tolido
Roger Wannee
Sjoukje Zaal
Gerard Peters
Han Konings
Fina Piazza
Thijs Geurts

Over Capgemini

Capgemini is een wereldwijde partner voor bedrijfs- en technologie-transformatie, die organisaties helpt hun duale transitie naar een digitale en duurzame wereld te versnellen, terwijl het tastbare impact creëert voor bedrijven en de samenleving. Het is een verantwoordelijke en diverse groep van 340.000 teamleden in meer dan 50 landen. Met zijn sterke erfgoed van meer dan 55 jaar wordt Capgemini door zijn klanten vertrouwd om de waarde van technologie te ontsluiten en in te spelen op de gehele breedte van hun bedrijfsbehoeften. Capgemini biedt end-to-end diensten en oplossingen, waarbij het gebruik maakt van sterke capaciteiten op het gebied van strategie en ontwerp tot engineering, allemaal aangedreven door de marktleidende capaciteiten in AI, generatieve AI, cloud en data, gecombineerd met diepgaande industrie-expertise en een partnerecosysteem. De Group rapporteerde een wereldwijde omzet in 2024 van €22,1 miljard.

Get the Future You Want | www.capgemini.nl

Capgemini Nederland B.V.
Postbus 2575 - 3500 GN Utrecht
Tel. + 31 30 203 05 00
www.capgemini.nl

De informatie in dit document is eigendom van Capgemini. Copyright © 2025 Capgemini.
Alle rechten voorbehouden.

2A-020.24