



AI Principes

Hanteer principes voor de inzet van AI

Inleiding

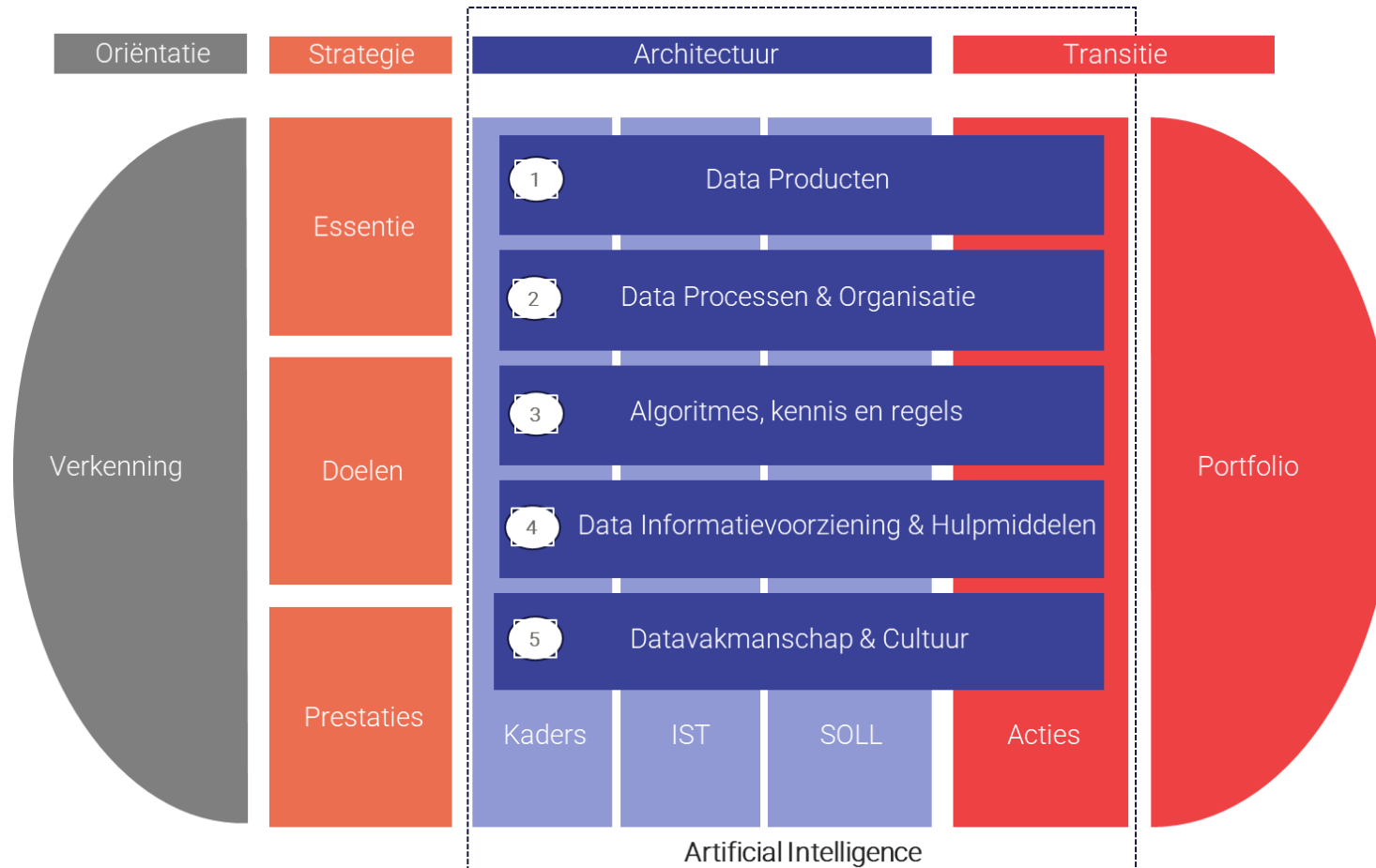
- Solventa benadert AI met een sterke nadruk op ethiek, verantwoordelijkheid en strategische inzet.
- Als organisatie stellen we dat AI en datagedreven werken alleen effectief kunnen zijn wanneer ze worden ondersteund door professioneel datamanagement en duidelijke ethische kaders. Dit houdt in dat organisaties zorgvuldig omgaan met data en transparant zijn over hoe AI-toepassingen worden ingezet.
- Wij als Solventa staan voor de volgende zaken en AI mag daarin niet ontbreken:
 - Ethische omgang met data: Respecteer de privacy en wensen van gebruikers, en stel duidelijke grenzen aan data-gebruik.
 - Verantwoord gebruik van AI: Zorg dat AI-technologieën bijdragen aan duurzame, betrouwbare besluitvorming.
 - Betrouwbaarheid van AI: De uitkomsten van AI moeten kloppen en dat heeft gevolgen voor het vertrouwen in AI en besluiten die daarmee gemoeid zijn.
 - Transparantie en verantwoording: Wees helder over hoe AI werkt en welke impact het heeft, vooral bij zelflerende technologie.



Solventa denkraam

Referentie voor ons werken en denken in architectuur & transformaties op het gebied van data

Het Solventa denkraam is een referentiekader voor denken en werken in architectuur en transformaties. Het biedt tevens een kader voor het omgaan met AI-vraagstukken. Zoals je van ons als architecten mag verwachten hebben we voor ieder van de vijf lagen van de architectuur, die in onderstaand denkraam te zien zijn, een principe gedefinieerd.



Principe #1: Data Producten

Titel	Zet AI enkel in voor (data) producten t.a.v. het bereiken van maatschappelijke of economische meerwaarde
Toelichting	Het grotere bereik van nieuwe technologieën, waaronder AI, raakt steeds meer de samenleving als geheel. Vooruitgang op het gebied van AI zal transformerende gevolgen hebben op een groot aantal gebieden, waaronder gezondheidszorg, energie, transport, productie en onderwijs, mits deze gericht zijn op het bereiken van maatschappelijke en economische meerwaarde. Wanneer we de potentiële ontwikkeling en het gebruik van AI-technologieën overwegen, zullen we ons richten op toepassingen die duidelijk bijdragen aan het welzijn van de samenleving of economische vooruitgang bevorderen. Hierbij houden we rekening met de sociale en economische factoren en kiezen we alleen voor inzet waar de verwachte voordelen aanzienlijk groter zijn dan de voorzienbare risico's en nadelen.
Rationale	Als mensen geloven dat AI-systemen veilig en betrouwbaar zijn, zijn ze eerder geneigd ze te accepteren en te gebruiken. AI wordt ingezet om complexe maatschappelijke problemen aan te pakken voor het verbeteren van het welzijn van de samenleving en ongelijkheid verminderen. AI draagt niet alleen bij aan de winstgevendheid, maar ook aan het versterken van de winstgevendheid en concurrentiekracht.
Implicatie	<p>De impact van het gebruik van kunstmatige intelligentie (AI) op een organisatie kan aanzienlijk zijn en heeft zowel positieve als uitdagende aspecten.</p> <ul style="list-style-type: none">• Maatschappelijke vooruitgang: dit kan leiden tot meer inclusiviteit, betere toegang tot basisvoorzieningen en een hogere kwaliteit van leven.• Automatisering en efficiëntie: AI kan repetitieve of routinematige taken automatiseren, waardoor medewerkers meer tijd hebben voor complexere en strategische en creatieve activiteiten. Dit leidt tot hogere productiviteit en kostenbesparingen voor organisaties.• Data-analyse en besluitvorming: AI kan grote hoeveelheden gegevens analyseren en waardevolle inzichten bieden. Dit helpt bij betere besluitvorming en strategische planning.• Toegankelijkheid: Er wordt gewerkt aan het democratiseren van AI-technologieën, zodat deze beschikbaar worden voor een breder publiek en niet alleen voor grote bedrijven.• AI in Creativiteit: AI wordt ook steeds meer gebruikt voor de creatie van bijvoorbeeld (architectuur)modellen en oplossingsrichtingen, waar het nieuwe vormen van expressie en samenwerking mogelijk maakt.• Verdere Integratie: AI zal waarschijnlijk steeds meer geïntegreerd worden in zowel het persoonlijke als professionele leven, met toepassingen in smart homes, gezondheidszorg op afstand, en gepersonaliseerde media.• Door processen te optimaliseren en fouten te verminderen, kunnen kosten worden bespaard.• Innovatie en nieuwe werkwijzen: AI stelt organisaties in staat om nieuwe manieren van werken te ontdekken, vooral in combinatie met algoritmen. Creatieve toepassingen van AI, zoals ChatGPT, kunnen leiden tot nieuwe vormen van communicatie en contentgeneratie.• Personalisatie: AI kan diensten en producten aanpassen aan individuele voorkeuren, wat leidt tot een meer gepersonaliseerde gebruikerservaring.



Principe #2: Data processen en organisatie

Titel	Menselijke controle over AI-systemen is vereist
Toelichting	AI-systemen kunnen een indrukwekkende ondersteuning bieden bij de uitvoering van processen en specifiek bij het menselijk beslissingsproces. Mensen moeten echter in staat zijn om controle te houden over AI-systemen en moeten de mogelijkheid hebben om beslissingen van deze systemen te begrijpen, te beïnvloeden en te corrigeren.
Rationale	<p>Complexe algoritmes die gebruikt worden voor AI-systemen kunnen feitelijk niet in detail door een mens gecontroleerd worden, dit is ook juist de kracht van AI-systemen (het overstijgt de capaciteit van de mens).</p> <p>AI-systemen hebben de potentie om besluitvorming en procesoptimalisatie aanzienlijk te verbeteren door snelheid, precisie en datagestuurde inzichten te bieden. Echter, zonder menselijke controle kan er een risico ontstaan op verkeerde interpretaties, vooroordelen, of onbedoelde gevolgen. Terwijl AI sterk is in patroonherkenning en dataverwerking, begrijpen mensen de bredere context en morele implicaties beter, waardoor de besluitvorming rijker en zorgvuldiger wordt. Het principe dient dus als een fundamentele waarborg dat AI-systemen een hulpmiddel blijven, in plaats van autonome entiteiten die de menselijke verantwoordelijkheid verdringen. Dit stelt mensen in staat om controle te behouden over de impact en het gebruik van AI in de samenleving.</p>
Implicatie	<p>Ontwerp AI-systemen of pas AI-systemen toe die passende mogelijkheden bieden voor menselijke leiding en controle, feedback en relevante uitleg.</p> <ul style="list-style-type: none">• Beslissingsmacht: Mensen blijven uiteindelijk verantwoordelijk voor de impact van beslissingen die AI-systemen nemen. Daarom moeten mensen altijd de uiteindelijke controle behouden over belangrijke beslissingen die door AI-systemen worden genomen. Dit betekent dat AI-systemen aanbevelingen kunnen doen, maar dat mensen de eindbeslissing moeten maken, vooral in gevoelige domeinen zoals gezondheidszorg, rechtspraak en financiën.• Transparantie: Begrijpelijke en beïnvloedbare AI-besluiten stellen gebruikers in staat om volledig inzicht te krijgen in de werking van het systeem, waardoor onduidelijkheid en wantrouwen worden voorkomen.• Interventiemogelijkheden: Mensen moeten in staat zijn om in te grijpen in de werking van AI-systemen, vooral als er signalen zijn dat het systeem foutieve of ongewenste resultaten oplevert. Dit kan betekenen dat er mechanismen moeten zijn om een AI-systeem te pauzeren, te corrigeren of zelfs uit te schakelen.• Feedback: Menselijke handelingen automatiseren ligt voor de hand bij de inzet van AI, maar creëer een feedback loop om de systemen te laten leren van eerder gemaakte foutieve beslissingen, vooroordelen en onbedoelde gevolgen die zijn bijgestuurd door menselijk handelen.



Principe #3: Algoritmes, kennis en regels

Titel	AI wordt ingezet bij het voldoen aan wetgeving, privacy, ethiek en menselijke waardigheid
Toelichting	<p>AI-algoritmen en datasets zullen onrechtvaardige gevolgen voor mensen vermijden en met name rekening houden met gevoelige kenmerken op een manier die eerlijk, onpartijdig en vrij van discriminatie is. Naast wat de wet van ons eist, is het onze plicht AI te allen tijde ethisch verantwoord toe te passen, AI-systemen worden ontworpen, ontwikkeld en ingezet op een manier die de veiligheid en privacy van gebruikers en de samenleving als geheel waarborgt. Deze systemen moeten streven naar hoge kwaliteit en nauwkeurigheid in hun besluitvorming en werking. Daarvoor is het nodig dat AI-systemen zowel een democratische en gelijkwaardige samenleving mogelijk maken door de gebruikersautonomie te garanderen, als fundamentele rechten respecteren, wat dient te worden afgedwongen door middel van menselijk toezicht.</p>
Rationale	<ul style="list-style-type: none">AI-systemen moeten eerlijk, juridisch correct en onpartijdig zijn om te voldoen aan ethische normen en wettelijke vereisten. Discriminatie in AI kan leiden tot ongelijke toegang tot kansen, diensten en middelen voor verschillende bevolkingsgroepen. Helpt het vertrouwen van gebruikers en stakeholders in AI-systemen te vergroten. Mensen moeten erop kunnen vertrouwen dat AI-systemen eerlijke en rechtvaardige beslissingen nemen.AI-systemen gebruiken grote volumes data om via algoritmes tot waardevolle informatie te komen. Indien de data kwaliteit te wensen over laat van de input data zal dit zijn uitwerking hebben op de resultaten van de AI-systemen. De uitkomsten kunnen daarmee waardeloos zijn en/of verkeerde beelden opwekken.Bij het ontwerpen AI-systemen wordt rekening gehouden met het beschermen van persoonlijke gegevens zodat deze niet misbruikt of onthuld worden zonder toestemming van de individu zelf. Individuen hebben het recht om controle te hebben over hun persoonlijke informatie en om te beslissen wie toegang heeft tot deze informatie. Verder moeten we vertrouwen hebben in de AI-systemen. Om vrijblijvendheid te vermijden stellen we per risicocategorie maatregelen voor die we moeten treffen bij het gebruik van AI.AI-systemen worden ontworpen met het oog op het minimaliseren van de kans op stereotypering op basis van demografie, cultuur of andere factoren. Het inzetten van AI binnen deze domeinen versterkt de transparantie en vergroot het vertrouwen van gebruikers, overheden en het publiek in technologie en organisaties die AI toepassen. Wel moet er met gezond verstand naar de uitkomst van een algoritme gekeken worden.
Implicatie	<p>De impact van wetgeving, privacy, ethiek en menselijke waardigheid is als volgt:</p> <ul style="list-style-type: none">Gegevensminimalisatie: Zorg ervoor dat AI alleen de minimale hoeveelheid gegevens verzamelt die nodig zijn voor het beoogde doel en zorg dat deze niet langer bewaard worden dan nodig.Vermijden van vooroordelen<ul style="list-style-type: none">Begint met het zorgvuldig selecteren en opschonen van datasets die worden gebruikt om AI-systemen te trainen. Als de trainingsdata vooroordelen bevatten, zullen de AI-resultaten dat ook doen. Het bevorderen van diversiteit en inclusie opnemen in AI-teams om culturele en sociale perspectieven in het ontwikkelingsproces te integreren.AI-ontwikkelaars moeten methoden en technieken implementeren om vooroordelen en ethische AI-beleid in AI-systemen op te sporen, te handhaven en te verminderen. Dit omvat het regelmatig controleren en aanpassen van de modellen om eerlijkheid te waarborgen.Organisaties moeten medewerkers en gebruikers bewust maken en hen voorzien van de nodige kennis en hulpmiddelen om deze te identificeren en te corrigeren.AI-ontwikkelaars moeten methoden en technieken implementeren om vooroordelen in AI-systemen op te sporen en te verminderen. Dit omvat het regelmatig controleren en aanpassen van de modellen om eerlijkheid te waarborgen.Voldoen aan bestaande wet- en regelgeving met betrekking tot de veiligheid van AI-systemen en organisaties moeten op de hoogte blijven van nieuwe regelgeving die kan worden ingevoerd.Innovatie en nieuwe werkwijzen: AI stelt organisaties in staat om nieuwe manieren van werken te ontdekken, vooral in combinatie met algoritmen. Creatieve toepassingen van AI, zoals Copilot, ChatGPT en Bard kunnen leiden tot nieuwe vormen van communicatie en contentgeneratie.Transparantie in AI-systemen:<ul style="list-style-type: none">AI-ontwikkelaars, organisaties, gebruikers moeten verantwoordelijkheid nemen bij ontwikkelen van AI-systemen om open te zijn in hoe en waarom een bepaalde besluit genomen wordt.Het vereist methoden voor het uitleggen van beslissingen, het auditen van datasets en het waarborgen van traceerbaarheid in algoritmen.Het delen van informatie van AI-systemen kan privacy-kwesties met zich meebrengen. De balans vinden tussen transparantie en het beschermen van persoonlijke gegevens is hier van belang.Toestemmingsvereisten: Zorg dat de individuen geïnformeerd worden over hoe hun gegevens worden gebruikt en vraag om hun toestemming voordat gegevens worden verzameld.

Principe #4: Data informatievoorziening & Hulpmiddelen

Titel	Informatievoorziening met daarin AI-functionaliteit is transparant en verantwoord opgezet
Toelichting	<p>Niet iedere applicatie die AI-functionaliteit in zich heeft moet op gelijke manier behandeld worden. Het is aan de organisatie hoe hieraan vorm te geven. Applicaties kunnen geclassificeerd worden naar risico-categorie. De categorieën zijn vervolgens bepalend voor de eisen die gesteld worden aan de applicaties en de maatregelen die genomen moeten worden, dit in relatie tot transparantie en verantwoord gebruik. Voor de toepassing van AI kijken wordt gekeken naar het type van de toepassing en de impact daarvan op de omgeving van onze organisatie (intern en extern).</p> <p>T.a.v. de te nemen maatregelen:</p> <ul style="list-style-type: none">• Systemen moeten veilig zijn en bescherming bieden tegen kwaadaardig gebruik, hacking en andere bedreigingen voor de veiligheid en het waarborgen van vertrouwelijkheid van informatie. Dit draagt bij aan het beschermen van de mentale en fysieke integriteit van natuurlijke personen.• Adequate mechanismen voor data governance moeten worden gewaarborgd, waarbij rekening wordt gehouden met de kwaliteit en integriteit van de gegevens en een gelegitimeerde toegang tot de gegevens wordt gewaarborgd.
Rationale	<p>De hoeveelheid data die verwerkt is en het algoritme dat daarover is toegepast is niet door het menselijk brein te overzien. Het garanderen van de veiligheid van AI-systemen (op het gebied van transparantie en verantwoord handelen) helpt in het vertrouwen dat gebruikers en stakeholders hebben in het systeem. De EU AI Act is risico-gebaseerd opgezet en stelt maatregelen voor per risicocategorie. Voor een organisatie die grip wil krijgen op een applicatielandschap dat rijk is aan AI-functionaliteit ligt een risico gebaseerde aanpak daarmee voor de hand, en daarmee moet de organisatie de maatregelen die vanuit de EU AI Act gesteld worden omarmen.</p>
Implicatie	<p>De impact voor de organisatie op het gebied van informatievoorziening i.r.t. de inpasbaarheid van AI-systemen (of systemen die dergelijke functionaliteit bevatten):</p> <ul style="list-style-type: none">• De organisatie is in staat om de toepassing van AI te classificeren naar zijn risico-categorie (verboden, hoog, beperkt, minimaal of geen).• Per risicocategorie worden maatregelen gedefinieerd.• AI-ontwikkelaars en organisaties moeten beveiligingsexpertise in huis halen om AI-systemen te beschermen tegen aanvallen en inbreuken. Dit omvat het identificeren van kwetsbaarheden en het implementeren van beveiligingsmaatregelen.• Het gaat niet alleen om het voorkomen van schade, maar ook om het waarborgen van eerlijke en rechtvaardige (ethische) resultaten, met name in geautomatiseerde besluitvorming. Het vermijden van vooroordelen begint met het zorgvuldig selecteren en opschonen van datasets die worden gebruikt om AI-systemen te trainen. Als de trainingsdata vooroordelen bevatten, zullen de AI-resultaten dat ook doen.• Zorg er dus voor dat je niet zomaar persoonlijke data van klanten, ketenpartners of werknemers gebruikt voor AI; vraag altijd toestemming. En check ook het privacybeleid van de technologie – bijvoorbeeld of gegevens worden gedeeld met andere partijen.



Principe #5: Datavakmanschap & Cultuur

Titel	Menselijke gedrag met AI-systemen wordt bepaald door de kennis en expertise van de medewerkers
Toelichting	De kennis en vaardigheden die professionals hebben om effectief te werken in een omgeving waar kunstmatige intelligentie (AI) een steeds grotere rol speelt. Daarbij gaat het om dat de AI-ready medewerker een basisbegrip heeft van datavakmanschap en de cultuur kent om van AI te weten wat en hoe het wordt ingezet om effectief en verantwoord gebruik te maken van AI.
Rationale	<ul style="list-style-type: none">• AI-systemen zijn slechts zo effectief als de mensen die ermee werken. Als medewerkers niet over de juiste kennis en expertise beschikken, kunnen zij de systemen niet optimaal benutten of verkeerd interpreteren wat kan leiden tot verkeerde interpretaties, inefficiëntie of fouten.• Menselijke expertise is nodig om de ethische implicaties van AI-gebruik te begrijpen en verantwoorde beslissingen te nemen. AI-systemen hebben geen moreel besef; dat ligt in handen van de gebruikers.• De gebruikers beschermen tegen schadelijke gevolgen van fouten of misbruik van AI-systemen. Het is de verantwoordelijkheid van AI-ontwikkelaars en organisaties om de veiligheid van gebruikers te waarborgen. Het garanderen van de veiligheid van AI-systemen helpt vertrouwen op te bouwen bij gebruikers en stakeholders. Als mensen geloven dat AI-systemen veilig en betrouwbaar zijn, zijn ze eerder geneigd ze te accepteren en te gebruiken.• AI verandert constant door nieuwe ontwikkelingen en updates. Medewerkers met kennis en expertise kunnen zich sneller aanpassen aan deze veranderingen en blijven relevant in een technologische omgeving.
Implicatie	<p>De impact voor de medewerker binnen de organisatie om te werken met AI-systemen voor de ontwikkeling en welzijn van de werknemer.</p> <ul style="list-style-type: none">• Zorgen en onzekerheid. Werknemers maken zich soms zorgen over de ontwikkeling van AI. Vragen over veiligheid, banenverlies en vooroordelen spelen hierbij een rol. Het is belangrijk om deze zorgen serieus te nemen en transparantie te bieden over het gebruik van AI.• Welzijn van medewerkers beschermen. Organisaties moeten verantwoordelijk omgaan met AI om het welzijn van hun medewerkers te waarborgen. Dit omvat het minimaliseren van negatieve gevolgen, het bieden van bijscholing en hulpmiddelen en het creëren van een ethisch kader.• Mensen moeten echter in staat zijn om controle te houden over AI-systemen en moeten de mogelijkheid hebben om beslissingen van deze systemen te begrijpen, te beïnvloeden en te corrigeren. Zie ook toelichting principe 2 en 4.• Verantwoordelijkheid mens en organisatie. AI-systemen nemen beslissingen, maar de verantwoordelijkheid blijft bij de mens en organisatie. Dit kan ethische vraagstukken en extra verantwoordelijkheid met zich meebrengen.• Samenwerking tussen mens en AI. AI fungeert vaak als ondersteuning voor menselijke beslissingen, wat een nieuwe manier van samenwerken vereist.• Cultuurverandering binnen de organisatie. De introductie van AI kan leiden tot veranderingen in hoe teams met elkaar werken en communiceren.

