

De impact van AI op de financiële sector en het toezicht

DeNederlandscheBank

EUROSYSTEEM



Hoofdboodschappen

- Artificiële intelligentie (AI) is wereldwijd bezig met een opmars in de financiële sector, ook in Nederland. AI brengt naast substantiële kansen ook risico's met zich. Enerzijds kan AI leiden tot betere dienstverlening voor klanten en meer gepersonaliseerde producten, en biedt het instellingen de kans om inkomsten te verhogen en kosten te verlagen. Anderzijds zijn er risico's, onder andere op het gebied van datakwaliteit, gegevensbescherming, uitlegbaarheid, incorrecte resultaten, discriminatie en uitsluiting, en een hogere mate van afhankelijkheid van derde partijen.
- Van financiële instellingen mag verwacht worden dat zij AI verantwoord inzetten. De doelstellingen van het toezicht en de normen waaraan instellingen moeten voldoen zijn onafhankelijk van de gehanteerde technologie. Ze gelden onverkort ook als AI wordt toegepast. Het gebruik van AI maakt onderdeel uit van bestaande processen en dienstverlening waar door de AFM en DNB toezicht op wordt gehouden op grond van geldende financiële wet- en regelgeving.
- AI heeft gevolgen voor de manier waarop de AFM en DNB toezicht houden. Om het gebruik van AI door instellingen te kunnen beoordelen zullen de AFM en DNB hun kennis op dit terrein moeten uitbreiden. Toezichtmethoden en procedures zullen in bepaalde gevallen ontwikkeld of aangepast moeten worden om rekening te houden met nieuwe technieken. Het toezicht zal zich onder meer richten op het risicobeheer rond de inzet van AI-toepassingen, de modaliteiten van die toepassingen en de uitkomsten.
- Afhankelijk van de toepassing van AI kunnen aanvullende vereisten nodig zijn. Een bestaand voorbeeld hiervan zijn de regels die gelden voor algoritmische handel op de kapitaalmarkt. Op dit moment zijn dergelijke concretisering in wetgeving voor het verantwoord gebruik van AI beperkt. Naarmate het gebruik en het belang van AI voor de financiële sector toeneemt, dient het normenkader te worden verduidelijkt dan wel gespecificeerd. Bij voorkeur gebeurt dit geharmoniseerd op Europees niveau, mede ten behoeve van toezichtconvergentie en het gelijke speelveld.
- In wet- en regelgeving moet een goede balans worden gevonden tussen enerzijds verantwoorde inzet van AI en anderzijds ruimte voor innovatie.
- De Europese AI-verordening wijst AI-systemen, die gebruikt worden voor kredietwaardigheidstoetsen van natuurlijke personen en voor premiebepaling en risicobeoordeling voor levens- en ziektekostenverzekeringen, aan als hoog risico. Dit betekent dat nadere vereisten zullen gelden voor de ontwikkeling en het beheerst en verantwoord gebruik van deze AI-toepassingen. De vereisten sluiten grotendeels aan op bestaande regels voor risicobeheer en governance. Voor andere AI-toepassingen worden instellingen aangemoedigd om vrijwillig aan vereisten voor hoog risico AI-systemen te voldoen, onder meer door het volgen van Europese richtsnoeren voor betrouwbare AI. De AFM en DNB sluiten zich hierbij aan.
- De AI-verordening zal van financiële instellingen die AI toepassen extra aandacht voor een goede bescherming van grondrechten vergen. Voor hoog risico toepassingen wordt het verplicht voor instellingen om een beoordeling te maken van het mogelijke effect op grondrechten van personen of groepen.
- Voor zover AI wordt ingezet voor financiële dienstverlening, is uitgangspunt van de AI-verordening dat bestaande financieel toezichthouders ook het toezicht op de regels van de verordening op zich nemen. In Nederland zijn dat de AFM en DNB. Zowel Europees als nationaal is brede coördinatie en samenwerking tussen toezichthouders op AI wenselijk, mede gezien de bijzondere risico's van AI-systemen ten aanzien van grondrechten.

Samenvatting

De snelle ontwikkeling van artificiële intelligentie (AI) stelt AFM en DNB voor uitdagingen in het toezicht. AI is wereldwijd bezig met een opmars, ook in de financiële sector. AI brengt naast veel kansen, ook aanzienlijke risico's voor instellingen en de samenleving. Wanneer risico's van AI zich materialiseren kan dit een impact hebben op de financiële soliditeit en integriteit van instellingen, en het klantbelang en de verhoudingen tussen marktpartijen op financiële markten raken. Deze ontwikkelingen stellen de AFM en DNB voor de vraag hoe verantwoord gebruik van AI in de financiële sector kan worden geborgd en het toezicht op AI kan worden vormgegeven. Het toezicht is erop gericht dat instellingen de risico's van AI-toepassingen adequaat beheersen terwijl er voldoende ruimte blijft voor innovatie binnen de wet- en regelgeving. AI verandert snel en regelgeving verandert mee. De uitkomsten van dit rapport bieden daarom geen pasklare antwoorden op de weergegeven vraagstukken. Het doel van het rapport is om uitgangs- en aandachtspunten te formuleren voor het vormgeven van het toezicht op AI.

AI gaat in essentie om computersystemen die taken verrichten waar gewoonlijk menselijke intelligentie voor nodig is, of die overtreft.

AI-toepassingen zijn geavanceerde statistische benaderingen van grote hoeveelheden data waarmee voorspellingen, aanbevelingen en nieuwe inhoud gegenereerd kunnen worden. Onder invloed van een exponentiële toename van rekenkracht en data en vooruitgang in de computerwetenschap, hebben AI-toepassingsmogelijkheden een grote vlucht genomen. De opkomst van generatieve AI, waarbij nieuwe inhoud kan worden gecreëerd, markeert een nieuwe fase in de evolutie van AI.

Nederlandse financiële instellingen gebruiken al langer AI en experimenteren met geavanceerdere AI-modellen, wat doet vermoeden dat het gebruik van AI in de komende jaren verder toeneemt. AI wordt bijvoorbeeld ingezet voor fraudepreventie en -detectie, het tegengaan van witwassen en terrorismefinanciering en cybercriminaliteit, kredietbeoordelingen en identiteitsverificatie. Ook helpt AI medewerkers om efficiënter te werken. Zo worden er bij instellingen automatische samenvattingen gemaakt van telefoongesprekken met klanten. Nederlandse instellingen geven aan dat zij groot belang hechten aan het op ethische wijze gebruiken van AI, en dat ze niet de grenzen opzoeken van wat er mag en mogelijk is volgens huidige en aanstaande wet- en regelgeving. Ook geven veel instellingen aan voorlopig terughoudend te zijn met het gebruik van generatieve AI, al zien ze wel de potentie van deze techniek. Sommige zetten generatieve AI langzaam in voor ondersteunende processen, en in de toekomst zullen er naar verwachting nieuwe gebruiksmogelijkheden worden ontdekt die financiële instellingen gaan inzetten in hun bedrijfsvoering. Ook concurrentieoverwegingen zouden ertoe kunnen leiden dat financiële instellingen in de toekomst meer AI gaan inzetten. Als ze hierin te terughoudend zijn, kan dat mogelijk nadelig uitpakken voor hun concurrentiepositie ten opzichte van buitenlandse financiële instellingen en niet-financiële partijen die zich ook op de Nederlandse markt begeven.

AI biedt substantiële kansen voor zowel klanten als voor financiële instellingen. AI biedt mogelijkheden om slimmer (efficiënter) gebruik te maken van data. Hierdoor kunnen bijvoorbeeld producten worden aangeboden die beter aansluiten op de individuele behoeften van

de klant. Ook kan AI een snellere dienstverlening mogelijk maken en kan het mogelijk de drempel voor toegang tot financiële diensten verlagen. AI biedt financiële instellingen kansen om inkomsten te verhogen, door nieuwe klantengroepen aan te boren, cross-selling mogelijkheden te vergroten en prijzen te optimaliseren. Achter de schermen kan AI leiden tot efficiëntere processen, wat de kosten voor instellingen kan verlagen. Daarnaast kan AI financiële instellingen helpen om risico's beter en consistent in te schatten, mits zij beschikken over kwalitatief goede en complete data en modellen die uitlegbaar en controleerbaar zijn. AI heeft tot slot de potentie om de verdediging tegen cyberaanvallen te verbeteren en de bestrijding van financieel-economische criminaliteit effectiever en efficiënter te maken.

Het gebruik van AI-toepassingen gaat echter ook gepaard met risico's die een impact kunnen hebben op financiële instellingen en hun klanten. Deze risico's bevinden zich daarmee ook in het toezichtdomein van de AFM en DNB. AI kan bijvoorbeeld leiden tot maatschappelijk ongewenste uitkomsten, door discriminatie en uitsluiting, wat het gevolg kan zijn van het gebruik van biased (vooringenomen) data of data waarvan de kwaliteit onvoldoende is. Het feit dat grote hoeveelheden data moeten worden verwerkt leidt ook tot risico's op het gebied van beveiliging en gegevensbescherming. Geavanceerde AI-modellen kunnen bovendien complexe beslissingen nemen die niet duidelijk te verklaren zijn, waardoor de uitlegbaarheid en transparantie van financiële uitkomsten kan worden ondermijnd. Verder kan het gebruik van AI door financiële instellingen leiden tot afhankelijkheidsrisico's, mede omdat er vooralsnog slechts een beperkte groep (vooral niet-Europese) technologiepartijen is die AI-technologie kan leveren. Andere risico's zijn gebrekkige governance raamwerken en het grote stroomverbruik van AI.

In wet- en regelgeving moet een goede balans worden gevonden tussen enerzijds verantwoorde inzet van AI en anderzijds ruimte voor innovatie. Wat de impact van AI op de financiële sector zal zijn in de komende jaren is moeilijk in te schatten. Een scenario-analyse naar ontwikkelingen ten aanzien van de ontwikkeling van regelgeving en innovatie in AI-toepassingen onderstreept het belang van evenwicht. Een scenario waarin een evenwicht wordt bereikt tussen groeiende innovatie en adapterende regelgeving lijkt de meest gunstige uitkomst te bieden. Daarentegen lijken scenario's waarin er een onevenwichtigheid is tussen regelgeving en innovatie tot minder gunstige resultaten te leiden. Dit benadrukt het belang van een proactieve en adaptieve houding van beleidsmakers en toezichthouders. Regelgeving moet in staat zijn om een kader te creëren voor verantwoordelijkheid en transparantie waar tegelijkertijd ook genoeg ruimte is voor innovatie en snelle ontwikkelingen in AI producten en diensten.

De doelstellingen van het toezicht en de normen waaraan instellingen moeten voldoen zijn onafhankelijk van technologie die instellingen hanteren en gelden onverkort ook als AI wordt toegepast. Van instellingen mag worden verwacht dat ze AI op een verantwoorde manier inzetten. AI-systemen mogen de financiële soliditeit en integriteit van financiële instellingen niet in gevaar brengen. Evenmin mag het gebruik van AI ertoe leiden dat het klantbelang en de zuivere verhoudingen tussen marktpartijen worden aangetast. Het bestaande toezichtrechtelijke kader wordt door de AFM en DNB ook op AI-systemen toegepast. Zo dienen instellingen te voldoen aan de vereisten van een integere en beheerste bedrijfsvoering. Hiertoe dienen ze een adequaat risicomanagement in

te richten, dat ook de risico's van het gebruik van AI omvat. Ook normen ten behoeve van het klantbelang, zoals de zorgplicht, normen wat betreft productontwikkeling en -distributie en normen ter voorkoming van overkreditering zijn van toepassing bij het gebruik van AI.

AI zal gevolgen hebben voor de manier waarop de AFM en DNB toezicht houden. Om het gebruik van AI door instellingen te kunnen beoordelen zullen de AFM en DNB hun kennis op dit terrein moeten uitbreiden. Toezichtmethoden en procedures zullen in bepaalde gevallen ontwikkeld of aangepast moeten worden om rekening te houden met nieuwe technieken. Het toezicht zal zich onder meer richten op de besluitvorming over en het risicobeheer rond de inzet van AI-toepassingen, de modaliteiten van die toepassingen en de uitkomsten. Naarmate het gebruik en het belang van AI voor de financiële sector toeneemt, zullen de AFM en DNB hun toezicht op de wijze waarop instellingen de risico's van gebruik van AI beheersen intensiveren. Zo voert DNB dit jaar een thema-onderzoek naar het gebruik van AI door verzekeraars uit.

Nadere guidance, en in bepaalde gevallen specifieke regelgeving, kan nodig zijn om instellingen duidelijkheid te bieden en effectief toezicht mogelijk te maken. Specifieke verwijzingen naar AI-systemen in de huidige regelgeving zijn beperkt. Op enkele terreinen is AI expliciet onderdeel van financiële (sectorale) regelgeving. Voorbeelden zijn onder andere de regels inzake algoritmische handel op kapitaalmarkten, specifieke vereisten aan het gebruik van AI bij kredietverlening aan consumenten en regels voor geautomatiseerd advies. Met het oog op onder meer de rechtszekerheid kan het wenselijk zijn om ook op andere terreinen nader te verduidelijken in

beleidsuitingen, en waar nodig in specifieke regelgeving, wat mag worden verwacht van een instelling. In regelgeving voor interne modellen van banken en verzekeraars worden veel specifieke vereisten gesteld waar ook AI-modellen aan moeten voldoen, wat de toepassing van bepaalde AI-modellen soms lastig maakt. Wanneer deze modellen meerwaarde blijken te hebben kan aanpassing van het huidige normenkader nodig zijn.

De Europese AI-verordening geeft voor kredietwaardigheidstoetsen van natuurlijke personen en premiebepaling en risicobeoordeling voor levens- en ziektekostenverzekeringen nadere vereisten voor het beheerst en verantwoord gebruik van AI. Deze toepassingen van AI zijn aangewezen als hoog risico. Dit betekent dat hiervoor eisen gelden voor onder meer risicomanagement, datakwaliteit, technische documentatie, menselijk toezicht, robuustheid en transparantie voor gebruikers, in aansluiting op bestaande regels voor risicomanagement en governance. Ook moeten instellingen een beoordeling maken van de effecten van deze AI-systemen op grondrechten van personen of groepen. Voor toepassingen die volgens de AI-verordening geen hoog risico met zich brengen wordt in de verordening aanbevolen om zoveel mogelijk ook aan de eisen voor hoog-risicosystemen te voldoen. De AFM en DNB sluiten zich hierbij aan.

Regelgeving voor gebruik van AI in de financiële sector zal gefragmenteerd blijven, hetgeen vraagt om verduidelijking van normenkaders en nauwe samenwerking tussen de toezichthouders. Op Europees niveau dient een goede aansluiting te worden gemaakt tussen de sectorale toezichtkaders en de AI-verordening die hier een aanvulling

op vormt. De Europese toezichthoudende autoriteiten (ESA's) kunnen waar nodig nadere verduidelijking bieden. Om te voorkomen dat het toezicht zelf ook gefragmenteerd wordt, en ook vanwege de opgebouwde sectorale expertise en ervaring, wordt de toezichtrol die uit de AI-verordening voortvloeit in beginsel neergelegd bij de bestaande financieel toezichthouders, voor zover die ziet op financiële dienstverlening. De verwachting is dat hierbij de huidige toezichtverdeling tussen de AFM en DNB zal worden gevolgd.

Ook op nationaal niveau is brede coördinatie van toezicht op AI wenselijk. De toepassingen van AI zijn overal in de publieke en private sector te vinden. Ook buiten de financiële sector zijn al bestaande regelgevende kaders die worden toegepast, zoals AI-toepassingen in de gezondheidszorg. De risico's van AI-systemen kunnen het beste worden beoordeeld in hun specifieke context. Dit vergt sectorale expertise, wat grenzen stelt aan de mate waarin invulling van normen kan worden geüniformeerd. Met inachtneming van die grenzen is voor effectief en efficiënt toezicht samenwerking tussen toezichthouders wenselijk.

Inhoudsopgave

| | |
|---|-----------|
| Hoofdboodschappen | 2 |
| Samenvatting | 3 |
| Inleiding | 8 |
| 1 AI in de financiële sector | 9 |
| 1.1 Wat verstaan we onder AI? | 9 |
| 1.2 Huidige toepassingen in de financiële sector | 11 |
| 1.3 Kansen van AI in de financiële sector | 14 |
| 1.4 Risico's van AI in de financiële sector | 17 |
| 1.5 Toekomstbeelden voor AI in de financiële sector | 21 |
| 2 Toezicht op AI en wetgeving | 24 |
| 2.1 Doelstellingen van het toezicht op AI | 24 |
| 2.2 Wettelijk kader & toezicht op AI | 25 |
| 2.2.1 Toezichtrechtelijk kader algemeen | 26 |
| 2.2.2 Toezichtrechtelijk kader AI specifiek | 27 |
| 2.3 AI-verordening | 29 |
| 2.4 AI in de financiële sector en de bescherming van grondrechten | 34 |

Inleiding

De afgelopen jaren heeft artificiële intelligentie (AI) aanzienlijke ontwikkelingen doorgemaakt.

Met name door de doorbraak van generatieve AI, die nieuwe inhoud kan genereren, zijn veel mensen de directe meerwaarde van AI in hun werk of privé-omgeving gaan inzien. De mogelijkheden van productiviteitsverhoging en innovatieve productontwikkeling leidden ertoe dat innovatieafdelingen van organisaties over de hele wereld gingen kijken naar wat AI ze kan bieden, en talloze jonge bedrijven de markt opkwamen met nieuwe toepassingen. Financiële instellingen zetten al langer systemen en modellen in waarin AI wordt toegepast. AI-gebruik brengt naast kansen ook risico's mee voor financiële instellingen en de maatschappij als geheel. Zo kunnen AI-modellen die inadequate data gebruiken leiden tot een onjuiste inschatting van risico's en kunnen modellen 'biases' bevatten die leiden tot benadeling van klantengroepen. Gebruik van AI kan dus een impact hebben op de financiële soliditeit en integriteit van instellingen en op het belang van de klant. Het bevindt zich daarmee in het domein van het toezicht van de AFM en DNB op de financiële sector. Het toezicht van de AFM en DNB richt zich op een verantwoorde inzet van AI waarbij de voordelen, die AI kan bieden, kunnen worden benut terwijl de risico's adequaat worden beheerst.

De AFM en DNB staan voor vergelijkbare vraagstukken wat betreft het toezicht op AI.

De technologische ontwikkelingen in de financiële sector, en de toekomstige invoering van de Europese AI-verordening zetten AFM en DNB als toezichthouders aan tot nadenken over de eisen die aan de inzet van AI in de financiële sector gesteld worden, en de inrichting van het toezicht daarop. AI verandert snel en wet- en regelgeving verandert mee. De uitkomsten van dit rapport bieden daarom geen pasklare antwoorden op

de weergegeven vraagstukken. Het doel van het rapport is om uitgangs- en aandachtspunten te formuleren voor het vormgeven van het toezicht op AI. AFM en DNB gaan graag de dialoog aan met de sector, hun klanten, autoriteiten en andere stakeholders over kansen en risico's van AI.

De opbouw van dit rapport is als volgt. In hoofdstuk 1 wordt een algemene beschrijving gegeven van wat in het rapport onder AI wordt verstaan en welke vormen AI kan aannemen. Vervolgens wordt er een overzicht gegeven van de huidige AI-toepassingen in de financiële sector en de kansen en risico's die zich hierbij voordoen. Hoofdstuk 2 gaat in op de toezichttaken rondom AI. De huidige toezichtkaders worden geschetst en de toekomstige veranderingen van het toezicht op AI met de komst van de AI-verordening worden hier uitgelegd. Het eigen gebruik van de AFM en DNB van AI-toepassingen (toezicht *met* AI), en de impact die brede veranderingen in de economie als gevolg van AI kunnen hebben op de financiële sector blijven in dit rapport buiten beschouwing.

1 AI in de financiële sector

Er zijn veel verschillende manieren om uit te leggen wat AI inhoudt, maar de AFM en DNB sluiten zich aan bij de definitie die in de AI-verordening wordt gehanteerd (paragraaf 1.1). Financiële instellingen passen AI-modellen op dit moment al op verschillende gebieden toe (paragraaf 1.2), waaronder bij het monitoren van transacties ter voorkoming van witwassen en terrorismefinanciering en bij het efficiënter maken van processen. Ze experimenteren ook veel met AI, wat doet vermoeden dat het gebruik van AI in de komende jaren flink toe zal nemen. Het gebruik van AI in de financiële sector brengt veel kansen met zich, zowel voor klanten als voor financiële instellingen (paragraaf 1.3). Tegelijkertijd leidt het gebruik van AI ook tot nieuwe risico's die een impact kunnen hebben op de manier waarop financiële instellingen hun klanten behandelen en op hun financiële soliditeit (paragraaf 1.4). De toekomstige invloed van AI hangt onder meer af van de ontwikkelingen in innovaties en regelgeving. Verschillende toekomstbeelden hebben ook diverse implicaties voor het toezicht op de financiële sector (paragraaf 1.5).

1.1 Wat verstaan we onder AI?

AI gaat in essentie om computersystemen die taken verrichten waar gewoonlijk menselijke intelligentie voor nodig is. Over de precieze definitie van AI, ook wel aangeduid als kunstmatige intelligentie, bestaat in de wetenschap, politiek en in de praktijk, discussie. AFM en DNB volgen de definitie die is voorgesteld door de OESO¹ (Organisatie voor Economische Samenwerking en Ontwikkeling) en is overgenomen in de AI-verordening van de Europese Unie, omdat deze in belangrijke mate de reikwijdte van beleid en regelgeving zal bepalen.

De definitie² van een AI-systeem in de AI-verordening is: *“a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments”*

Er zijn verschillende vormen van AI, die worden toegelicht in box 1. De vormen verschillen in de mate van autonomie en het lerende element in het AI-systeem.

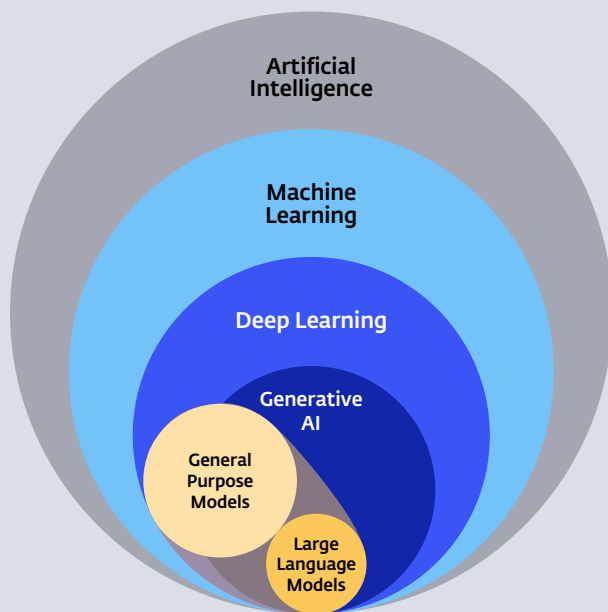
¹ OESO (2023), *Updates to the OECD's definition of an AI system explained*

² De Nederlandse tekst van de AI-verordening is nog niet definitief. Waarschijnlijk zal de vertaling van de definitie langs deze lijnen vorm krijgen: “een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na de uitrol aanpassingsvermogen kan vertonen en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren zoals voorspellingen, inhoud, aanbevelingen of beslissingen die fysieke of virtuele omgevingen kunnen beïnvloeden”.

Box 1 Uitleg over de verschillende soorten AI

AI is een breed begrip dat zich in verschillende vormen kan manifesteren. Feitelijk zijn AI-toepassingen geavanceerde statistische benaderingen van data waarmee voorspellingen, aanbevelingen en nieuwe inhoud gegenereerd kunnen worden. Onder meer wordt onderscheid gemaakt naar machine learning en deep learning. Het recent sterk opgekomen Generative Artificial Intelligence (GenAI), dat in staat is nieuwe inhoud te genereren (zoals ChatGPT en Midjourney), is een stroming binnen deep learning. Zie figuur 1.

Figuur 1 Samenhang tussen verschillende vormen van AI



Machine learning is een vorm van AI waarbij het model leert van bestaande data en deze data kan verbeteren waardoor beslissingen of voorspellingen mogelijk zijn. Het model herkent als het ware patronen en kan deze verder ontwikkelen. Het laten leren door het model kan plaatsvinden op verschillende manieren. Zo kan er gebruik worden gemaakt van gelabelde data (supervised learning), niet-gelabelde data (unsupervised learning) of kan er feedbackgebaseerde data worden gebruikt om het AI-systeem te verbeteren (reinforcement learning).

Deep learning is een variant van machine learning waarbij verschillende *learning* lagen (layers) worden gebruikt om het leerproces van de menselijke hersenen na te bootsen door zogenaamde 'neurale netwerken'. Deze neurale netwerken worden gebruikt om data te analyseren en beslissingen te maken. Neurale netwerken kennen een input layer (de input data voor het netwerk), hidden layer(s) (tussenlaag waar de combinaties en transformaties plaatsvinden) en een output layer (de uitkomst van de berekeningen gebaseerd op de gegeven input). De

complexiteit van deep learning modellen maakt het moeilijk of onmogelijk de uitkomsten te relateren aan specifieke input, waardoor ze ook wel aangeduid worden als black box modellen.

Generatieve AI maakt gebruik van deep learning waarbij er, gebaseerd op aanwijzingen (ook wel prompts genoemd) en bestaande data, nieuwe inhoud wordt gecreëerd. Hierbij valt bijvoorbeeld te denken aan het genereren van nieuwe teksten, afbeeldingen, audio en video.

Large language models (LLMs) zijn generatieve AI-modellen die werken met taal, op basis van zeer grote hoeveelheden data van menselijk taalgebruik (zoals boeken, online forums, e.d.) waarmee deze modellen leren teksten te genereren. De AI-verordening spreekt over AI-modellen voor algemene doeleinden (General Purpose AI), dit zijn AI-modellen met brede mogelijkheden die ingezet kunnen worden voor meerdere doelen in plaats van één specifiek doel. Deze modellen kunnen de basis vormen van vele uiteenlopende concrete toepassingen, zoals bijvoorbeeld LLMs die ook zelf kunnen programmeren.

1.2 Huidige toepassingen in de financiële sector

AI is wereldwijd bezig met een opmars in de financiële sector. Uit een enquête van The Economist die in 2022 is uitgevoerd blijkt dat banken wereldwijd AI voornamelijk gebruiken voor fraudedetectie (58% gebruikt het intensief, en nog eens 32% tot op zekere hoogte).³ Bijna alle banken gaven in deze enquête aan AI te gebruiken, of van plan te zijn het in de komende drie jaar op verschillende gebieden te gaan gebruiken (zie ook figuur 2). Ook in de verzekeringswereld rukt AI snel op. In een Amerikaanse enquête uit 2023 gaf bijvoorbeeld 88% van de autoverzekeraars en 58% van de levensverzekeraars aan AI te gebruiken of dat te willen doen.⁴ Op de kapitaalmarkt is AI terug te vinden in de machine learning modellen die worden gebruikt voor algoritmische handel.⁵

Ook binnen Europa passen financiële instellingen AI volop toe. Uit een enquête van de ECB in 2022 blijkt dat 60% van de 105 grote Europese banken gebruik maakt van AI.⁶ AI wordt volgens deze enquête het meeste toegepast in chatbots, kredietbeoordeling en algoritmische handel.⁷ Ook Europese verzekeraars zetten AI in. Een enquête van EIOPA, de Europese Autoriteit voor verzekeringen en bedrijfspensioenen, uit 2023 laat zien dat AI wordt ingezet door de helft van de respondenten in de schadeverzekeringsbranche, en door 24% in de levensverzekeringsbranche. AI wordt door Europese verzekeraars het meest ingezet voor klantenservice (41%), fraudedetectie (39%) en claimafhandeling (37%). Chatbots zijn de meest gebruikte toepassing. AI wordt ook ingezet om cross-selling aanbevelingen te geven en producten beter te beprijzen.⁸ Het zijn vooral eenvoudige en uitlegbare AI-modellen die

³ The Economist Intelligence Unit (2022), [Banking on a game-changer: AI in financial services](#)

⁴ Insurance Newsnet (2023), [58% of life insurers use artificial intelligence or are interested: NAIC survey](#)

⁵ AFM (2023), [HERs passen machine learning op grote schaal toe in handelsalgoritmes](#)

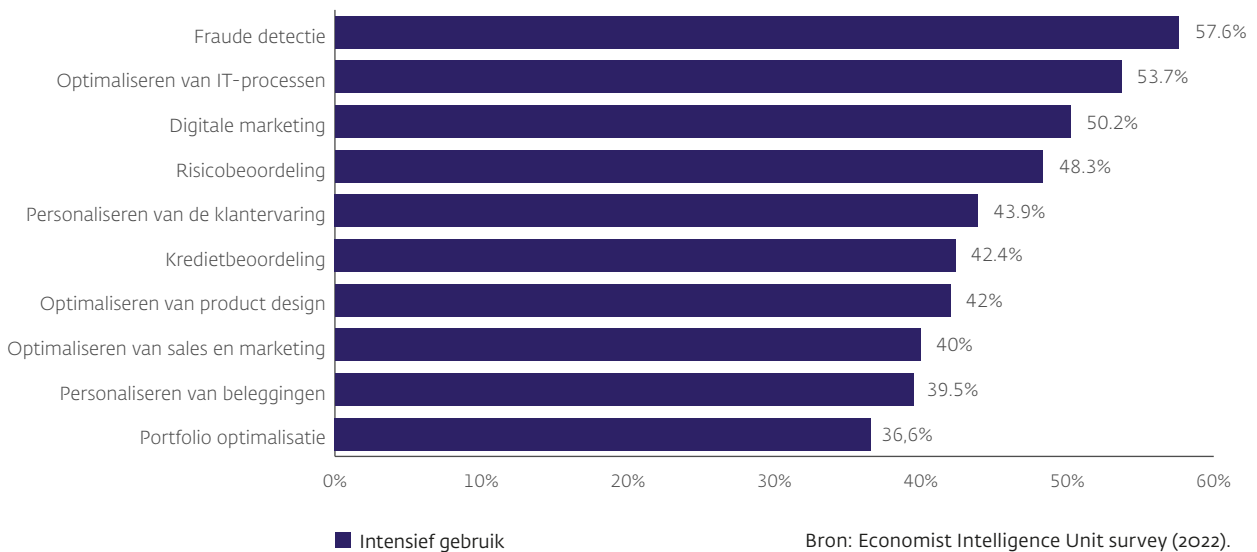
⁶ ECB (2023), [Supervising the future of banking: navigating the digital transformation](#)

⁷ ECB (2023), [Take-aways from the horizontal assessment of the survey on digital transformation and the use of fintech](#)

⁸ EIOPA (2024), Report on the digitalization of the European insurance sector (publicatie van dit rapport volgt later dit jaar)

Figuur 2 Gebieden waarop AI wordt ingezet door banken wereldwijd

Percentage banken dat AI intensief inzet op een bepaald gebied



momenteel worden toegepast, maar verzekeraars verwachten in de nabije toekomst ook meer complexe AI-modellen in te gaan zetten.

Nederlandse financiële instellingen geven aan al langer systemen en modellen te gebruiken waarin AI wordt toegepast; zij experimenteren met geavanceerdere modellen.

Bijna 60% van de significante Nederlandse banken geeft in de enquête van de ECB aan één of meerdere AI-modellen te gebruiken. In interviews stellen de meeste Nederlandse banken dat ze vooral nog de mogelijkheden aan het verkennen zijn, wat doet vermoeden dat het AI-gebruik in de komende jaren zal toenemen.⁹ De Nederlandse banken die op dit moment AI gebruiken, zetten het in bij kredietbeoordelingen, fraudepreventie en de bestrijding van cybercriminaliteit en financieel-economische criminaliteit. Natural language processing (NLP) wordt gebruikt om de

kwaliteit van chatbots te verbeteren. Ook helpt AI medewerkers om efficiënter te werken. Zo worden er automatische samenvattingen gemaakt van telefoongesprekken met klanten. Uit de enquête van EIOPA blijkt dat AI wordt gebruikt door 56% van de Nederlandse respondenten in de schadeverzekeringsbranche en door 50% in de levensverzekeringsbranche. AI wordt volgens deze enquête het vaakst toegepast in chatbots, en bij gerichte online marketing en fraudedetectie. De meeste verzekeraars verwachten dat hun gebruik van AI in de komende drie jaar zal toenemen.

Nederlandse instellingen zien de potentie van generatieve AI, maar veel van hen zijn voorlopig terughoudend met het gebruik ervan. Dit beeld rijst op uit een reeks interviews met Nederlandse instellingen. Het gebruik van de publieke versie van ChatGPT is bij veel financiële instellingen verboden, om te voorkomen dat vertrouwelijke

⁹ Voor dit rapport is gesproken met banken, betaalinstanties, pensioenuitvoerders en verzekeraars.

interne data op externe servers terecht komt en bijvoorbeeld wordt gebruikt om het model verder te trainen. Sommige instellingen gebruiken wel een variant waarbij de ingevoerde data niet op een publieke server wordt geplaatst. Ook zetten sommige instellingen generatieve AI langzaam in voor ondersteunende processen. Onderzoek van de OESO bevestigt dit. Zo experimenteren financiële instellingen volgens de OESO momenteel voornamelijk met offline versies van LLM's die worden gebruikt als ondersteuning bij interne processen en activiteiten, zoals het maken van vertalingen en het opzoeken van informatie.¹⁰ Veel banken en verzekeraars benoemen dat de ontwikkelingen op het gebied van generatieve AI de afgelopen periode sneller zijn gegaan dan verwacht. Tegelijkertijd zijn veel instellingen momenteel terughoudend om het zelf in te zetten doordat de complexiteit met generatieve AI snel toeneemt. Zo gaf een bank aan dat "alles dat met generatieve AI te maken heeft verboden is", terwijl een pensioenuitvoerder stelde dat generatieve AI (nog) niet in scope van de AI ontwikkelteams valt. In de toekomst zullen er naar verwachting wel meerdere nieuwe gebruiksmogelijkheden worden ontdekt die financiële instellingen kunnen inzetten in hun dagelijkse bedrijfsvoering.¹¹ Dan zou de houding van financiële instellingen tegenover generatieve AI kunnen veranderen.

Binnen vermogensbeheer zet een toenemend aantal beheerders AI in bij beleggingsstrategieën, risicomanagement en compliance.

Dit blijkt uit een studie van de Europese Autoriteit voor effecten en markten (ESMA).¹² Tot nu toe hebben slechts een paar vermogensbeheerders

een volledig AI-gebaseerd beleggingsproces ontwikkeld. AI kan beleggingsondernemingen helpen om de handelsuitvoering en post-trade processen te optimaliseren. Daarbij reduceert AI de marktimpact van grote orders en het aantal fouten in het afwikkelingsproces.¹³ De AFM heeft de toepassing van machine learning in handelsalgoritmes bij de vier grote handelaars die voor eigen rekening handelen onderzocht, en constateert dat deze partijen veelvuldig gebruik maken van machine learning modellen. Vooral nog gaat het vooral om supervised learning modellen. Deze modellen voorspellen de prijs van een financieel instrument waarop wordt gehandeld. De complexiteit van deze modellen brengt echter wel risico's met zich mee. Dit maakt dat modellen goed uitlegbaar moeten zijn en vraagt om goede controles.

Sommige instellingen geven aan in andere landen, zoals de Verenigde Staten, al meer AI in te zetten.

Gebrek aan schaal binnen Europa, en daarmee gebrek aan voldoende data om AI-modellen mee te trainen, lijkt daarbij een factor. De strengere privacywetgeving in Europa wordt ook als een verklaring gegeven. Daarnaast lijkt onzekerheid over regelgeving een rol te spelen bij het feit dat instellingen in Europa terughoudender zijn met de inzet van AI. Risicomanagement zou in Europa ook grondiger kunnen zijn dan in de Verenigde Staten. Voorstellen voor AI-toepassingen moeten vaak worden goedgekeurd door het bestuur, ook waar het gaat om toepassingen die alleen beogen niet-essentiële repetitieve taken efficiënter te maken.

¹⁰ OESO (2023), *Generative Artificial Intelligence in Finance*

¹¹ Ibid.

¹² ESMA (2023), *Artificial intelligence in EU securities markets*

¹³ Ibid.

Financiële instellingen stellen dat zij hun AI-modellen het liefst zo eenvoudig mogelijk houden. Instellingen maken een afweging tussen de prestaties van een model en de uitlegbaarheid. Hoe risicovoller de activiteit, hoe eerder de afweging uitvalt richting een eenvoudig model, ook als geavanceerdere modellen tot betere uitkomsten leiden. Uitlegbaarheid is dan immers extra belangrijk. Sommige complexe AI-modellen worden beperkt of niet ingezet vanwege de onzekerheid rondom het uiteindelijke toezicht. Enkele partijen geven aan principieel tegen black box modellen zoals deep learning modellen te zijn, en zeggen deze in de toekomst ook niet te zullen gebruiken. Of financiële instellingen er over een paar jaar nog zo over denken is de vraag. Zo zouden concurrentieoverwegingen ertoe kunnen leiden dat financiële instellingen in de toekomst meer geavanceerde AI gaan inzetten. Als ze hierin te terughoudend zouden zijn, kunnen ze mogelijk de concurrentie niet aan met buitenlandse financiële en niet-financiële partijen.

Instellingen hechten belang aan het behouden van een 'human in the loop'. Dit betekent dat AI-modellen of systemen niet altijd volledig geautomatiseerd zijn. Een medewerker van de financiële instelling beoordeelt de door AI gegenereerde uitkomst voordat deze definitief wordt vastgesteld en wordt gebruikt. De medewerker (en uiteindelijk het management) blijft verantwoordelijk voor de inhoud. Veel instellingen hebben bijvoorbeeld ook een *Chief Technology Officer* (of vergelijkbare functie) in hun bestuur, die verantwoordelijk is voor zaken rondom innovatie en digitalisering.

Om de risico's die met het gebruik van AI gepaard gaan te beperken worden bestaande

kaders voor risicomanagement toegepast, soms aangevuld met specifieke AI-kaders.

Instellingen zijn gewend om te werken met modellen. Ook voor AI-modellen kunnen zij hun bestaande kaders voor modelrisicomanagement toepassen. Sommige instellingen hebben ook specifieke vereisten voor AI-modellen opgesteld. Zo werkt een verzekeraar met 'model score cards', waarin per model een aantal vragen moet worden beantwoord. Denk bijvoorbeeld aan vragen over waartoe het model dient, welke data het model gebruikt en hoe voorkomen wordt dat het model discrimineert.

Financiële instellingen geven aan dat zij groot belang hechten aan het op ethische wijze toepassen van AI, en dat ze niet de grenzen opzoeken van wat er mag en mogelijk is volgens huidige en aanstaande wet- en regelgeving.

Om met ethische dilemma's om te gaan heeft het merendeel van de instellingen geen specifieke interne AI-ethische commissie ingesteld, maar zetten zij bestaande ethische commissies in op AI-gerelateerde onderwerpen. Instellingen uit verschillende sectoren geven aan behoefte te hebben aan gedeelde definities en interpretaties van ethische principes (fairness, explainability, etc.), om meer houvast te hebben in de grijze gebieden. Verzekeraars maken hiervoor al gebruik van het Ethisch Kader van het Verbond van Verzekeraars.¹⁴ Dit kader vraagt verzekeraars een aantal extra checks te doen als zij AI inzetten.

1.3 Kansen van AI in de financiële sector

AI biedt zowel kansen voor klanten als voor financiële instellingen zelf. Hieronder gaan we eerst in op de kansen die AI biedt voor betere en snellere dienstverlening voor de klant, en daarna

¹⁴ Verbond van Verzekeraars (2020), [Ethisch Kader](#)

op de mogelijkheden die AI de financiële sector biedt om hogere inkomsten te genereren en kosten te verlagen.

Betere en snellere dienstverlening

AI kan leiden tot betere dienstverlening aan klanten door meer gepersonaliseerde producten en snellere dienstverlening mogelijk te maken. Door het efficiëntere gebruik van data kunnen producten en diensten worden aangeboden die beter aansluiten bij de individuele behoeften van de klant. AI heeft ook de potentie om in verschillende klantinteracties, zoals het aanvragen van een lening of een verzekering, bepaalde stappen te versnellen, wat leidt tot een snellere dienstverlening.

Hogere inkomsten

AI kan financiële instellingen helpen om nieuwe klantgroepen aan te boren. De inzet van AI kan het bijvoorbeeld mogelijk maken om risicobeoordelingen te maken van klanten met een beperkte of geen kredietgeschiedenis.¹⁵ Als dit ertoe leidt dat producten als consumentenkrediet of verzekeringen op een verantwoorde wijze aangeboden kunnen worden aan mensen die daar eerder niet voor in aanmerking kwamen, verbetert de financiële inclusie. Of en hoe nieuwe klantgroepen kunnen worden aangesneden hangt ervan af of de financiële instellingen over voldoende geschikte data beschikken en of zij in staat zijn modellen effectief te gebruiken. Verder spelen privacy- en ethische overwegingen hierbij een rol. **AI kan helpen om cross-selling mogelijkheden te vergroten.** Bestaande klanten kunnen beter worden bereikt en diensten kunnen gericht worden verkocht. AI kan bijvoorbeeld helpen

voorspellen welk product een klant als volgende zal kopen, of inschatten hoe waarschijnlijk het is dat een klant een bepaald product zal kopen. McKinsey schetst hoe in Azië fintechs, waaronder banken die alleen via internet opereren, gebruikmaken van klantinformatie zoals gedrag, geschiedenis bij een bank en klanttevredenheid om een profiel op te stellen waarmee toekomstige acties en doelen van klanten worden voorspeld.¹⁶ Ook in de verzekeringssector kan AI worden ingezet om verkoopkanalen te optimaliseren en kunnen nieuwe omzetkanalen worden gecreëerd.

AI kan instellingen helpen om hun prijzen te optimaliseren. Verzekeraars kunnen AI bijvoorbeeld inzetten voor gedragsbeprijzing: het bepalen van de hoogte van premies op basis van gedrag van de klant. Dat kan ze helpen de verwachte schade beter in te schatten of te verlagen, en hun winst te maximaliseren. Het kan ook leiden tot een perceptie van een eerlijkere markt, omdat risicomijdende consumenten minder 'meebetalen' aan risicovol gedrag van anderen.¹⁷ Dergelijke beprijzingsstrategieën zijn overigens niet nieuw, maar met behulp van AI kunnen verzekeraars gericht aan gedragsbeprijzing doen door gebruik te maken van nieuwe data – vaak een combinatie van 'nieuwe' interne data (bijvoorbeeld door klik- en afsluitinformatie vanuit de website te analyseren) en externe data.¹⁸ Het personaliseren van prijzen kent overigens niet alleen kansen, maar ook risico's (zoals ondermijning van de solidariteit, discriminatie en uitsluiting, zie sectie 1.4).

¹⁵ OESO (2023), *Generative Artificial Intelligence in Finance*

¹⁶ McKinsey (2022), *Building a winning AI neobank*

¹⁷ AFM (2021), *Personaliseren van prijs en voorwaarden in de verzekeringssector*

¹⁸ AFM-DNB (2019), *Artificiële intelligentie in de verzekeringssector, een verkenning*

Lagere kosten

AI kan processen efficiënter maken, waardoor financiële producten en diensten tegen lagere kosten kunnen worden aangeboden. AI helpt financiële instellingen om processen verder te automatiseren en te versnellen. AI biedt ook kansen om IT-operaties te verbeteren, door verschillende databronnen te integreren en het bouwen en testen van softwaresystemen efficiënter te maken.¹⁹ Handelaren op financiële markten kunnen AI inzetten om fricties te reduceren en de efficiëntie en handelssnelheid te verhogen. Handelsalgoritmes zijn in staat om potentiële handel te identificeren en uit te voeren zonder menselijke tussenkomst. AI helpt om risico's en het orderboek beter te managen om zo de uitvoering te stroomlijnen en efficiënter te werk te gaan.²⁰ Ook bij kredietbeoordelaars, stemadviesbureaus²¹ en andere financiële marktpartijen kan AI worden ingezet om bepaalde activiteiten te ondersteunen en optimaliseren. Zo kan NLP helpen bij het systematisch ontsluiten van publiekelijk beschikbare documenten om analisten te ondersteunen.²²

AI kan financiële instellingen helpen om risico's beter en consistent in te schatten, mits zij beschikken over kwalitatief goede en complete data en modellen die uitlegbaar en controleerbaar zijn. Hoewel gestructureerde data altijd al werd gebruikt, helpen de data als input voor AI-modellen om een zogenaamde

“sentimentanalyse” te doen en om extra inzichten te creëren gebaseerd op patroonherkenning.²³ Waar mensen inconsistent kunnen zijn in hun risico-inschattingen, zijn AI-modellen dat niet. Daarbij moet wel worden opgemerkt dat er ook systematische fouten in AI-modellen kunnen zitten die tot incorrecte resultaten leiden (zie voor meer informatie hierover sectie 1.4).

AI kan ook worden ingezet om environmental, social and governance (ESG) beleggingen te ondersteunen. Beleggers kunnen natural language processing gebruiken om gerichte ESG-evaluaties te maken op basis van corporate social responsibility rapporten van bedrijven.²⁴ Ook binnen DNB is met behulp van NLP onderzoek gedaan naar de awareness en implementatie van duurzaam beleggen bij pensioenfondsen.²⁵

De inzet van AI kan het bestrijden van financieel-economische criminaliteit effectiever en efficiënter maken door betere inschattingen te maken en de administratieve belasting voor instellingen en klanten te verminderen.²⁶ Klantonderzoeken zijn tijdrovend en arbeidsintensief voor financiële instellingen en hun klanten. AI kan uitkomst bieden door de (administratieve) belasting voor de klant te verminderen, terwijl de risico-analyse verder wordt verbeterd. Volgens de Europese Bankautoriteit (EBA) hebben modellen op basis van AI de potentie om verdachte actoren en

19 TU Delft (2023), [The link between artificial intelligence \(AI\) and software engineering](#)

20 OESO (2021), [Artificial Intelligence, Machine Learning and Big Data in Finance Opportunities, Challenges and Implications for Policy Makers](#)

21 Stemadviesbureaus, ook wel proxy advisory firms, zijn organisaties die aan institutionele beleggers, dat wil zeggen beleggingsfondsen, pensioenfondsen en verzekeringsmaatschappijen, adviseert wat te stemmen op een aandeelhoudersvergadering.

22 ESMA (2023), [Artificial Intelligence in EU securities markets](#)

23 OESO (2021), [Artificial Intelligence, Machine Learning and Big Data in Finance](#)

24 ESMA (2023), [Artificial Intelligence in EU securities markets](#)

25 Rob Bauer, Dirk Broeders, Annick van Ool, DNB Working paper 770, [Walk the green talk? A textual analysis of pension funds' disclosures of sustainable investing](#)

26 DNB (2022), [Van herstel naar balans](#)

activiteiten beter te identificeren dan modellen zonder AI.²⁷ Ook transactiemonitoring kan met behulp van AI worden verbeterd. Samenwerking, zoals nu gebeurt tussen vijf Nederlandse banken onder de naam Transactie Monitoring Nederland (TMNL) maakt dit proces nog effectiever. Door het combineren van transactiegegevens van zakelijke klanten van verschillende banken kunnen er verbanden worden gelegd die afzonderlijke banken niet kunnen leggen. Analyse van particuliere transacties is nog niet mogelijk; hiervoor moeten eerst goede (wettelijke) privacy-waarborgen worden opgenomen.

AI kan de verdediging tegen cyberaanvallen verbeteren. AI kan helpen bij het detecteren van zwakke schakels in een (complex) systeem, zodat aanpassingen kunnen worden gedaan waarmee de cyberveiligheid wordt verhoogd. AI-technieken zoals machine learning en deep learning zullen een steeds grotere rol spelen in AI-gedreven cyberveiligheid. Door gebruik te maken van deze technieken kunnen bepaalde cyberveiligheidsprocessen geautomatiseerd worden en 'intelligenter' worden dan de huidige systemen. Hoewel AI de verdediging tegen cyberaanvallen kan verbeteren, kan AI er tegelijkertijd ook voor zorgen dat aanvallen geavanceerder worden (zie volgende sectie).

1.4 Risico's van AI in de financiële sector

AI kan significante voordelen opleveren in de financiële sector, maar deze voordelen gaan gepaard met verschillende risico's die een impact kunnen hebben op financiële

instellingen en hun klanten. Deze risico's vallen daarmee ook binnen de reikwijdte van het toezicht door de AFM en DNB op financiële instellingen en de financiële stabiliteit. Het toezicht op financiële instellingen door de AFM en DNB focust zich op het beheersen van deze risico's binnen instellingen.

Er doen zich verschillende risico's voor in de verschillende fasen van een AI-model. Voor de duiding van deze risico's helpt het om het proces van een AI-model onder te verdelen in drie fasen. Een AI-model heeft data nodig, die worden geselecteerd en in een model worden opgenomen. Dit is de input fase. De geselecteerde data gaan daarna in een AI-model bepaalde operaties en bewerkingen door. Dit is de throughput fase, ook wel 'het systeem' genoemd. Na deze bewerkingen in het systeem komen er resultaten uit het model; dit is de output fase. In elk van deze drie fasen komen specifieke risico's voor. Daarnaast zijn er ook overkoepelende risico's die in alle fasen voorkomen. Zie figuur 3.

Input risico's

AI-modellen worden getraind op grote hoeveelheden data: bij gebruik van data in het risicomanagement of richting de klant zijn de resultaten van AI-modellen zo goed als de data die ze gebruiken. Hierbij geldt het principe '*garbage in, garbage out*': als een AI-model wordt getraind met incorrecte of incomplete data, dan komen er ook incorrecte of incomplete resultaten uit. Een belangrijke voorwaarde voor het gebruiken van efficiënte en effectieve AI-systemen is het op orde brengen van de datakwaliteit en het moderniseren van de oudere *legacy* systemen binnen instellingen. Vanuit het

²⁷ EBA (2020), [Final Report on Big Data and Advanced Analytics](#)

financiële toezicht wordt al enige tijd aandacht gevraagd voor gebreken in de datakwaliteit van instellingen.²⁸ Meer in het bijzonder is bij AI-toepassingen het probleem dat de data biased kunnen zijn; het gebruik van historische data kan ertoe leiden dat bestaande ongelijkheden en historische biases worden overgenomen door het model. Een voorbeeld hiervan is het gebruik van historische sollicitatiedata om toekomstige sollicitanten te selecteren. Als deze data biased is, kan dit leiden tot discriminatie van bepaalde bevolkingsgroepen. Zo bleek een AI-model dat techbedrijf Amazon voor de selectie van sollicitanten gebruikte op basis van de historische data zichzelf te hebben geleerd vrouwen praktisch uit te sluiten.²⁹ De AI-toepassingen moeten dan ook periodiek gevalideerd worden om tijdig bij te kunnen sturen wanneer dit soort risico's zich voordoen.

Bij het gebruik van data ontstaan ook zorgen over privacy en gegevensbescherming.

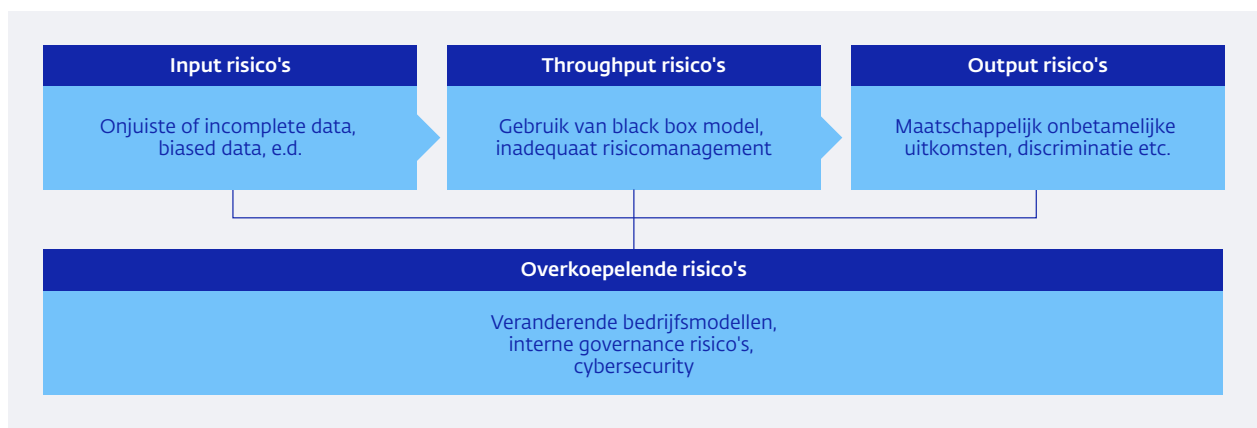
Het gebruik van data moet zorgvuldig gemotiveerd worden en datahouders dragen de verantwoordelijkheid om data goed te

beschermen. Financiële instellingen beschikken in potentie over veel informatie over hun klanten; van de aanschaf van een nieuwe auto of huis tot het afrekenen van de dagelijkse boodschappen in de supermarkt. Deze informatie kan, onder bepaalde voorwaarden, worden gebruikt om een algoritme mee te ontwikkelen. Ook kunnen klantgegevens dienen als input bij een bestaand algoritme. De verwerking van grote hoeveelheden data betekent eveneens dat, wanneer er sprake is van persoonsgegevens, het waarborgen van privacy belangrijk is. Hiervoor is het van belang dat de data governance binnen een instelling voldoet aan de eisen van de Algemene Verordening Gegevensbescherming (AVG). Hierbij kan een spanningsveld ontstaan tussen enerzijds de voorspellende kracht van een algoritme en anderzijds het belang van een verantwoord, beheerst en integer proces – waarbij de privacy van klanten adequaat wordt geborgd.

Throughput risico's

Systematische fouten in het AI-model, ook wel 'model bias' genoemd, kunnen leiden

Figuur 3 De risico's van de AI-toepassingsfasen



²⁸ DNB (2020), *Veranderen voor vertrouwen*; ECB (2020), *ECB report on banks' ICAAP practices*; DNB (2022), *Good Practice: Robuuste pensioenadministratie*
²⁹ BBC News (2018), *Amazon scrapped 'sexist AI' tool*. Dergelijke toepassingen vallen overigens buiten het financieel toezicht.

tot ongewenste of incorrecte resultaten.

Model bias kan optreden als de selectie van trainingsdata niet goed was of de kenmerken van het model niet goed gekozen waren. Hierdoor wordt het algoritme verkeerd getraind. Wanneer hetzelfde model later met nieuwe en betere data wordt gevoed, kunnen dezelfde systematische fouten zich blijven herhalen door de training van het model. Dit kan leiden tot ongewenste of incorrecte resultaten. Om dit risico te ondervangen dienen AI-modellen te worden verbeterd in feedbackloops. Hierbij kan bijvoorbeeld de output worden gemeten aan de hand van gestandaardiseerde maatstaven, en de resultaten worden getest voor geïdentificeerde risico's. Vervolgens zullen de gewenste aanpassingen doorgevoerd moeten worden en het model worden verfijnd. Als dit niet gebeurt, dan lopen modellen het risico om dezelfde fouten eindeloos te herhalen of kunnen fouten zichzelf steeds verder versterken.

De implementatie van AI in de financiële dienstverlening kan leiden tot een probleem met de uitlegbaarheid. Een AI-model bestaat vaak uit een netwerk van meerdere lagen die aan elkaar verbonden zijn en bevat soms wel honderden kenmerken die bijdragen aan het leveren van een resultaat. Het is daardoor niet altijd mogelijk om resultaten van AI-modellen te herleiden tot één kenmerk waarmee aan een klant uitgelegd kan worden waarom een bepaalde beslissing is genomen. Een voorbeeld hiervan is een afgewezen aanvraag voor een krediet of verzekering waarbij de klant recht heeft op motivering. Het kan lastig zijn om de beslissing van het AI-model uit te leggen aan de klant omdat er geen duidelijk datapunt is dat de beslissing verklaart. Het gebruik van dit soort black box modellen kan ten koste gaan van transparantie en uitlegbaarheid. In de wetenschap wordt gewerkt

aan technische oplossingen om de uitlegbaarheid van AI-modellen te vergroten.

Output risico's

Instellingen kunnen AI-toepassingen inzetten die leiden tot maatschappelijk onwenselijke uitkomsten. Zo kan een biased model leiden tot ongewenste financiële exclusie van bepaalde marktsegmenten en kan het gebruik van biased data leiden tot discriminatie van bepaalde bevolkingsgroepen. Dit kan het geval zijn bij AI-modellen die kredietbeoordelingen maken. Als deze modellen zijn getraind met historische gegevens waarin bepaalde groepen minder toegang hadden tot krediet, dan kunnen deze modellen deze bias doorzetten door lagere kredietscores toe te kennen aan klanten uit deze groepen. Dit kan leiden tot financiële exclusie en ongelijke toegang tot financiële diensten.

Het gebruik van AI voor het personaliseren van prijzen kan markttransparantie en solidariteit verstoren. AI-gebaseerde prijsbepaling vereist vaak grote hoeveelheden klantgegevens. Dit kan ertoe leiden dat klanten onbewust 'betalen' met hun gegevens voor gepersonaliseerde prijzen, wat privacy- en ethische vraagstukken oproept. Daarnaast kan het de transparantie van de markt verminderen, waardoor het voor klanten moeilijker wordt om producten te vergelijken. In de verzekeringssector kan het personaliseren van premies het solidariteitsprincipe aantasten. Dit principe, waarbij alle verzekerden bijdragen aan de gemeenschappelijke kosten, kan in het gedrang komen als premies volledig op het individu gebaseerd berekend worden. Dit kan leiden tot situaties waarin sommige klanten aanzienlijk meer betalen dan anderen, waardoor het gevoel van solidariteit wordt ondermijnd en bepaalde groepen potentieel onverzekerbaar worden.³⁰

³⁰ DNB (2023), *Verzekeraars in een veranderende wereld*

Overkoepelende risico's

Een belangrijk overkoepelend risico bij het gebruik van AI is een mogelijk gebrek aan fairness. In maatschappelijke discussies over de risico's van AI-systemen worden vaak eisen genoemd zoals 'eerlijkheid' en 'verantwoord gebruik' van AI. Het gebruik van een AI-model mag bijvoorbeeld niet leiden tot ongewilde ongelijke behandeling van bepaalde individuen of groepen. Een probleem hierbij is dat er (nog) geen maatschappelijke consensus is over wat termen zoals 'eerlijke AI' of 'verantwoorde AI' precies inhouden en hoe instellingen hun AI-toepassingen zouden moeten aanpassen om eraan te voldoen. Instellingen ontwikkelen hier doorgaans wel zelf beleid op, gebaseerd op hun variërende interpretaties, wat leidt tot de behoefte om gezamenlijk uniforme kaders vast te stellen.

Een gebrekkig governance raamwerk binnen instellingen kan leiden tot inefficiënties en gemiste risico's. De interne governance van financiële instellingen speelt een cruciale rol in het mitigeren van de risico's van AI. Een goede strategische organisatie van het risicomanagement is van belang voor een beheerste bedrijfsvoering en zorgt ervoor dat risico's van AI systemen tijdig geïdentificeerd en beheerst kunnen worden. Een gestandaardiseerde documentatie van de gebruikte AI-modellen binnen een financiële instelling, waarmee rekening wordt gehouden met de mogelijke risico's, kan daarbij helpen. Verder is het van belang dat werknemers worden ingezet met voldoende kennis en vaardigheden van AI en van de inhoud van de terreinen waarop AI wordt toegepast.

Daarnaast kleven er risico's aan de inzet van AI, bijvoorbeeld doordat ze tot reputatieschade en onjuiste risico-inschattingen kunnen leiden;

maar ook het niet benutten van het potentieel van AI kent nadelen. Door de inzet van AI-toepassingen veranderen de bedrijfsmodellen en werkprocessen van financiële instellingen. Een risico van AI is dat AI-toepassingen bij een instelling de winstgevendheid kunnen ondermijnen als ze leiden tot reputatieschade of onjuiste voorspellingen en aanbevelingen, bijvoorbeeld doordat kredietrisico's verkeerd worden ingeschat. Daarnaast kan de vervanging van personeel met specifieke expertise door AI-modellen ertoe leiden dat menselijke expertise niet meer beschikbaar is wanneer het AI-model niet naar behoren werkt, een AI-model wordt verboden of wordt vervangen. Maar ook in het geval er géén AI wordt ingezet door een instelling zijn er risico's voor de bedrijfsmodellen van financiële instellingen. Door het innovatiepotentieel van AI niet te benutten kunnen instellingen inboeten aan concurrentievermogen.

Naast de risico's van de implementatie van AI-toepassingen door één instelling, zijn er ook risico's die ontstaan door het gebruik van AI in de gehele sector, waaronder afhankelijkheid van derde partijen. AI-systemen die online uitgevoerd worden kunnen onbedoeld in contact met elkaar komen en door *algoritmische collusie* resulteren in oneerlijke handelspraktijken.³¹ Een ander risico betreft het gebruiken van dezelfde ingekochte AI-software door meerdere financiële instellingen tegelijkertijd. Zij zijn dan allemaal afhankelijk van dezelfde partij. Deze marktconcentratie maakt dat risico's zich op grote schaal kunnen manifesteren. Een voorbeeld hiervan is de massale inkoop van de GPT toepassingen van OpenAI als gevolg van de sterke publiciteit rond ChatGPT eind 2022. Dit betekent dat bepaalde potentiële gebreken of onjuistheden van de Large Language Models zich bij veel instellingen tegelijkertijd kunnen voordoen,

³¹ AFM (2023), [Algoritmische collusie op kapitaalmarkten, verdiepende analyse](#)

wat tot systeemrisico's kan leiden. Daarnaast leidt de afhankelijkheid van een vooralsnog beperkte groep technologiepartijen die AI-technologie kunnen leveren tot risico's. Als de partijen die deze technologie leveren om welke reden dan ook hun diensten niet meer (goed) kunnen leveren, kan niet zomaar worden overgestapt op een andere leverancier en ontstaan problemen bij financiële instellingen die deze technologie gebruiken. Dit geldt des te meer als financiële instellingen AI gebruiken in hun kernprocessen. De omstandigheid dat het gaat om vooral niet-Europese partijen die deze technologie kunnen leveren geeft dit ook een geopolitieke dimensie.

Het gebruik van generatieve AI, waaronder LLM's zoals ChatGPT, kent een aantal additionele risico's, onder andere door verspreiding van onjuistheden en desinformatie. De gegenereerde content kan onjuistheden (ook wel aangeduid als hallucinaties van het model) bevatten en hierdoor leiden tot het verspreiden van verkeerde informatie. Verder kan het gebruik van online en publiek toegankelijke AI-toepassingen ertoe leiden dat vertrouwelijke interne data op externe servers terecht komt; het is dus belangrijk dat de generatieve AI alleen in beveiligde omgevingen wordt gebruikt. Andere vormen van generatieve AI zijn onder andere deep fakes en voice generators waarmee vervalste beelden en geluiden gebruikt kunnen worden om reputatieschade aan te richten bij instellingen. Vanuit een toezichtperspectief is er ook een risico dat de rapporten en formulieren die instellingen leveren aan toezichthouders gegenereerd worden door AI waardoor de juistheid of precisie niet vast te stellen is. Duidelijk moet zijn dat persoonlijke verantwoordelijkheid voor documentatie of andere handelingen van de instelling niet door inzet van AI weggenomen wordt.

Ook cyberrisico's kunnen zich in verschillende fasen van het proces manifesteren. Cyberaanvallen door hackers of penetratie-aanvallen kunnen hele systemen platleggen, gevoelige informatie stelen of de output van het model manipuleren. AI-modellen kunnen ook getraind worden om effectievere cyberaanvallen uit te voeren. De opkomst van generatieve AI heeft nieuwe vormen van criminaliteit en fraude met zich meegebracht. Zo zijn bijvoorbeeld *scammers* AI stemvormers gaan gebruiken (om bijvoorbeeld zichzelf de stem te geven van een vertrouwd persoon) om daarmee wachtwoorden en pincodes van klanten te achterhalen. Werknemers van financiële instellingen moeten over de kennis en vaardigheden bezitten die benodigd zijn om valse berichten en onjuiste informatie te herkennen.

Het stroom- en koelwaterverbruik van AI kan in de toekomst een aanzienlijk deel van het wereldwijde verbruik zijn en een grote kostenpost voor financiële instellingen worden. Dit brengt zowel prudentiële zorgen over de afhankelijkheid aan volatiele en instabiele prijzen van stroom als integriteitszorgen over klimaatbeleid en duurzaamheid van de sector mee.

1.5 Toekomstbeelden voor AI in de financiële sector

Om een beeld te krijgen van de mogelijke richtingen waarin deze technologische innovatie zich kan bewegen, is een aantal scenario's ontwikkeld over de toekomst van AI in de financiële sector. Wat de impact van AI op de financiële sector zal zijn in de komende jaren is nog moeilijk in te schatten. Met een scenarioanalyse kan een concreter beeld worden geschetst van mogelijke veranderingen die te verwachten zijn in de financiële sector en de omstandigheden die kunnen leiden tot die veranderingen. De scenario's

worden gedreven door twee factoren rondom de inzet van AI: enerzijds de mate waarop AI toepassingen gereguleerd zullen worden en hoe die regelgeving vorm krijgt en anderzijds de mate waarop innovatieve financiële diensten met AI ontwikkeld zullen worden. Regelgeving kan AI in goede banen leiden maar ook belemmerend werken. Innovatie kan achterblijven vanwege belemmerende regelgeving maar ook door gebrek aan innovatiekracht in de financiële sector. Innovatie komt vooral door marktkrachten tot stand, maar vertrouwenwekkende regelgeving kan adoptie stimuleren. De resulterende vier scenario's geven mogelijke toekomstbeelden van de financiële sector (figuur 4).

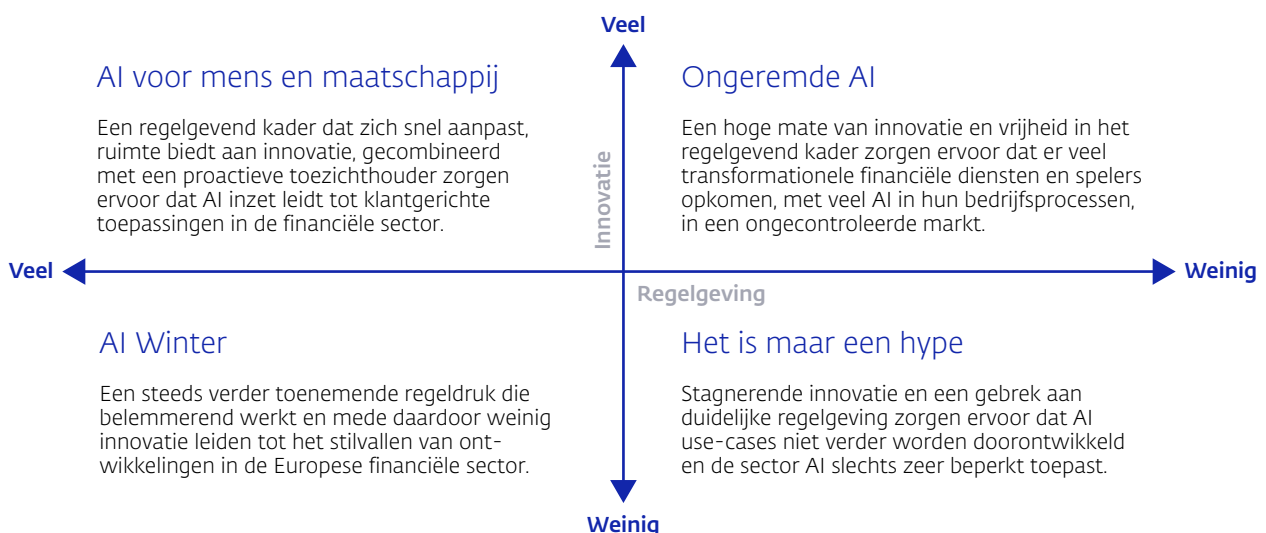
De scenario-analyse onderstreept het belang van evenwicht. Een scenario waarin een evenwicht wordt bereikt tussen groeiende innovatie en adapterende regelgeving lijkt de meest gunstige uitkomst te bieden. Dit scenario 'AI voor mens en maatschappij' zou leiden tot klantgerichte toepassingen en een gezonde groei van de sector. Daarentegen lijken scenario's waarin er een onevenwichtigheid is

tussen regelgeving en innovatie minder gunstige resultaten te bieden. De scenario's waarin veel regelgeving gepaard gaat met weinig innovatie laat zien dat de regelgeving belemmerend kan werken en daarmee kan leiden tot een 'AI winter'. Een ander onevenwichtig scenario is het ontstaan van 'ongeremde AI' wanneer er veel innovatie is met weinig regelgeving om de risico's in een ongecontroleerde markt te beheersen. In beide gevallen van onevenwichtigheid kunnen de beheerste bedrijfsvoering en het klantbelang negatieve gevolgen ervaren.

De scenario-analyse benadrukt daarmee het belang van een proactieve en adaptieve houding van beleidsmakers en toezichthouders.

Regelgeving hoort een kader te creëren voor verantwoordelijkheid en transparantie waar tegelijkertijd ook genoeg ruimte wordt geboden voor snelle ontwikkelingen in AI producten en diensten. Het is een uitdaging die zowel beleidsmakers als innovatieve instellingen samen moeten aangaan, met het oog op het creëren van een toekomst waarin AI ten goede komt aan zowel de financiële sector als de samenleving als geheel.

Figuur 4 Toekomstscenario's AI in de financiële sector



2 Toezicht op AI en wetgeving

Dit hoofdstuk gaat in op het toezicht op het gebruik van AI door financiële instellingen. Voor het toezicht van de AFM en DNB staan financiële soliditeit en integriteit van instellingen, het borgen van het belang van de klant en zuivere verhoudingen tussen marktpartijen voorop. Dit is niet anders wanneer AI wordt ingezet (paragraaf 2.1). Het bestaande wettelijke kader voor het financiële toezicht kan dan ook worden toegepast bij het gebruik van AI-systemen door financiële instellingen (paragraaf 2.2). Belangrijke pijlers van financieel toezicht, zoals adequaat risicomanagement en klantbelang, bieden aangrijpingspunten om het toezicht op AI vorm te geven (2.2.1). Bestaande normen moeten worden geïnterpreteerd in een nieuwe context waarin AI wordt toegepast. Ten behoeve van onder meer de rechtszekerheid kan het wenselijk zijn om in regelgeving en beleidsuitingen nader te verduidelijken wat mag worden verwacht van een instelling, zoals nu al in de eisen voor algoritmische handel op de kapitaalmarkt (2.2.2). De toepassing van AI vraagt ook om een bredere wettelijke bescherming van waarden en grondrechten van de Europese Unie. Dit zien we terug in de totstandkoming van de Europese AI-verordening. Paragraaf 2.3 licht toe hoe de AFM en DNB aankijken tegen de toepasselijkheid van de verordening op de financiële sector. Gegeven de relatie van AI met grote datasets, de ethische vraagstukken en risico's voor grondrechten wordt ook stilgestaan bij de rollen van andere toezichthouders (paragraaf 2.4).

2.1 Doelstellingen van het toezicht op AI

De bestaande doelstellingen van financieel toezicht gelden onverkort voor het toezicht op gebruik van AI door financiële instellingen.

AI-systemen mogen de financiële soliditeit en integriteit van financiële instellingen niet in gevaar brengen. Evenmin mag het gebruik van AI ertoe leiden dat het klantbelang en de zuivere verhoudingen tussen partijen op financiële markten worden aangetast. Innovatieve systemen waarin gebruik wordt gemaakt van AI kunnen ook ondersteunend zijn aan het voldoen aan prudentiële, integriteits- en gedragsnormen.

Bovendien is innovatie van belang voor het concurrentievermogen van financiële instellingen en daarmee voor een gezonde financiële sector. Onder de voorwaarde dat risico's in voldoende mate beheerst worden, zou toezicht voldoende ruimte moeten bieden aan innovatie. Daarnaast is het wenselijk dat duidelijkheid wordt geboden over de toepassing van de regulering en de manier waarop instellingen moeten aantonen dat hun technieken voldoen.³² Ten behoeve van toezichtconvergentie en het gelijke speelveld gebeurt dit bij voorkeur op Europees niveau.

AI zal gevolgen hebben voor de manier

³² Bonnin Roca, J. & O'Sullivan, E. (2022), The role of regulators in mitigating uncertainty within the Valley of Death, Technovation, januari.

waarop de AFM en DNB toezicht houden.

Om het gebruik van AI door instellingen te kunnen beoordelen zullen de AFM en DNB hun kennis op dit terrein moeten uitbreiden. Toezichtmethoden en procedures zullen in bepaalde gevallen moeten worden aangepast of ontwikkeld om rekening te houden met nieuwe technieken. Het toezicht zal zich onder meer richten op de besluitvorming over en het risicobeheer rond de inzet van AI-toepassingen, de modaliteiten van die toepassingen en de uitkomsten. Hierbij kunnen verschillende onderwerpen aan bod komen. Er is aandacht voor HR-, data- en IT-capaciteiten, zodat kan worden bepaald of financiële instellingen voldoende deskundigheid in huis hebben om AI in te zetten. Ook model-governance en model risico management zijn aandachtspunten, evenals het beleid om van AI-uitkomsten af te kunnen wijken: medewerkers (en uiteindelijk senior management) blijven verantwoordelijk voor de uitkomsten en dienen deze wanneer nodig aan te kunnen passen. Verder zal er verhoogde aandacht zijn voor strategische besluitvorming bij verschuivende concurrentieverhoudingen

en (mondiale) concurrentiedruk, omdat dit de ethische toepassing van AI onder druk zou kunnen zetten. Toezichthouders zullen eveneens moeten beoordelen of het regelgevend kader hen genoeg handvatten biedt voor de uitoefening van hun toezicht. Naar verwachting dienen normen te worden verduidelijkt. Daarnaast vraagt de inwerkingtreding van de AI-verordening om het aanwijzen van toezichthouders en een verduidelijking van de interactie tussen het huidige regelgevend kader en de verordening (zie verder 2.3).

2.2 Wettelijk kader & toezicht op AI

Financiële toezichtvereisten kunnen worden gekoppeld aan het proces en de risico's van een AI-model, zoals geïntroduceerd is in hoofdstuk 1.

Figuur 5 geeft de hierboven genoemde verschillende onderdelen van het AI-proces weer, en schetst de relatie met belangrijke onderwerpen van regelgeving. In de sub-paragrafen die volgen worden de toezichtsvereisten voor verschillende fasen van AI-toepassingen binnen het wettelijk kader toegelicht.

Figuur 5 Verschillende fasen van AI-toepassingen en het toezicht



2.2.1 Toezichtrechtelijk kader algemeen

Het toezicht op AI-toepassingen wordt intensiever. Financiële instellingen gebruiken al langer systemen en modellen waarin AI wordt toegepast. Het gebruik van AI maakt deel uit van bestaande processen en dienstverlening waar door de AFM en DNB toezicht op wordt gehouden op grond van geldende wet- en regelgeving. Hoe specifiek dat toezicht is en welke eisen worden gesteld, hangt af van de betreffende toepassing. De aandacht vanuit het toezicht wordt bepaald door het belang van de processen waarop AI wordt ingezet, en de risico's die zich daarbij kunnen voordoen. Daarbij is van belang dat het normenkader voldoende duidelijk is. Naarmate het gebruik en het belang van AI voor de financiële sector toeneemt, zullen de AFM en DNB hun toezicht op de wijze waarop instellingen de risico's van gebruik van AI beheersen intensiveren. Zo voert DNB dit jaar een thema-onderzoek naar gebruik van AI door verzekeraars uit.

Belangrijke pijlers van financieel toezicht zijn bepalingen gericht op integriteit en risicobeheersing. Zo dienen instellingen te voldoen aan de vereisten van een integere en beheerste bedrijfsvoering, die overkoepelend op AI-systemen worden toegepast. Beheersing van risico's die voortvloeien uit gebreken in de data (input) en in de gebruikte modellen (throughput) vragen bij de inzet van AI bijzondere aandacht. Ook moeten instellingen, uit hoofde van een integere en beheerste bedrijfsvoering, tegengaan dat wetsovertredingen worden begaan of maatschappelijke normen worden overschreden, waardoor het vertrouwen in de financiële onderneming of in de financiële markten kan worden geschaad. Hieruit vloeit voort dat instellingen ervoor moeten zorgen dat gebruik van hun AI-systemen er niet toe

leidt dat een instelling discrimineert of in strijd handelt met andere grondrechten of wezenlijke maatschappelijke normen (output). Instellingen moeten hun bedrijfsprocessen en bedrijfsrisico's beheersen, alsmede hun financiële risico's en andere risico's die de soliditeit van de instelling kunnen aantasten. Met andere woorden: financiële ondernemingen dienen een adequaat risicomangement in te richten. Ook de systemen die financiële instellingen gebruiken om risico's op witwassen en terrorismefinanciering in te schatten, en die in toenemende mate gebruik maken van AI-technieken, moeten aan de eisen van een integere en beheerste bedrijfsvoering voldoen.

Hoe het risicomangement wordt ingericht is afhankelijk van de aard van de betreffende onderneming, en wordt vaak nader ingevuld door specifieke regelgeving per (sub)sector. Deze is veelal gebaseerd op Europese regelgeving en geïmplementeerd in de Wft. Voorbeelden zijn de Europese richtlijn en verordening kapitaalvereisten (CRD en CRR). Hierin wordt een aantal risico's waaraan banken zijn blootgesteld benoemd, zoals kredietrisico en marktrisico, en worden eisen gesteld aan de beheersing van deze risico's en aan de organisatorische inbedding van de risicobeheersfunctie. Soortgelijke bepalingen zijn voor verzekeraars opgenomen in de Europese richtlijn voor verzekeraars (Solvency II). Een ander voorbeeld zijn de regels voor beleggingsondernemingen (MiFID2) met onder andere aandacht voor specifieke marktrisico's om het ordentelijk verloop van de handel te borgen. In het algemeen kan gesteld worden dat wanneer AI wordt gebruikt, de risico's die daarmee gepaard gaan binnen ditzelfde kader adequaat dienen te worden beheerst door de instelling. Wat precies van instellingen wordt verwacht dient waar nodig nader te worden uitgewerkt op EU-niveau.

Naast normen voor de bedrijfsvoering, gelden normen ten behoeve van het klantbelang.

Hierbij kan gedacht worden aan de zorgplicht, normen wat betreft productontwikkeling en distributie, en normen ter voorkoming van overkreditering. Ongeacht of een product of dienst tot stand is gekomen met behulp van AI, dient het betreffende klantbelang voldoende te zijn meegewogen in het advies, dient de klant niet te zijn overgekrediteerd en dient voorafgaand aan de distributie van een product op een passende wijze de doelgroep te zijn gedefinieerd. Onder meer is van belang dat AI-toepassingen consumenten stimuleren (bewust of onbewust) beslissingen te nemen die hun financiële welzijn ten goede komen.³³ Ook mag dynamische prijsvorming niet ten kosten gaan van markttransparantie.

Van financiële instellingen mag worden verwacht dat zij AI verantwoord inzetten.

Naarmate het gebruik van AI belangrijker wordt in de besluitvorming van een financiële instelling, en de potentiële gevolgen hiervan voor de instelling en haar klanten groter worden, komt de lat voor een verantwoorde en uitlegbare inzet van AI hoger te liggen. AFM en DNB zullen hier in het toezicht op financiële instellingen ook nadrukkelijk naar kijken. Reeds in eerdere publicaties hebben de AFM en DNB aangegeven welke onderwerpen relevant zijn, zie bijvoorbeeld het gezamenlijke rapport 'AI in de verzekeringssector'.³⁴ De komende AI-verordening (zie 2.3) biedt een kader met normen voor het verantwoord gebruik van AI.

2.2.2 Toezichtrechtelijk kader AI specifiek

Specifieke verwijzingen naar algoritmes, AI of AI-systemen in financiële wet- en regelgeving zijn beperkt. Op enkele terreinen is AI expliciet onderdeel van financiële (sectorale) wet- en regelgeving. Voorbeelden zijn: regels inzake algoritmische handel op kapitaalmarkten (MiFID2/ MiFIR), specifieke vereisten aan het gebruik van AI bij kredietverlening (opgenomen in de vernieuwde Consumer Credit Directive, CCD2) en regels voor geautomatiseerd advies. Deze op AI gerichte normen laten zien dat afhankelijk van het type toepassing specifieke vereisten kunnen worden gesteld. Voor het regelgevende kader voor interne modellen van banken (onderdeel van de CRR) en in zekere mate ook dat voor verzekeraars (onderdeel van Solvency II) geldt dat daarin veel specifieke vereisten worden gesteld, waar ook AI-modellen aan moeten voldoen. Het ligt in de lijn der verwachtingen dat in de toekomst op meer terreinen specifieke regels worden vastgesteld of aangepast.

Algoritmische handel

De regels inzake algoritmische handel op de kapitaalmarkt³⁵ stellen eisen aan waarborgen bij het gebruik van AI en verbieden dat AI wordt ingezet op een wijze die bijdraagt of kan leiden tot een onordelijke markt. De vereisten zien onder andere op de weerbaarheid van handelssystemen, afdoende handelscapaciteit, inclusief gepaste handelsdrempels en -limieten. De vereisten worden nader ingevuld door reguleringsstandaarden die zijn vastgesteld op EU-niveau. Een greep uit die standaarden leert

³³ AFM & DNB (2019), Aandachtspunten AFM en DNB bij artificiële intelligentie in de verzekeringssector, p.28

³⁴ AFM & DNB (2019), Aandachtspunten AFM en DNB bij artificiële intelligentie in de verzekeringssector. Zie ook DNB komt met richtlijnen voor gebruik kunstmatige intelligentie en DNB (2019), General principles for the use of Artificial Intelligence in the financial sector.

³⁵ Zie in dit kader onder andere artikelen: 17, 18, 26, 27, 45, 47 en 48 MiFID2.

dat systemen en handelsstrategieën uitvoerig getest moeten zijn, personeel voldoende begrip moet hebben van de systemen, er pre- en post handelsnormen gelden, evenals realtime monitoring verplichtingen. Dergelijke vereisten zijn gerelateerd aan het verantwoord gebruik van het AI-systeem (zie figuur 5).

Consumentenkrediet

De vernieuwde Consumer Credit Directive 2 (CCD2)³⁶ bevat handvatten voor het reguleren van het gebruik van AI inzake de kredietwaardigheidstoets. De vereisten zien voornamelijk op de data (input, zie figuur 5) die wel of niet gebruikt mogen worden en transparantie aangaande de geautomatiseerde dienstverlening (output). Wanneer de prijs van het aangeboden product of dienst tot stand is gekomen door middel van AI (geautomatiseerde processen, inclusief profilering) dient dit te worden medegedeeld aan de consument. De consument kan zo bij de aankoopbeslissing rekening houden met de potentiële risico's.³⁷ Daarnaast krijgen kredietverstrekkers en bemiddelaars de verplichting om consumenten op een duidelijke en begrijpelijke manier te informeren wanneer een gepersonaliseerd aanbod wordt gedaan dat tot stand is gekomen op basis van de geautomatiseerde verwerking van persoonlijke gegevens (transparantieplichting; processen en waarborgen). Daarnaast bevat artikel 18 CCD2 beperkingen met betrekking tot het type persoonsgegevens dat mag worden gebruikt (ongeacht of sprake is van AI) en het recht op menselijke tussenkomst als de kredietwaardigheidstoets het gebruik van geautomatiseerde verwerking van persoonsgegevens omvat.

Geautomatiseerd advies

Ook de regels over geautomatiseerd advies³⁸ zien op de waarborgen die moeten gelden binnen de omgeving waarin een systeem een geautomatiseerd advies genereert. Er wordt verduidelijkt hoe adviesregels – oorspronkelijk geschreven voor fysieke situaties met adviseur en klant aan dezelfde tafel – moeten worden geïnterpreteerd in een geautomatiseerde omgeving. De vereisten omvatten in termen van figuur 5 onder andere toezicht op het systeem, het integreren van vakbekwaamheidseisen in het systeem en output testen door middel van scenarioanalyses. Onder andere wordt bepaald dat personen die worden aangewezen als verantwoordelijke voor de geautomatiseerde adviezen, ook beschikken over de benodigde vakbekwaamheid om zelf te adviseren over het betreffende financiële product. Het geautomatiseerde systeem moet worden getest alvorens het in gebruik wordt genomen. Ook moet de financiële dienstverlener periodiek controleren of de geautomatiseerde adviezen passen bij en aansluiten op de door de consument verstrekte informatie. Tot slot wordt bepaald welke acties moeten worden ondernomen indien blijkt dat de geautomatiseerde adviezen niet voldoen aan de wettelijke regels rondom advies.

Interne modellen banken en verzekeraars

Banken kunnen voor het bepalen van hun kapitaalvereisten onder voorwaarden gebruikmaken van de zogenoemde 'interneringbenadering' (IRB), onder meer voor het bepalen van het kredietrisico. Met behulp van IRB-modellen kunnen banken hun kredietrisico schatten en mogelijk op een lagere

³⁶ EUR-Lex - 32023L2225 - EN - EUR-Lex (europa.eu). Op 30 oktober 2023 is de definitieve tekst van de CCD2 gepubliceerd. Uiterlijk 20 november 2025 moet de richtlijn zijn omgezet in nationale regelgeving en vanaf 20 november 2026 worden de regels toegepast.

³⁷ Zie ook overweging 46 van de CCD2.

³⁸ Artikel 32da Besluit gedragstoezicht financiële ondernemingen. Dit artikel is opgenomen in het Wijzigingsbesluit Financiële Markten 2023 en wordt van

kapitaaleis uitkomen dan zou gelden bij het gebruik van de standaardmethode. AI, en dan met name machine learning, kan mogelijk een nauwkeurigere schatting geven van dit risico dan de traditionele modellen. IRB-modellen moeten aan een uitgebreide set specifieke voorschriften voldoen. Ook instellingen die AI gebruiken moeten binnen deze voorschriften blijven. Deze voorschriften zijn weliswaar niet geschreven met het oog op de toepassing van AI, maar raken wel de verschillende fasen van AI-toepassingen uit figuur 5. Zo is de eis dat het indelen van instellingen in vergelijkbare risicoklassen consequent en te begrijpen voor derden moet zijn (art. 171 CRR). Met AI kan dit in principe erg verfijnd gedaan worden, maar het kan lastig zijn om de consistentie aan te tonen en de uitlegbaarheid aan derden te waarborgen. Het is daardoor niet zonder meer mogelijk AI toe te passen. Een ander voorbeeld is dat bij het maken van ramingen met modellen, de ramingen ‘aannemelijk en intuïtief’ moeten zijn (art. 179 CRR). Het gebruik van AI kan soms leiden tot uitkomsten die niet direct intuïtief zijn, maar toch een betere inschatting blijken te geven van de blootstelling aan kredietrisico dan een traditionele aanpak.³⁹ Er zal dus verduidelijking of mogelijk aanpassing van het kader nodig zijn om bepaalde vormen van AI toe te passen.

Voor het gebruik van interne modellen door verzekeraars gelden vergelijkbare (zij het iets minder gedetailleerde) bepalingen op grond van Solvency II. Met name op de punten van uitlegbaarheid en modelvalidatie kan het gebruik van geavanceerde methoden van machine learning knelpunten opleveren. Zo moeten

gevoeligheden voor wijziging in onderliggende aannames kunnen worden uitgelegd, en moet elke onafhankelijke derde deskundige in staat zijn om de opzet en operationele bijzonderheden van het interne model te begrijpen.

2.3 AI-verordening

De AI-verordening is een cross-sectorale Europese verordening gericht op de bescherming van belangrijke waarden en grondrechten van de EU, zoals veiligheid en non-discriminatie. De toepasselijkheid van de verordening is gebaseerd op de definitie van AI en het onderscheid tussen “onaanvaardbare risico’s”, “hoge risico’s” en “lage of minimale risico’s”. Daarnaast komen er regels voor AI-modellen voor algemene doeleinden. Figuur 6 geeft dit onderscheid in risico-categorieën grafisch weer. Over de AI-verordening is in december 2023 een akkoord bereikt. Naar verwachting wordt de verordening voor de zomer van 2024 gepubliceerd, waarna voor de meeste bepalingen een implementatieperiode van twee jaar volgt.

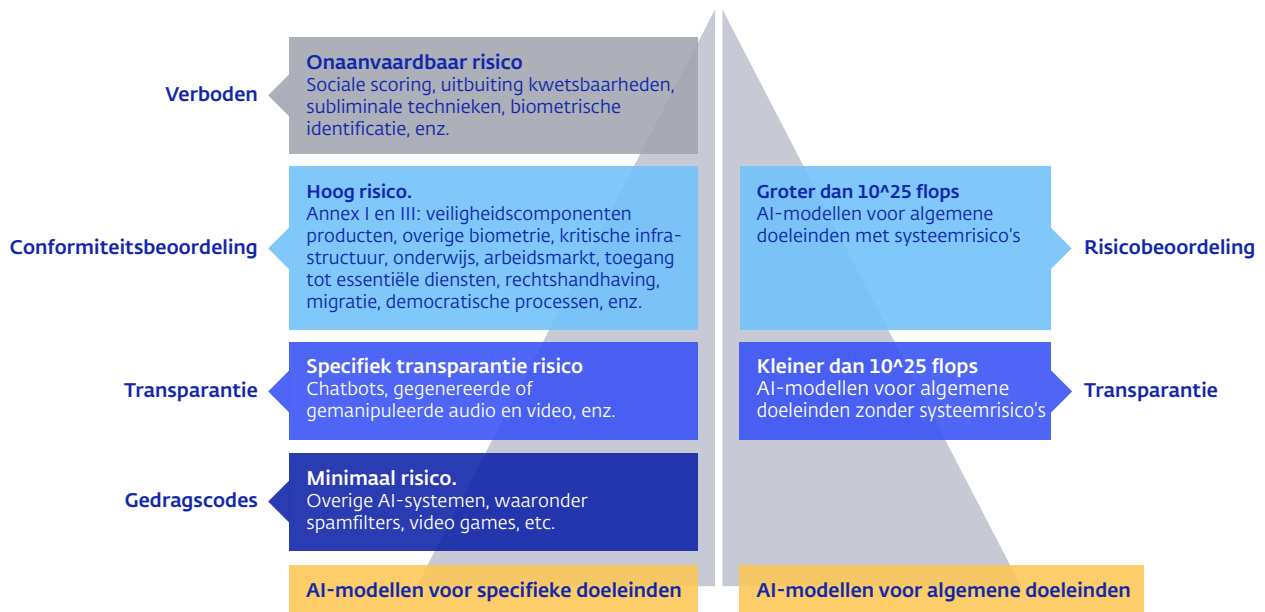
AI-toepassingen die bedoeld zijn om te worden gebruikt voor kredietwaardigheidstoetsen van natuurlijke personen en voor premiebepaling en risicobeoordeling voor levens- en ziektekostenverzekeringen, zijn in de AI-verordening aangewezen als ‘hoog risico’.

Deze toepassingen moeten voldoen aan eisen ten aanzien van onder meer risicomanagement, datakwaliteit, technische documentatie, menselijk toezicht, robuustheid en transparantie voor gebruikers. Ook moeten instellingen een beoordeling maken van de effecten van het AI-

kracht per 1 juli 2024.

39 Zie uitgebreider EBA (2021), [Discussion paper on machine learning for IRB models](#)

Figuur 6 Risicogebaseerde vereisten in de AI-verordening



systeem op de grondrechten van personen of groepen. Voor AI-systemen die niet als hoog risico zijn gekwalificeerd, maar waar wel mensen mee in aanraking komen, zoals chatbots, of die inhoud genereren (generatieve AI) gelden eveneens transparantieverplichtingen, hetgeen ook de financiële sector zal aangaan.

Voor toepassingen die geen hoog risico met zich brengen zullen de Europese Commissie en de lidstaten stimuleren dat gedragscodes worden ontwikkeld om zoveel mogelijk ook aan de vereisten voor hoogrisicosystemen te voldoen. Deze gedragscodes kunnen door de instellingen zelf of brancheorganisaties worden ontwikkeld. AFM en DNB sluiten zich hierbij aan, in lijn met

eerdere uitingen.⁴⁰ De verordening verwijst hierbij naar de Europese 'Ethische richtsnoeren voor betrouwbare KI'.⁴¹ In Nederland heeft het Verbond van Verzekeraars hier zijn 'Ethisch Kader' op gebaseerd. Hoewel kaders vaak op belangrijke punten vergelijkbaar zijn kan het nuttig zijn als de ESA's hier een coördinerende rol gaan spelen om fragmentatie te voorkomen. EIOPA heeft al een rapport gepubliceerd over AI governance principes, eveneens voortbouwend op de genoemde Europese Ethische richtsnoeren.⁴²

De AI-verordening is een aanvulling op de bestaande (sectorale) wetgeving, en moet in haar toepassing zoveel mogelijk aansluiten bij de financiële sector wetgeving.⁴³ De vereisten

⁴⁰Zie bijvoorbeeld: DNB komt met richtlijnen voor gebruik kunstmatige intelligentie

⁴¹ Deskundigengroep op hoog niveau inzake kunstmatige intelligentie (2019), *Ethische richtsnoeren voor betrouwbare KI*

⁴² EIOPA (2021), *AI Governance principles*

⁴³ Dit blijkt uit overweging 158 van de AI-verordening.

zijn afgeleid van bestaande productregelgeving, die voornamelijk betrekking heeft op tastbare producten ondergebracht in het zogenoemde 'Nieuwe Wetgevingskader'. De vereisten hebben betrekking op AI-systemen als product en leggen verplichtingen op aan aanbieders en exploitanten van die systemen. Kort gezegd kunnen AI-systemen met hoog risico slechts op de markt worden gebracht als a priori een conformiteitstoets heeft plaatsgevonden die aantoont dat aan bovengenoemde vereisten wordt voldaan. De aanbieder van het systeem dient hiervoor zorg te dragen. Ook voor de exploitant⁴⁴ van een AI-systeem met hoog risico gelden verplichtingen, onder meer ten aanzien van het gebruik van een systeem overeenkomstig de voorschriften en de competenties van degenen die toezien op het systeem, voor zover ze daar controle over hebben. Financiële instellingen worden geacht hier al deels aan te voldoen als ze zich houden aan de bestaande regels in het financieel toezicht op het gebied van interne governance. Exploitanten moeten ook de genoemde effectbeoordeling op grondrechten maken. Op Europees niveau dient een goede aansluiting te worden gemaakt tussen de sectorale toezichtkaders en de AI-verordening. Hier is een coördinerende taak weggelegd voor de Europese toezichthoudende autoriteiten, die waar nodig nadere guidance kunnen bieden.

Op basis van financiële sector wetgeving blijft een financiële instelling bij uitbesteding van activiteiten zelf verantwoordelijk en dat verandert niet. Onder de AI-verordening is sprake van een gelaagde benadering waarbij afhankelijk van het type normadressant (zoals aanbieder of exploitant) een andere subset aan verplichtingen geldt. Onduidelijkheid over waar

precies verantwoordelijkheden liggen zal moeten worden weggenomen in de samenloop van financiële sector wetgeving en de AI-verordening. Financiële instellingen die zelf de rol van aanbieder of exploitant van hoog-risicosystemen vervullen zullen aanvullend aan de financiële regelgeving aan de vereisten van de AI-verordening moeten voldoen. Waar sprake is van uitbesteding door een financiële instelling gelden voor de instelling de uitbestedingsregels, en voor de aanbieder aan wie is uitbesteed de betreffende regels van de AI-verordening. Op deze manier ontstaat dus meer grip op aanbieders van AI-systemen voor de financiële sector. Deze hoeven onder de bestaande financiële regelgeving alleen indirect vanuit de aanbestedingsrelatie aan toezichteisen te voldoen. De beheersing van risico's vergt bijzondere aandacht als een beperkt aantal aanbieders van AI-systemen dominant worden in de markt.

De AI-verordening reguleert naast AI-systemen voor specifieke toepassingen ook AI-modellen voor algemene doeleinden.

AI-modellen voor algemene doeleinden, waaronder grote generatieve AI-modellen, kunnen voor allerlei verschillende taken worden ingezet. Voor deze modellen gelden specifieke eisen, onder meer ten aanzien van het bekend maken van informatie zodat aanbieders die deze modellen gebruiken om eigen toepassingen te ontwikkelen er voor kunnen zorgen dat die toepassingen veilig zijn en in overeenstemming met de vereisten van de verordening. Ook moeten aanbieders zorgen dat ze het auteursrecht eerbiedigen. Voor modellen die systeemrisico's met zich brengen, wat afgemeten wordt aan de rekenkracht waarmee ze zijn getraind, gelden additionele eisen op het gebied van risicomanagement en cyberbeveiliging. Hiervoor geldt een drempel van 10^{25} 'flops' ('floating

⁴⁴ Een gebruiker van een AI-systeem voor professionele doeleinden (in het Engels aangeduid als 'deployer').

point operations per second', een maatstaf voor computerrekenkracht). Alleen de meest geavanceerde AI-modellen komen hier boven.

De AI-verordening beoogt naast het beheersen van risico's en het beschermen van grondrechten ook innovatie te bevorderen.

De verordening biedt een aantal maatregelen om innovatie te stimuleren, waaronder de opzet van testomgevingen voor regelgeving. Hierin kunnen innoverende partijen advies krijgen van toezichthouders over de interpretatie van de AI-verordening ter voorbereiding op de marktintroductie. Ook komen er buiten de testomgevingen mogelijkheden om toepassingen te testen onder reële omstandigheden. Meer algemeen zijn de regels van de verordening risicogebaseerd en laten ze veel ruimte voor toepassingen die geen hoge risico's meebrengen. En waar dat wel zo is, kunnen die AI-toepassingen met voldoende waarborgen wel worden aangeboden. Regels en toezicht moeten ook het vertrouwen bevorderen. Dit is zeker niet overbodig, zoals blijkt uit een enquête uit 2023 onder Nederlandse huishoudens (zie box 2). Slechts een kwart van de respondenten was positief over gebruik van AI door financiële instellingen. Bijna tweederde van de respondenten is positiever als ze weten dat er toezicht wordt gehouden op het gebruik van AI.

Er worden coördinatiemechanismen opgezet voor het toezicht. Lidstaten zullen nationale autoriteiten aanwijzen die het toezicht zullen uitvoeren. Voor AI-systemen met hoog risico die worden aangeboden of gebruikt door financiële instellingen ten behoeve van financiële diensten wordt in de verordening de toezichtrol toebedeeld aan de bestaande financieel toezichthouders; voor Nederland betreft dat de AFM en DNB (lidstaten hebben de mogelijkheid hiervan af te wijken). De nationale toezichthouders zullen

samenwerken in een op te richten AI-comité. Ook wordt een Europees AI-bureau opgericht binnen de Europese Commissie, dat een coördinerende rol vervult en toezicht houdt op AI-modellen voor algemene doeleinden. De Europese instanties worden bijgestaan door een adviesforum van organisaties van belanghebbenden en een wetenschappelijk panel van onafhankelijke deskundigen. AI-toepassingen zijn niet altijd aan landsgrenzen gebonden, bijvoorbeeld algoritmische handel, waardoor ook mondiale afstemming nodig is. De ECB is vanuit haar prudentiële toezichtmandaat verantwoordelijk voor AI-toepassingen van systeemrelevante banken, maar neemt geen toezichtrol op basis van de AI-verordening op zich.⁴⁵ Er zal voor het toezicht op AI-gebruik van deze banken afstemming nodig zijn tussen ECB en de nationale financieel toezichthouders.

Ook op nationaal niveau is brede coördinatie van het toezicht op AI wenselijk. De toepassingen van AI zijn overal in de publieke en private sector te vinden. Ook buiten de financiële sector zijn al bestaande regelgevende kaders die worden toegepast, zoals AI-toepassingen in de gezondheidszorg. De risico's van AI-systemen kunnen het beste worden beoordeeld in hun specifieke context. Beoordeling vergt sectorale expertise, wat grenzen stelt aan de mate waarin invulling van normen kan worden geüniformeerd. Met inachtneming van die grenzen is voor effectief en efficiënt toezicht samenwerking tussen toezichthouders wenselijk. De Autoriteit Persoonsgegevens (AP) met haar Directie Coördinatie Algoritmes in haar rol als algoritmetoezichthouder en de Rijksinspectie Digitale Infrastructuur (RDI) spelen hierin een coördinerende rol. Ten aanzien van toenemende cyberrisico's is goede samenwerking met het Nationaal Cyber Security Centrum (NCSC) nodig.

⁴⁵ Zie voor de achtergrond hiervan het advies van de ECB: [Publications Office \(europa.eu\)](https://www.ecb.europa.eu/publications/office/2023/01/ai-impact-on-the-financial-sector-and-supervision).

Box 2 Meningen over AI-gebruik in de financiële sector lopen uiteen

In 2023 zijn door DNB aan een panel van Nederlandse huishoudens vragen gesteld over het beeld dat ze hebben over AI. De vragenlijst is voorgelegd aan 3.054 panelleden, 2.256 respondenten vulden de vragen volledig in (responspercentage 73,9%).

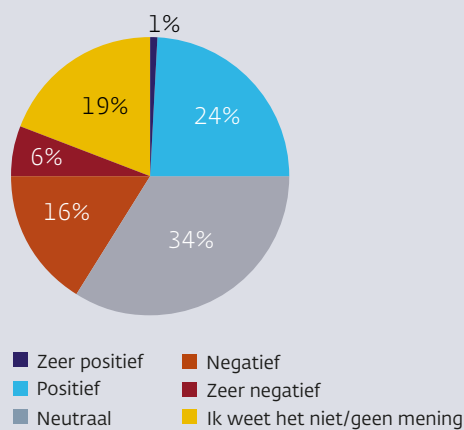
Uit de enquêteresultaten blijkt dat de meningen over AI erg uiteen lopen. Een derde is in het algemeen positief over AI, 15% heeft een negatief beeld. De helft van de respondenten heeft nog niet echt een mening over AI of is neutraal, wellicht mede omdat ze nog niet zo'n goed beeld hebben van wat AI nu precies inhoudt. Slechts vier op de tien mensen geven namelijk aan een (redelijk) goed beeld te hebben van wat AI is.

Een kleine meerderheid van de respondenten heeft over het gebruik van AI door financiële instellingen nog geen mening of is neutraal. Een kwart is positief. Opvallend is dat de groep respondenten die negatief is over AI-gebruik door financiële instellingen (22%) groter is dan de groep die negatief is over AI in het algemeen (15%). Slechts een paar respondenten hebben een goed beeld van in hoeverre hun bank (6%), verzekeraar (5%) of pensioenfonds (4%) AI gebruikt.

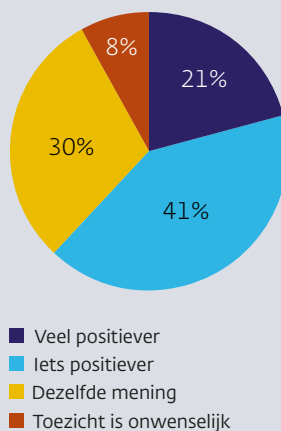
Over de vraag of de voordelen of de risico's van AI-gebruik door financiële instellingen groter zijn lopen de meningen uiteen: 28% denkt dat de voordelen groter zijn, terwijl 29% de risico's groter vindt. De rest is neutraal of weet het niet. Een meerderheid (62%) van de mensen is positiever over AI-gebruik door financiële instellingen als er toezicht wordt gehouden.

Figuur 7 Wat vindt u van het gebruik van AI door financiële instellingen?

Wat vindt u van het gebruik van AI door financiële instellingen?



Wat vindt u ervan als uw bank, verzekeraar of pensioenfonds AI gebruikt als u weet dat er toezicht wordt gehouden op het gebruik van AI?



Noot: De enquête is afgenomen tussen 13 en 28 maart 2023.

2.4 AI in de financiële sector en de bescherming van grondrechten

Bescherming van grondrechten vergt bij toepassing van AI bijzondere aandacht. Bij de toepassing van AI bestaat het risico dat grondrechten van burgers worden geschonden, zoals in paragraaf 1.4 aan de orde kwam. Bij het toepassen van AI door financiële instellingen zijn voor wat betreft grondrechten in het bijzonder het discriminatieverbod en de bescherming van persoonsgegevens van belang. Dit vergt samenwerking tussen financiële toezichthouders en andere toezichthouders. Het overkoepelende EU-grondrechtenkader is ook van toepassing op het gebruik van AI in de EU. Dit kader bestaat uit het Handvest van de grondrechten van de Europese Unie (hierna het “Handvest”) en het Europees Verdrag voor de rechten van de mens (EVRM). In secundaire EU-wetgeving, zoals de AVG en non-discriminatiewetgeving van de EU, is uitgewerkt hoe bepaalde grondrechten dienen te worden beschermd. Ook de AI-verordening past in deze context, en bevat een nadere uitwerking van regels ter bescherming van grondrechten en Uniewaarden. Op nationaal niveau zijn de grondrechten opgenomen in de Grondwet, en is het discriminatieverbod van artikel 1 van de Grondwet uitgewerkt in de Algemene Wet gelijke behandeling. In het kader van de uitvoering van de AI-verordening is het van belang dat door de Nederlandse wetgever wordt verduidelijkt hoe de bepalingen uit de verordening die zien op de bescherming van EU grondrechten en waarden zich verhouden tot de bestaande mandaten van toezichthouders (waaronder de AFM en DNB).

Instellingen dienen ervoor te zorgen dat door hen gebruikte AI-systemen niet discrimineren.

Modellen die worden gebruikt om risico's van transacties in te schatten, of het nu gaat om bijvoorbeeld kredietrisico, verzekeringsrisico of witwasrisico, zullen al snel bepaalde kenmerken van de klant meewegen, naast kenmerken van de transactie zelf. Dit onderscheid tussen klanten mag gemaakt worden, maar niet op grond van godsdienst, levensovertuiging, politieke gezindheid, ras, geslacht, nationaliteit, seksuele gerichtheid of burgerlijke staat. Dit geldt voor zowel direct onderscheid als voor indirect onderscheid. Bij indirect onderscheid gaat het om een ogenschijnlijk neutrale bepaling die een bepaalde groep benadeelt. Zo kan het stellen van de eis dat een aanvrager van een bankrekening een woonadres in Nederland moet hebben gezien worden als indirect onderscheid naar nationaliteit.⁴⁶ Indirect onderscheid is wel toegestaan indien dat onderscheid objectief gerechtvaardigd, passend en proportioneel is. Zo kan in het aangehaalde geval het stellen van een woonvereiste een legitieme manier zijn voor een bank om anti-witwasmaatregelen te nemen. Maar het is niet noodzakelijk om aan het niet-voldoen aan het woonvereiste een automatische afwijzing te koppelen; de aanvrager zou ook op andere wijze kunnen aantonen dat witwasrisico's worden gemitigeerd. Wanneer modellen zelflerende algoritmen toepassen, kan het gebeuren dat er onbedoeld op een verboden manier indirect onderscheid wordt gemaakt. Een instelling zal dit moeten voorkomen, of wanneer het zich toch voordoet maatregelen dienen te treffen om het algoritme vrij van discriminerende effecten te maken.⁴⁷

⁴⁶ Oordeel College voor de Rechten van de Mens, 22-2-2023.

⁴⁷ Oordeel College voor de Rechten van de Mens, 1-8-2023.

Gegevensbescherming is een aandachtspunt, aangezien het ontwikkelen, trainen en gebruiken van AI gepaard gaat met de verwerking van grote hoeveelheden data. Hoe meer informatie en data een AI-model als input kan gebruiken, hoe beter het voorspellingen kan doen over bijvoorbeeld de kredietwaardigheid van een klant. De AVG stelt eisen aan de verwerking van persoonsgegevens, en geeft de klant het recht om niet onderworpen te worden aan een uitsluitend op geautomatiseerde verwerking gebaseerd besluit. Deze regels gelden ook voor het trainen en gebruiken van AI-modellen door financiële instellingen. Voor de verwerking van bijzondere persoonsgegevens, zoals gezondheid, politieke voorkeur en religie, geldt een verbod, tenzij een uitzondering van toepassing is zoals expliciete toestemming of vermelding in de wet. Voor een financiële instelling betekent dit dat niet zonder meer alle beschikbare informatie kan worden aangewend om een AI-model mee te ontwikkelen of te gebruiken.

Tussen het belang van gegevensbescherming en het voorkomen van discriminatie kunnen zich dilemma's voordoen. Ook als bijzondere persoonsgegevens nadrukkelijk niet worden gebruikt bij de ontwikkeling van een algoritme, kan het toch voorkomen dat een algoritme bepaalde bijzondere persoonsgegevens achterhaalt. Zo kan een AI-model aan de hand

van het betaalgedrag van een klant afgeleiden (ook wel 'proxies' genoemd) ontwikkelen voor bijzondere persoonsgegevens waar het algoritme, oorspronkelijk, helemaal geen beschikking over heeft. Voorbeelden hiervan zijn wanneer betalingen voor een medische behandeling als 'proxy' worden gebruikt door het AI-model om een inschatting te maken van de gezondheid van een klant. Een ander voorbeeld is wanneer woonadressen als 'proxy' dienen voor inschattingen van het inkomen van een klant. Zo kan het gebruik van een AI-model – ongewenst, en mogelijk ongemerkt – resulteren in een vorm van discriminatie. Om te borgen dat een AI-model 'eerlijk handelt' moet worden gecontroleerd, waar nodig met specifieke modellen, dat het model niet structureel bepaalde groepen discrimineert. Het is belangrijk dat financiële instellingen – met inachtneming van de handvatten geboden in de diverse wettelijke kaders – AI-modellen zo ontwerpen dat effectief gecontroleerd kan worden of het model discrimineert zodat hiertegen kan worden opgetreden. De AI-verordening voorziet in uitzonderlijke gevallen en onder strikte voorwaarden in een mogelijkheid om voor dit doel toch bijzondere persoonsgegevens te gebruiken. Op die manier kan een modelvalidatie worden uitgevoerd en kan het model zo nodig worden aangepast, of als dat niet goed mogelijk blijkt kan worden besloten het model niet te gebruiken.

Werkwijze

De inhoud van dit rapport is tot stand gekomen in samenwerking tussen diverse experts van de AFM en DNB. Vanuit multidisciplinaire teams zijn interviews gehouden met onder toezicht staande instellingen, brancheorganisaties, Europese toezichthouders, technologieleveranciers en academici. Daarnaast zijn diverse onderzoeken en analyses geraadpleegd (zowel intern als extern).

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl

Volg ons op:



Autoriteit Financiële Markten
PO Box 11723, 1001 GS Amsterdam
020 797 20 00
afm.nl

Follow us on:



DeNederlandscheBank

EUROSYSTEM

