



Rapport

Zicht op de digitale toekomst

Forecast Digitale Technologieën

TNOvector
Centre for Societal Innovation and Strategy

Inhoudsopgave

1. Inleiding	3
2. Onderzoeksaanpak	4
3. Forecast digitale technologie	5
3.1 Veertien opkomende digitale technologieën	6
3.2 Tijd tot grootschalige maatschappelijke adoptie	10
3.3 Maatschappelijke impact	10
3.4 Momentum	11
3.5 Combinatie van technologieën	11
3.6 Digitale autonomie	11
3.7 Conclusies	11
4. Verkenning publieke waarden	12
4.1 Agentic AI	12
4.2 AI-gestuurde robots	16
4.3 Reflectie – Overkoepelende lessen	20
Bijlage I: Maatschappelijke impactdomeinen	21
Bijlage II: Impact technologie per domein	22
Bijlage III: Expertworkshop en deelnemende experts	23

1. Inleiding

Samenvatting

- De forecast digitale technologie laat zien welke digitale technologieën in opkomst zijn en op welke maatschappelijke domeinen deze effect hebben. Veel rapporten over toekomstige technologieën geven beperkt aandacht aan de maatschappelijke dimensie van technologie. Deze forecast geeft daar invulling aan.
- Vooral agentic AI, digital twins, multiparty computing en quantum safe cryptography zullen op relatief korte termijn effect hebben op een breed spectrum aan maatschappelijke domeinen.
- Een combinatie van technologieën kan een groter maatschappelijk effect hebben dan een individuele technologie. Denk aan AI in combinatie robotics, quantum computing en digital twins.
- Welk maatschappelijk effect technologie daadwerkelijk heeft zal sterk afhangen van de mate waarin we grip krijgen op de ontwikkeling van de technologie.

De ontwikkelingen van digitale technologie volgen elkaar in rap tempo op. Digitalisering kan bijdragen aan het verbeteren van de dienstverlening van de overheid en het oplossen van maatschappelijke uitdagingen, zoals in de energietransitie en in de druk op de zorg. Tegelijkertijd zijn er zorgen over de impact van nieuwe technologieën op publieke en democratische waarden zoals privacy, autonomie en veiligheid. De digitale geletterdheid van burgers en gelijkheid in toegang tot en gebruik van nieuwe technologie zijn belangrijke factoren voor een goed functionerende digitale samenleving. De overheid heeft de taak om verschillende waarden en belangen bij elkaar te brengen en de inbedding van nieuwe, mogelijk disruptieve digitale technologieën in goede banen te leiden.

Goede informatie is belangrijk om tijdig een juiste keuze te kunnen maken tussen reguleren, stimuleren of anderszins vormgeven van innovatieprocessen voor nieuwe digitale technologieën. Er is daarom behoefte aan een technologiemonitor die op reguliere basis actuele informatie geeft over nieuwe en opkomende digitale technologieën en hun mogelijke maatschappelijke impact.

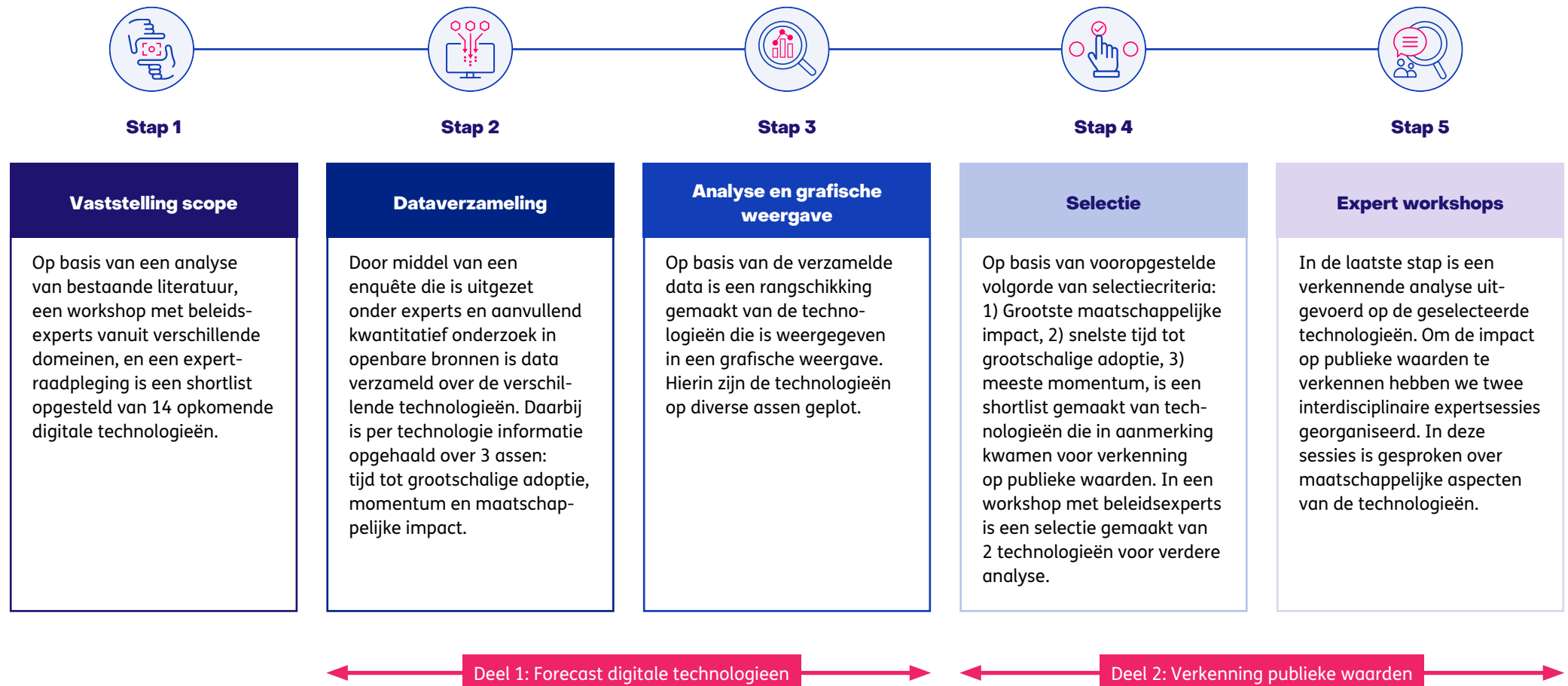
TNO Vector is gevraagd om hieraan bij te dragen door het opzetten van een forecast van opkomende technologieën en de potentiële impact daarvan op publieke waarden. In deze forecast is gekeken naar opkomende, domeinoverstijgende digitale technologieën met een potentieel voor grootschalige adoptie en impact.

De forecast digitale technologie (deel 1) geeft inzicht in de digitale technologieën die in de komende 10 jaar op ons afkomen en de reikwijdte van de maatschappelijke domeinen die hier effect van kunnen ondervinden. Deel 2 is een verkenning van publieke waarden. Voor twee van de technologieën uit de forecast verkennen we de mogelijke positieve en negatieve gevolgen voor publieke waarden: agentic AI en AI-gestuurde robots.

Op de volgende pagina vindt u een overzicht van de onderzoeksopzet. Op pagina 5 is de forecast digitale technologie weergegeven en in de navolgende pagina's toegelicht. Op pagina 12 vindt u de verkenning publieke waarden.

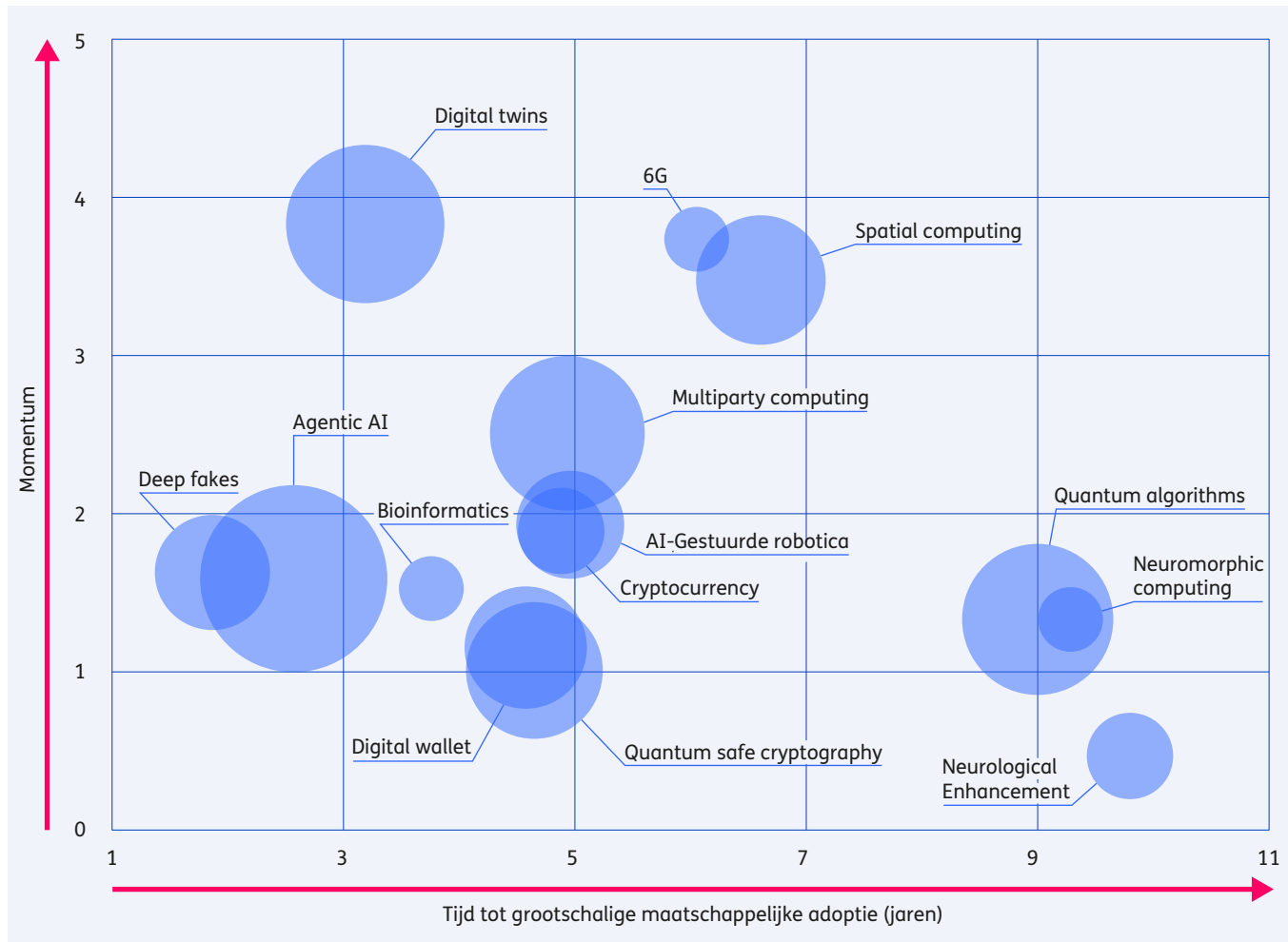


2. Onderzoeksaanpak



Figuur 1: Onderzoeksaanpak in vijf stappen

3. Forecast digitale technologie



Figuur 2: Grafische weergave van de forecast digitale technologie. Deze geeft weer welke digitale technologieën de komende tien jaar naar verwachting op ons afkomen. Op de horizontale as de verwachte tijd in jaren tot grootschalige maatschappelijke adoptie. De verticale as geeft het momentum weer. De grootte van de bol geeft de reikwijdte van de maatschappelijke impact weer.

De drie assen van de forecast digitale technologie:

Tijd tot grootschalige maatschappelijke adoptie

Verwachte tijd in jaren tot de technologie grootschalig in de maatschappij is geadopteerd. Hierbij wordt de technologie niet meer alleen door voorlopers gebruikt, maar is de technologie al meer gemeengoed geworden. De x-as geeft het gewogen gemiddelde weer van de tijdsduur in jaren die door experts werd aangegeven via een enquête.

Momentum

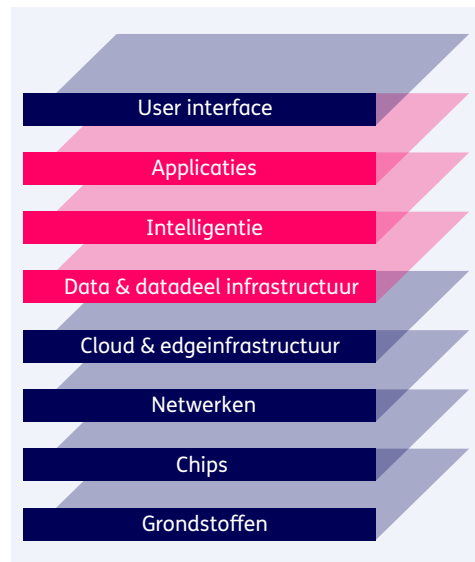
De as momentum geeft weer hoeveel interesse er voor de technologie is. Dit hangt samen met de waarschijnlijkheid dat een technologie grootschalige maatschappelijke adoptie zal bereiken. Om dit te meten zijn de volgende indicatoren gebruikt: investeringen, geregistreerde patenten, publicaties en zoektrends. Op de Y-as is dit weergegeven als de som van scores van genormaliseerde indicatoren.

Maatschappelijke impact

De grootte van de bol geeft inzicht in de reikwijdte van maatschappelijke domeinen die de digitale technologie naar verwachting zal raken. Vooraf zijn 15 maatschappelijke domeinen gedefinieerd. Experts kregen via een enquête de vraag op welke domeinen de technologie naar verwachting effect zou hebben. Een uitwerking van de impactdomeinen vindt u in bijlage 1.

3.1 Veertien opkomende digitale technologieën

Figuur 2 is een grafische weergave van digitale technologieën die de komende tien jaar naar verwachting op ons afkomen. De forecast geeft inzicht in de ingeschatte tijd voor grootschalige adoptie door de samenleving, het momentum en de maatschappelijke domeinen die naar verwachting effect zullen ondervinden.



Figuur 3: Het stack model: Digitale technologieën zijn opgebouwd uit diverse lagen, ook wel 'de stack' genoemd. Het stack model voor digitale technologie geeft dit overzichtelijk weer.

Digitale technologieën zijn opgebouwd uit diverse lagen, ook wel 'de stack' genoemd. Een stack model voor digitale technologie geeft dit overzichtelijk weer. Voor dit onderzoek maken we gebruik van het stack model van Freedom Lab, dat we enigszins versimpeld hebben (zie figuur 3).

De aandacht binnen dit onderzoek is gericht op het maatschappelijke effect van opkomende digitale technologieën. Er is voor gekozen om het onderzoek voornamelijk te richten op de bovenste lagen van het stack model; data & data-deel infrastructuur, intelligentie en applicaties, omdat deze naar verwachting de meest directe impact op de samenleving zullen hebben. De bovenste stacklaag, de user interface, hebben we buiten beschouwing gelaten. In die laag vindt interactie tussen mens en technologie plaats. Het is daarmee vooral de toegang van mensen tot de onderliggende technologieën, die een dienst of intelligentie leveren. Ook de onderste lagen lieten we buiten beschouwing. Op basis van een analyse van de bestaande literatuur over nieuwe technologie en met behulp van input van domeinexperts hebben we een shortlist gemaakt van veertien (opkomende) digitale technologieën (in willekeurige volgorde):

1 Spatial computing

Spatial computing is een technologie die het mogelijk maakt om de driedimensionale fysieke ruimte en de digitale omgeving met elkaar te vermengen. Spatial computing wordt ondersteund door en maakt gebruik van andere technologieën zoals bijvoorbeeld digital twins, augmented reality, virtual reality en AI. Door spatial computing wordt digitale content 'verankerd' in de fysieke wereld, waardoor gebruikers er op een natuurlijke manier mee kunnen interacteren, bijvoorbeeld met behulp van bewegingen die worden gemeten door bewegingssensoren.

Voorbeelden van toepassingen:

- Onderwijs op afstand.
- Verbeteren van toegankelijkheid van digitale of fysieke omgevingen voor mensen met een beperking.

2 Agentic AI

Agentic AI zijn AI-systemen die zelfstandig complexe taken kunnen plannen en uitvoeren, met beperkt menselijk toezicht. Agentic AI is in staat om lange termijn doelen na te streven, een plan op te stellen om die doelen te bereiken en zelfstandig acties te ondernemen om het plan uit te voeren, waar nodig met behulp van andere tools (denk bijvoorbeeld aan zoekmachines, chatbots en beeldgeneratoren). Daarbij is agentic AI adaptief en flexibel, leert het van eerdere ervaring en kan het inspelen op onverwachte gebeurtenissen. Agentic AI kan worden gezien als de volgende stap in de ontwikkeling van AI. Grotere datavolumes, geavanceerde algoritmen en verbeteringen in rekenkracht en opslag dragen bij aan de snelle evolutie en ontwikkeling van agentic AI.

Voorbeelden van toepassingen:

- Agentic AI kan in de toekomst marktonderzoek doen, complexe software bouwen en uitrollen, op basis van eigen inschatting online boodschappen doen, en eigen subagenten voortbrengen.

3 AI-gestuurde robotica

Fysieke robots geïntegreerd met geavanceerde AI systemen, zoals computer vision en large language models¹, waardoor de robot met behulp van natuurlijke taal met mensen kan communiceren en de context kan 'interpreteren' (inclusief het navigeren door fysieke ruimtes). Daardoor kunnen mensen en robots eenvoudiger samenwerken, in complexe omgevingen. In tegenstelling tot traditionele robots (die vooraf geprogrammeerde instructies volgen in een gecontroleerde omgeving), kunnen AI-gestuurde robots hun omgeving waarnemen, interpreteren en erop reageren op een flexibele manier.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Companion robots voor mensen met een beperking.
- Autonome voertuigen zoals drones en auto's.
- Robots die samenwerken met mensen in verschillende omgevingen, zoals de supermarkt horeca, ziekenhuis, school of bij de politie.
- Defensie-robots die taken uitvoeren onder gevaarlijke omstandigheden.

4 Quantumalgoritmen

Quantumalgoritmen zijn algoritmen die zijn ontworpen om de unieke eigenschappen van quantummechanica te benutten om complexe problemen efficiënter op te lossen. Ze worden specifiek ontwikkeld om te gebruiken op quantumcomputers voor snelle gegevensverwerking, optimalisatie en AI-toepassingen. Quantumcomputers bestaan nog niet, dus quantumalgoritmen worden op dit moment ontwikkeld en alleen in testomgevingen gebruikt. Zodra quantumcomputers commercieel beschikbaar komen, zullen deze algoritmen ook breder te gebruiken zijn.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- In de financiële sector kunnen quantumalgoritmes mogelijk helpen om rentes beter te voorspellen.
- Met quantumalgoritmes kunnen mogelijk patronen herkend worden in grote datasets, die nu nog verborgen zijn. Dit kan in veel sectoren leiden tot toepassingen, bijvoorbeeld het personaliseren van diagnose en behandeling in de zorg, fraude-detectie, in data-gedreven onderzoek en beleidsanalyses.

5 Deep fakes

Het genereren van valse digitale inhoud of manipulatie van echte inhoud met behulp van deep learning-technieken. De inhoud omvat onder andere video, afbeeldingen, audio en tekst. De populariteit komt voornamelijk voort uit de manipulatie van het uiterlijk van het gezicht (attributen, identiteit, uitdrukking), meestal ingedeeld in de volgende categorieën: (i) synthese van het hele gezicht, (ii) manipulatie van een attribuut, (iii) identiteitswisseling en (iv) expressieruil (d.w.z. re-enactment). Deepfakes bestaan al, maar worden snel geavanceerder. Ook veroorzaakt deze technologie een technologierace voor deepfake-detectietechnologieën.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- nieuwe mogelijkheden in het maken van films, betaalbare videoproductie, digitale avatars, enz.
- Social engineering aanvallen.
- Politieke desinformatie.
- (Niet-consensuele) pornografie.

6 Bioinformatica

In de bioinformatica worden databases met biomoleculaire gegevens gemaakt waarop berekeningen worden toegepast, met behulp van algoritmen en statistische technieken. Daarmee worden op basis van (digitale) biologische gegevens, zoals DNA- en aminozuursequenties of annotaties over die sequenties, inzichten verkregen over biomoleculaire mechanismen en biologische processen, bijvoorbeeld de rol van genen in fysiologische processen en het ontstaan van ziekten. De toenemende beschikbaarheid van technologieën die grote hoeveelheden (digitale) biologische gegevens verzamelen drijft de vraag naar (complexe) computertechnologie om deze gegevens te verwerken en te analyseren om wetenschappelijk relevante inzichten en informatie af te leiden.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Next Generation Whole Genome Sequencing.
- Multi-omics analyse (gezamenlijke analyse van verschillende profielen, zoals DNA (genomics), eiwitten (proteomics) en metabolisme (metabolomics)).
- Alpha Fold (protein folding prediction).

¹ Computer vision (CV) : AI systemen die informatie kunnen afleiden uit afbeeldingen en video's. Bijvoorbeeld het herkennen van personen, objecten, dieren of andere patronen. Large Language Models (LLM's) zijn AI systemen die machine learning technologie toepassen om natuurlijke tekst te interpreteren en genereren.

7 Digital Wallet

Toepassing die wordt gebruikt om privé- en publieke sleutels te genereren, beheren, opslaan of gebruiken. Deze sleutels zijn gekoppeld aan digitale inloggegevens, voor identiteitskenmerken (bijv. leeftijd, rijbewijs), voor (crypto-)valuta's of andere eigendomsclaims. Een wallet kan worden geïmplementeerd als een software- of hardware module. Belangrijk is dat de digitale portemonnee het mogelijk maakt om attributen selectief openbaar te maken. Hoewel digitale portemonnees al langer bestaan in de blockchain-ruimte, worden ze nu door de publieke sector uitgerold als middel voor burgers om zich veilig online te authenticeren en te identificeren.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Europese digitale identiteit: gebruikers krijgen toegang tot onlinediensten, voor het opslaan en delen van digitale documenten, en om handtekeningen te zetten.
- Metamask: stelt gebruikers in staat om te communiceren met de Ethereum-blockchain en om fungibele (inwisselbare) en niet-fungibele tokens (NFT's) op te slaan, te verhandelen en uit te wisselen.

8 Neuromorphic computing

Een IT-paradigma dat inspiratie haalt uit de biologische structuur en werking van hersenen, en dat nabootst in niet-biologische hardware. Vergeleken met bestaande elektronica verbruiken de hersenen zeer weinig energie voor het verwerken en opslaan van data. Dat komt doordat het opslaan en verwerken van data op dezelfde locatie in de hersenen plaatsvindt (in netwerken van hersencellen). Het is dus niet nodig om data van de opslaglocatie naar de verwerkingslocatie te verplaatsen, zoals nu vaak gebeurt in hardware (van het geheugen naar de processor). Dit maakt parallele rekenprocessen mogelijk. Het idee van neuromorphic computing is om daarmee de snelheid en energie-efficiëntie van informatieverwerking te vergroten.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Bijvoorbeeld autonome voertuigen (auto's en drones) die beelden moeten interpreteren en omzetten in actie, brein-computer interfaces, zoals neuroprothesen en neuro-implantaten en real-time monitoring van complexe (productie) processen.

9 Quantumveilige cryptografie

Cryptografie verwijst naar algoritmen en methodologiën die data versleutelen zodat niet-geautoriseerde personen er niet bij kunnen. Huidige, traditionele computers hebben jaren nodig om deze versleuteling te kraken. Quantumcomputers bestaan op dit moment nog niet, maar als ze op de markt komen zullen ze daar wel toe in staat zijn. Quantumveilige cryptografie verwijst naar cryptografische methoden die zijn ontworpen om bestand te zijn tegen de potentiële bedreigingen van quantumcomputers, zodat data nog steeds veilig te versleutelen is als quantumcomputers commercieel beschikbaar komen.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Gevoelige gegevens (bijv. financiën, gezondheid, persoonlijke gegevens) veilig houden in een wereld van quantumcomputing.
- Beveiligen van IP tegen (door de staat gesponsorde) hackers of kwaadwillende concurrenten om technologische superioriteit te behouden.

10 Cryptocurrency

Een cryptocurrency is een digitale valuta, die cryptografisch kan worden verzonden van de ene gebruiker op een blockchain netwerk naar de andere, zonder dat daar een centrale autoriteit (zoals bijvoorbeeld een bank) tussen zit.

Op dit moment zien we eerste institutionele bewegingen in de richting van adoptie. De twee belangrijkste trends zijn stablecoins en digitale valuta's van centrale banken (zoals de digitale euro). Dit betekent dat steeds meer mensen ermee in aanraking zullen komen, met name ook in de machine-economie (handel tussen robots, satellieten, enz.).

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Bitcoin, Ethereum, stablecoins, enz.
- In sommige gevallen digitale valuta's van de centrale banken (Central Bank Digital Currencies).

11 **Multiparty computing**

Secure multiparty computing (MPC) verwijst naar een verzameling tools waarmee meerdere partijen gezamenlijk berekeningen kunnen uitvoeren op elkaars gegevens. Cruciaal is dat de partijen, vanwege het gebruik van geavanceerde cryptografische methoden, nooit elkaars werkelijke gegevens kunnen bekijken. De deelnemende partijen bepalen wie de uitkomst van de berekening mag inzien. MPC maakt het mogelijk om te privacygevoelige gegevens te analyseren buiten de organisatorische silo's.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Analyseren van grote hoeveelheden zorggegevens van meerdere organisaties om de ontwikkeling van nieuwe behandelingen mogelijk te maken.
- De ontwikkeling van empirisch onderbouwd beleid, op basis van datasets die in het bezit zijn van verschillende overheidsdiensten, - zonder risico op inbreuk op de privacy.

12 **6G**

Zesde generatie van draadloos mobiel netwerk met THz-frequentiebanden, met hogere snelheden, lagere latentie en grotere bandbreedte dan eerdere generaties (3G, 4G, 5G). De verwachte bandbreedte van 6G ligt 10 keer hoger dan de downloadsnelheid van 5G. Door de inzet van AI kan het netwerk optimaal benut worden en real-time gereageerd op verstoringen. Mede daardoor is 6G betrouwbaar dan 5G. Tot slot is de dekking van 6G beter dankzij nieuwe radiotechnologie en het combineren van satelliettechnologie met mobiele netwerken.

NB: Hoewel deze technologie niet behoort tot de geselecteerde stack lagen is deze toch meegenomen in de analyse, vanwege de maatschappelijk belangstelling voor het onderwerp.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Real-time en privacy-veilig monitoren en sturen van verkeersstromen.
- Autonoom rijden.
- Industrie 4.0.
- Slimme steden.

13 **Digital Twins**

Een digital twin is een virtuele weergave van een fysiek object, een bestaand systeem of proces. Het is in feite een 'digitale replica' die het gedrag, de kenmerken en de prestaties van de fysieke wereld weerspiegelt met behulp van realtime data, simulatie en modellering. Deze digitale replica's kunnen bijvoorbeeld worden gebruikt voor simulaties, optimalisatie, monitoring en besluitvorming.

Door het gebruik van AI liggen er steeds meer kansen om digital twins in te zetten voor het voorspellen van bepaalde ontwikkelingen, of het simuleren van fictieve scenario's

Voorbeelden van (mogelijk toekomstige) toepassingen:

- gebouwde omgeving (smart city) en infrastructuur.
- Gepersonaliseerde medicijnen.
- Simulatie van leefomgeving zodat scenario's voor beleidskeuzen gemaakt kunnen worden.

14 **Neurological enhancement**

Het verbeteren van de cognitieve vaardigheden van personen, met behulp van (digitale) technologie die een verbinding leggen tussen de hersenen en een extern apparaat (zogenaamde brein-computer interfaces - BCI). Neurotechnologieën lezen, analyseren en interpreteren hersenactiviteit en zetten dat om in bruikbare signalen, bijvoorbeeld inzicht in concentratie of slaperigheid. Sommige neurotechnologieën manipuleren ook direct de hersenactiviteit. Andere manieren om de cognitieve vaardigheden te verbeteren is neurofarmacologie, waarbij medicatie of andere middelen (drugs, alcohol) worden gebruikt om de hersenactiviteit te beïnvloeden.

Voorbeelden van (mogelijk toekomstige) toepassingen:

- Hersenimplantaten (zoals van Neuralink).
- Neuroprotheses.
- EEG headbands.
- Neuromodulatie.

Nadat de technologieën waren vastgesteld is in de volgende stap informatie verzameld om ze verder in kaart te brengen. Hierbij hebben we gebruik gemaakt van openbare databronnen en een expert enquête die 64 experts van TNO hebben ingevuld. De drie assen van de digitale technologie forecast (figuur 2) geven inzicht in de tijd die een technologie naar verwachting nodig heeft om breed geadopteerd te worden in de maatschappij en het momentum dat er voor een technologie is. Daarnaast hebben we een analyse gedaan naar het aantal maatschappelijke domeinen waarop de verschillende technologieën effect zullen hebben.

3.2 Tijd tot grootschalige maatschappelijke adoptie

Technologieën worden geleidelijk geadopteerd. Het begint vaak met adoptie door een klein aantal 'pioniers', waarna het gebruik mogelijk uitbreidt naar het grote publiek of bedrijven. De impact op publieke waarden zal over het algemeen worden opgemerkt wanneer de adoptie verder gaat dan de 'pioniers' naar de meerderheid. We noemen dat tijd tot grootschalige maatschappelijk adoptie. Via een enquête is experts gevraagd om per technologie een inschatting te geven:

Wanneer verwacht je dat deze technologieën mainstream zullen worden? (in jaren)

De horizonscan laat zien dat met name van **Agentic AI**, **deep fakes**, **digital twins** en **bioinformatics** verwacht wordt dat ze op korte termijn gangbaar zijn. Een aantal technologieën, zoals deep fakes en digital twins worden gebruikt door voorlopers, maar maken nog een ontwikkeling door.

3.3 Maatschappelijke impact

Om een beeld te krijgen van de reikwijdte van het maatschappelijke effect zijn 15 domeinen gedefinieerd (zie figuur 4). Op verzoek van het ministerie zijn minderjarigen toegevoegd als doelgroep. Via een enquête is aan experts gevraagd om per technologie aan te geven op welke domeinen ze effect zullen hebben:

Selecteer de domeinen waarin de technologie(ën) waarschijnlijk een aanzienlijke maatschappelijke impact zullen hebben.

Technologieën werden geacht effect te hebben op een domein, als het domein werd geselecteerd door ten minste 50% van de respondenten (zie uitwerking in bijlage II).

Sommige technologieën hebben slechts in een beperkt aantal domeinen effect, zoals bijvoorbeeld **bioinformatics** waarbij het effect naar verwachting vooral binnen de gezondheidszorg merkbaar zal zijn.

Andere technologieën raken veel meer maatschappelijke domeinen, waardoor het effect van de technologie breder voelbaar zal zijn. Vooral **Agentic AI** springt eruit. Deze technologie zal 9 van de 15 domeinen raken, denk bijvoorbeeld aan de defensie, digitale infrastructuur en onderwijs.

Daarnaast scoren **digital twins** (7), **multiparty computing** (6) en **quantum algoritmen** (6) hoog op aantal maatschappelijke domeinen die de technologie naar verwachting zal beïnvloeden.

Respondenten geven vooral aan impact te verwachten van digitale technologieën in de domeinen **defensie en nationale veiligheid en digitale infrastructuur**, waar 9 van de 14 technologieën effect zullen hebben. Ook **economie en ondernemersklimaat** (7), **financiële infrastructuur** (6), **dienstverlening overheid** (5) en **gezondheidszorg** (5) zullen relatief vaak effect ervaren van digitale technologieën. Zie een volledig overzicht in bijlage II.

- Defensie en nationale veiligheid
- Digitale infrastructuur
- Economie en ondernemersklimaat
- Onderwijs en wetenschap
- Natuur en klimaat
- Energie
- Financiële infrastructuur
- Voedselzekerheid
- Justitie en rechtsstaat
- Overheidsdiensten
- Gezondheidszorg
- Openbare ruimte en ordening
- Sociale zaken en werkgelegenheid
- Telecommunicatie,
- Waterbeheer, infrastructuur en transport
- Groep: Minderjarigen

Figuur 4: De 15 maatschappelijke domeinen en 1 groep

In de expert enquête is ook gevraagd naar speciale aandachtspunten op het gebied van publieke waarden. Opvallend is dat er vooral veel gewezen wordt op het effect van agentic-AI. Vaak wordt hierbij gewezen op problematische afhankelijkheid van niet-Europese partijen.

“AI-agents hebben geen moreel kompas of notie van normen en waarden, tenzij ze worden onderwezen (welke? hoe? door wie?). Tegelijkertijd zullen ze zeer indringend zijn in de samenleving en zullen de “fabrikanten” onze normen en waarden mogelijk niet onderschrijven. Van alle opkomende digitale technologieën in de bovenstaande lijst verwacht ik dat AI-agents de meeste impact zullen hebben op publieke waarden.” Respondent 29

3.4 Momentum

Momentum gaat over de interesse die er op dit moment voor een technologie is. Het gaat dan om indicatoren als publieke en private investeringen, het aantal patenten en wetenschappelijke publicaties. We zien dat met name **digital twins**, **6G**, **spatial computing** en **multiparty computing** hoog op deze indicatoren scoren.

3.5 Combinatie van technologieën

De technologieprognose geeft inzicht in de status van individuele technologieën.

Maar geraadpleegde experts geven ook aan dat een combinatie van technologieën elkaar kunnen versterken.

“Sommige van deze opkomende technologieën hebben op zichzelf niet zo’n impact, het is de combinatie met (opkomende) technologieën die impact creëert.” Respondent 62

Zo is de ontwikkeling van AI in een stroomversnelling geraakt toen er meer rekenkracht beschikbaar kwam. Ook hadden huidige ontwikkelingen op het gebied van generatieve AI niet kunnen plaatsvinden zonder ontwikkelingen op het gebied van cloud.

“De combinatie van AI-agents en robotics zou groter kunnen zijn dan de som, omdat dit zou resulteren in autonome robots/systemen die de huidige mogelijkheden aanzienlijk overtreffen” Respondent 41

49 van de in totaal 64 respondenten in de expert survey bevestigen dat een combinatie van technologieën meer impact kan hebben dan effecten daarbij diverse voorbeelden. Maar ook andere combinaties komen terug.

“6G + AI + digital twinning zal een groot effect hebben op de convergentie tussen de fysieke en virtuele wereld, veel groter dan de individuele technologie” Respondent 32

3.6 Digitale autonomie

Ook wordt strategische afhankelijkheid van niet-Europese aanbieders bij bijna alle technologieën aangemerkt als risico.

“Voor alle genoemde technologieën zal Nederland afhankelijk zijn van een niet-Europese speler. Veel onderwerpen binnen de digitale thema’s worden bezet door grote techbedrijven uit voornamelijk de VS, die ook een groot deel van de infrastructuur leveren. Door afhankelijk te zijn van deze infrastructuur en kennis, is Nederland kwetsbaarder voor de geopolitieke situaties. Als vitale infrastructuur zoals overheidsdiensten en defensie- en veiligheidsdiensten gebaseerd zijn op niet-Europese leveranciers, is de toegang niet altijd gegarandeerd. Ook als de tech-gap met andere partijen te groot wordt, kan dat economische gevolgen hebben, zoals dat de Nederlandse economie niet in een gunstige concurrentiepositie verkeert.” Respondent 57

3.7 Conclusies

- Er zijn 14 digitale technologieën geïdentificeerd die de komende 10 jaar op de samenleving afkomen (zie figuur 2).
- Een aantal van deze technologieën zal effect hebben op een breed spectrum aan maatschappelijke domeinen.
- Uit de forecast komt naar voren dat vooral agentic AI (AI-agents) breed maatschappelijk effect zal hebben. Bovendien geven experts aan dat hier specifieke zorgen zijn over strategische afhankelijkheid van niet-Europese fabrikanten.
- Een combinatie van technologieën kan een groter maatschappelijk effect hebben dan een individuele technologie. Denk aan agentic AI in combinatie met robotics, quantum en digital twins.
- Digitale autonomie is bij alle technologieën een punt van zorg, omdat veel van deze technologie ontwikkeling niet in Europa plaatsvindt.

4. Verkenning publieke waarden

4.1 Agentic AI

Agentic AI is een opkomende technologie die binnen afzienbare tijd zijn weg zal vinden naar de samenleving. Welke gevolgen zal dat hebben? Om dat te onderzoeken hebben we een verkenning uitgevoerd naar de impact van agentic AI op de samenleving. Daarvoor hebben we onderzocht welke publieke waarden – waarden die we als samenleving belangrijk vinden om te waarborgen – onder druk komen te staan of juist worden gestimuleerd als agentic AI op grote schaal wordt gebruikt.

Om een beter beeld te krijgen van de technologie agentic AI en de kansen en risico's voor publieke waarden in kaart te brengen hebben we met experts gesproken en een expertworkshop georganiseerd. Zie bijlage III voor meer details.

Wat is agentic AI?

De termen AI agents en agentic AI zijn sterk in opkomst en kunnen worden gezien als de volgende stap in de ontwikkeling van AI. Onder de door ons consulteerde experts was er geen overeenstemming over een eenduidige definitie van agentic AI, maar ze waren het erover eens dat *agency* een belangrijk kenmerk is van dit type AI. Agency houdt in dat een systeem zelfstandig

complexe taken (taken die bestaan uit een reeks van opeenvolgende deeltaken) kan plannen en uitvoeren, met beperkt menselijk toezicht. Een AI systeem met veel *agency* ('agentic AI') is in staat om lange termijn doelen na te streven, een plan op te stellen om die doelen te bereiken en zelfstandig acties te ondernemen om het plan uit te voeren, waar nodig met behulp van andere tools (denk bijvoorbeeld aan zoekmachines, chatbots en beeldgeneratoren). Daarbij is het adaptief en flexibel, leert het van eerdere ervaring en kan het inspelen op onverwachte gebeurtenissen.

Hoewel het moeilijk is agentic AI helder af te bakenen, is het wel mogelijk om te spreken van AI systemen met meer of minder *agency*. De huidige beschikbare "AI agents" hebben vaak een vooraf gedefinieerde workflow of beslisboom die een AI systeem aanroept als daarom wordt gevraagd. Een voorbeeld is een chatbot die reageert op de vraag van een gebruiker door een proces in gang te zetten of de vraag door te zetten naar de juiste personen. De taken die ze uitvoeren zijn dus weinig complex, en de systemen zijn niet heel zelfstandig of adaptief. AI ontwikkelaars richten zich op dit moment op het ontwikkelen van AI systemen met meer *agency*. In plaats van dat die systemen vooral korte, simpele opdrachten uitvoeren, zijn dat doelgerichte actoren die in staat zijn om te leren, zich aan te passen en zelfstandig

en adaptief acties te ondernemen om complexe doelen te bereiken, zonder dat hun gedrag van tevoren is gespecificeerd.

Toepassingen van agentic AI

(Toekomstige) toepassingen van agentic AI die in onze gesprekken met experts naar boven kwamen zijn AI systemen die:

- als persoonlijke assistent of vriend fungeren en proactief suggesties kunnen doen ('ik zie dat je dit onderwerp onderzoekt, misschien moet je met deze persoon praten');
- zelfstandig de aandelenmarkt onderzoeken en aandelen kopen;
- op basis van verschillende modaliteiten (beeldanalyse, lab uitslagen, verhaal van de patiënt) een diagnose stellen;
- reisplannen, vluchtschema's en hotelovernachtingen uitwerken (en boeken);
- zelfstandig software bouwen en uitrollen;
- online boodschappen doen voor een huishouden; of
- iemands belastingaangifte doen.

Wat zeggen experts?

Tijdens de expertsessie bespraken we twee toekomstige toepassingen van agentic AI: AI als vriend en AI bij de belastingdienst om schulden te voorkomen. Voor beide toepassing bespraken we wat de belangrijkste (positieve en negatieve) gevolgen voor gebruikers zouden kunnen zijn.

Voor **AI als vriend** ging het daarbij bijvoorbeeld over de voor- en nadelen van het hebben een virtuele vriend. Een AI vriend is laagdrempelig, altijd beschikbaar, oordeelt niet, kan de efficiëntie van de gebruiker verbeteren en de toegang tot de (al bestaande) virtuele wereld vergroten. Tegelijk levert de gebruiker ook veel (intieme data) in, kunnen sociaal-emotionele vaardigheden ook verminderen en vergroot het de kans op stigmatisering en manipulatie.

Voor **AI bij de belastingdienst** waren efficiëntie en toegankelijkheid van de diensten en personalisering van de aanpak mogelijke voordelen. Negatieve gevolgen die werden genoemd waren dat het systeem minder goed werkt voor minderheden, dat de privé sfeer wordt ondermijnd en overmatig vertrouwen op metingen en data.

In het tweede deel van de sessie hebben we de blik verbreed en gekeken naar de gevolgen voor de samenleving in het algemeen. Daarin bespraken we ook wat volgens deelnemers de grootste kansen en risico's voor publieke waarden waren. De uitkomsten van dat gesprek zijn weergegeven op de volgende twee pagina's.

Agentic AI | Verkenning publieke waarden



Privacy

Agentic AI verzamelt en verwerkt veel **persoonlijke gegevens** van gebruikers. De gegevens worden gebruikt om acties en communicatie goed af te stemmen op de individuele gebruiker, en om acties uit te voeren in **andere systemen**. Bijvoorbeeld in een online reisbureau om een reis te boeken, of met een ander AI systeem om een gezamenlijke afspraak te maken.

Op die manier raken verschillende systemen met elkaar **verknoot**, en wordt de data die is verzameld door één agentic AI systeem verspreid over veel andere systemen. Omdat agentic AI zelfstandig werkt is het voor de gebruiker moeilijk te overzien **welke** data **waar** terechtkomt en **wat** er met de data gebeurt.

De grote hoeveelheid aan verzamelde persoonlijke data maakt het ook mogelijk om zeer gedetailleerd te **profiëren**. Dat kan uitsluiting of discriminatie in de hand werken, of manipulatie. Zie ook 'autonomie en transparantie.'



Autonomie en transparantie

Agentic AI is adaptief en leert gaandeweg over de gebruiker. Daarmee worden de ervaringen van gebruikers **gepersonaliseerd**. De interactie met agentic AI voelt voor de gebruiker als interactie met een persoon. Daardoor wordt het lastiger om onderscheid te maken tussen **communicatie** met een mens en agentic AI (gebrek aan transparantie.)

Agentic AI maakt ook **impliciete aannames** over de gebruiker die de uitvoering van taken beïnvloeden, zonder dat dit inzichtelijk is voor de gebruiker. Daarbij komt dat Agentic AI zelfstandig werken. Controle op redeneringen is lastig en uitkomsten ondoorzichtig.

Manipulatie wordt mogelijk eenvoudiger en krachtiger door de gedetailleerde profilering, aanpassing aan de leefwereld en overtuiging van de gebruiker. Daarmee kunnen gebruikers worden aangezet om bijvoorbeeld bepaalde producten te kopen of een stem uit te brengen, wat ze anders niet gedaan zouden hebben.



Toegankelijkheid en inclusiviteit

Omdat agentic AI adaptief is en leert van de interacties met gebruikers, kunnen agentic AI systemen ontwikkeld worden die zich qua **communicatie** aanpassen aan de **vaardigheden** en **leefwereld** van gebruikers, om zo de **toegankelijkheid tot de digitale wereld** te vergroten voor mensen met lage (digitale) geletterdheid. Als agentic AI systemen niet worden ontwikkeld om rekening te houden met verschillende soorten gebruikers, kan dat de digitale kloof vergroten. Het risico bestaat dan dat de voordelen van agentic AI vooral ten goede komen aan **digitaal geletterde en welvarende gebruikers**, omdat voor hen de drempel om de systemen te gebruiken lager is.

Net als andere soorten AI, baseert agentic AI zijn adviezen en handelingen op **gemiddelden** van beschikbare data. Daardoor kan deze slecht omgaan met *outliers*, personen die ver van het gemiddelde liggen. Omdat agentic AI vaak zelfstandig acties onderneemt kan het nadelige effect versterkt worden. Er is dan geen mens betrokken om fouten tijdig terug te draaien.

Agentic AI | Verkenning publieke waarden



Democratie

Commerciële aanbieders hebben er mogelijk baat bij dat gebruikers een zo lang mogelijke interactie aangaan met hun AI systeem. Het idee is dat gebruikers een systeem dat gebruikers niet tegensprekt en hun denkbeelden bevestigt of versterkt voor langere tijd zal gebruiken. Dit **echokamereffect** is al bekend van huidige algoritmen en AI systemen. Omdat agentic AI in staat is om verder te personaliseren, en het ook een directe ingang biedt tot andere informatie en websites (bijvoorbeeld door websites of evenementen aan te bevelen of zelfs tickets te regelen), wordt dit effect nog eens versterkt. Gelijkgestemden ontmoeten vooral gelijkgestemden, online en offline. Het gevolg kan een (verdere) **polarisatie** van de maatschappij zijn, wat democratische processen (gericht op deliberatie en consensus) bemoeilijkt.

Het is ook mogelijk om agentic AI toepassingen te ontwikkelen die kansen bieden om **connectiviteit** en **deliberatie** te vergroten. Bijvoorbeeld als het gericht is op het in contact komen met andersdenkenden en overwegen van andere perspectieven.



Sociale verbinding

Agentic AI biedt kansen voor inclusief gebruik (zie ook toegankelijkheid en inclusiviteit). Agentic AI kan daarbij ook een ingang bieden tot de virtuele wereld, en via die ingang ook helpen met het vinden en organiseren van mogelijkheden tot **sociale interactie**. Denk bijvoorbeeld aan het vinden van activiteiten die in de buurt of wijk worden georganiseerd, of het samenbrengen van eenzame personen. Waar iemand nu zelf gericht moet zoeken naar informatie op internet, kan agentic AI daarin bemiddelen.

Anderzijds bestaat er een risico dat gebruikers door het gebruik van persoonlijke agentic AI (zoals een persoonlijke assistent, vriend of collega) minder gezamenlijke ervaringen opdoen. Deze zijn belangrijk om de sociaal-emotionele vaardigheden te ontwikkelen die nodig zijn voor deelname aan **maatschappelijke processen**. Dit vergroot individualisering en verkleint sociale verbinding. Ook bestaat er het risico om sterk **afhankelijk** worden van deze systemen, bijvoorbeeld om met anderen te communiceren als vriend, assistent of collega. Daarmee het verlies aan menselijke (sociale en communicatieve) vaardigheden.



Duurzaamheid

Agentic AI vereist enorme **rekenkracht**, waarvoor grote datacentra nodig zijn die veel **elektriciteit** en **water** verbruiken. Dit gebeurt steeds meer buiten het zicht van gebruikers, waardoor die geen direct inzicht hebben in het verbruik.

Bovendien heeft ook de fabricage van de benodigde hardware (voor de apparatuur in de datacentra en de elektronische apparaten die nodig zijn om een agentic AI te gebruiken) een grote **ecologische impact**. Denk bijvoorbeeld aan het delven van de materialen voor batterijen en elektronica, en de (soms giftige) afval die ontstaat als de apparaten worden weggegooid.

Aandachtspunten

Context-afhankelijkheid

Uit de verkenning van de gevolgen van agentic AI voor publieke waarden wordt duidelijk dat agentic AI zowel positieve als negatieve gevolgen kan hebben. Of het gebruik van agentic AI positief of negatief uitpakt voor een bepaalde waarde (zoals democratie, toegankelijkheid of sociale verbinding) is sterk afhankelijk van hoe en waarvoor de technologie wordt ontwikkeld en hoe die wordt toegepast. De maatschappelijke gevolgen van agentic AI zijn dus sterk afhankelijk van de precieze toepassing en de context waarin deze wordt geïmplementeerd. Agentic AI biedt kansen om de toegankelijkheid tot de digitale wereld te vergroten, en voor mensen om zich te verplaatsen in anderen, maar kan juist ook de toegankelijkheid verkleinen of de afstand tot anderen vergroten.

Het ontwikkelen van prototypes waarmee wordt geëxperimenteerd in praktijk-situaties, waarbij zowel ontwikkelaars als toekomstige gebruikers en andere betrokkenen meewerken kan hier meer inzicht geven.

Wie ontwikkelt, beheert en bezit de technologie?

De gevolgen van agentic AI zijn dus afhankelijk zijn van de precieze toepassing en gebruikscontext. Hoe die toepassingen eruit zullen zien en hoe ze worden gebruikt wordt voor een groot deel bepaald door

partijen die de technologie zullen ontwikkelen, beheren en bezitten. Langdurig zicht en grip op de technologie is cruciaal om te waarborgen dat agentic AI wordt ontwikkeld en toegepast op basis van voor de Nederlandse samenleving centrale waarden.

Experts gaven aan dat er veel data, rekenkracht en financiële investeringen nodig zijn voor de ontwikkeling van agentic AI. Het ligt daarom voor de hand dat buitenlandse big tech bedrijven een grote speler zullen zijn. Volgens de experts die deelnamen aan de workshop verkleint dit de mogelijkheid om controle uit te oefenen op de ontwikkelingen en vergroot het de digitale afhankelijkheid van deze bedrijven. Bovendien concentreert macht over data en informatievoorziening zich in een kleine aantal commerciële bedrijven.

Onvoorstelbare toekomst

De toekomst van agentic AI is mogelijk complexer en verregaander dan wij ons op dit moment kunnen voorstellen. De ontwikkelingen volgen elkaar snel op, en de combinatie van verschillende typen AI systemen biedt wellicht mogelijkheden en toepassingen die nu nog buiten onze voorstellingsvermogen liggen. Dit vraagt om continue monitoring en deliberatie. Een voorbeeld van een moeilijk te voorspellen gevolg die door experts werd genoemd is dat het gebruik van agentic AI ook een grote impact kan hebben op de economie

van het internet. Gebruikers krijgen via één punt toegang tot het internet voor het aanschaffen van producten en diensten, in plaats van dat ze verschillende website-suggesties zien en bekijken. Een ander voorbeeld zijn gevolgen voor de arbeidsmarkt. Omdat veel toekomstige toepassingen in het werkdomein vallen, is de kans groot dat werkzaamheden zullen veranderen, maar het is moeilijk te overzien welke impact agentic AI zal hebben op welke beroepen en vaardigheden.

Conclusies

- Agentic AI biedt kansen, bijvoorbeeld om complexe taken te automatiseren en om de ervaring van gebruikers in de digitale wereld te personaliseren.
- Risico's ontstaan wanneer agentic AI buiten menselijke controle opereert.
- Langdurig zicht en grip op de technologie is cruciaal om te waarborgen dat agentic AI wordt ontwikkeld en toegepast op basis van voor de Nederlandse samenleving centrale waarden.
- Afhankelijk van de grip die we krijgen op de ontwikkeling van deze technologie en de specifieke toepassingen daarvan kan agentic AI in meer of mindere mate effect hebben op publieke waarden zoals privacy, autonomie, transparantie, toegankelijkheid en inclusie, democratie, sociale verbinding en duurzaamheid.

4.2 AI-gestuurde robots

AI-gestuurde robots zijn een opkomende technologie. Hoewel grootschalige toepassing nog ver weg lijkt, zullen robots op steeds meer plekken in de samenleving te zien zijn. Welke gevolgen zal dat hebben? Om dat te onderzoeken hebben we een verkenning uitgevoerd. Daarin hebben we onderzocht welke publieke waarden – waarden die we als samenleving belangrijk vinden om te waarborgen – onder druk komen te staan of juist worden gestimuleerd als AI-gestuurde robots op grote schaal wordt gebruikt.

Om een beter beeld te krijgen van de technologie AI-gestuurde robots en de kansen en risico's voor publieke waarden in kaart te brengen hebben we met experts gesproken en een expertworkshop georganiseerd. Zie meer informatie in bijlage III.

Wat zijn AI-gestuurde robots?

AI gestuurde robots combineren een fysieke (mechanische) component die de fysieke wereld beïnvloedt, met software die de fysieke component aanstuurt. De integratie van LLM's en andere geavanceerde AI zoals computervisie met fysieke robotica maken natuurlijk aanvoelende interactie met mensen mogelijk.

Ook kunnen AI-gestuurde robots zich aanpassen aan de omgeving. Daardoor

kunnen mensen en robots eenvoudiger samenwerken, in complexe omgevingen.

In tegenstelling tot traditionele robots (die vooraf geprogrammeerde instructies volgen in een gecontroleerde omgeving), kunnen AI-gestuurde robots hun omgeving waarnemen, interpreteren en erop reageren op een flexibele manier, vaak met behulp van sensoren, computervisie en natuurlijke taalverwerking.

Uit onze gesprekken met experts blijkt dat dit nieuwe toepassingen mogelijk maakt, waarbij robots in een niet gestandaardiseerde omgeving, zoals kassen, supermarkten, hotels en restaurants, scholen of ziekenhuizen verschillende taken uitvoeren. Bijvoorbeeld om schappen te vullen, tafels af te ruimen, inventariseren, schoon te maken of gezelschap te bieden. De nieuwe toepassingsgebieden zijn plekken waar ook mensen – die zich vaak op een voor robots moeilijk te voorspellen manier gedragen – aanwezig zijn.

In dit project hebben wij ons gericht op robots met het vermogen tot adaptief gedrag en het nemen van beslissingen in real-time op basis van AI-modellen, waaronder machine learning en deep learning.

Ook richten we ons op robots die emotionele signalen herkennen (zoals intonatie of gezichtsuitdrukkingen) en omgaan met nuances in menselijke spraak, zoals sarcasme of regionale uitdrukkingen.

De experts in de workshop gaven aan dat robots op dit moment vooral goed zijn in óf het uitvoeren van een gespecialiseerde, mechanische taak, zonder veel interactie met mensen, óf goed zijn in de interactie met mensen maar geen taken kunnen uitvoeren waarmee ze de fysieke wereld beïnvloeden. Voor de toepassingen die we hierboven benoemen moeten deze twee elementen worden gecombineerd. Volgens experts zal dat nog even (> 10 jaar) op zich laten wachten. Sociale robots die hun fysieke omgeving niet manipuleren kunnen op kortere termijn (< 5 jaar) worden verwacht. Experts zien dit eerder als 'ipads op wielen,' of virtuele avatars die eigenlijk geen fysieke behuizing nodig hebben om hun taken uit te voeren.

Wat zeggen experts?

Tijdens de expertsessie bespraken we twee toekomstige toepassingen van AI-gestuurde robots: een robotleraar en een robotpolitie. Voor beide toepassing bespraken we wat de belangrijkste (positieve en negatieve) gevolgen voor gebruikers zullen zijn.

In beide gevallen vonden experts het belangrijk dat de robots geen vervanging zijn van menselijke politieagenten of leraren, maar professionals ondersteunen. Op die manier nemen robots geen waardevol werk over van mensen, en worden veel risico's gemitigeerd.

Voor de **robotleraar** zagen de experts mogelijkheid om nieuwe manieren van

lesgeven te ontwikkelen, die beter aansluiten bij de individuele behoeften van kinderen. Aan de andere kant zagen ze het risico dat robots niet goed om zouden kunnen gaan met kinderen die anders zijn (zoals neurodiverse kinderen), en dat kinderen een band opbouwen met een machine. In het gesprek over de **robotpolitie** waren mogelijke gunstige effecten volgens experts een gevoel van veiligheid, anoniemere interacties (met robot in plaats van een persoon) en de zichtbaarheid van surveillance. Mogelijke nadelen zijn dat de voor aansturing gebruikte AI discrimineert, en een gevoel van overmatige surveillance bij burgers.

In het tweede deel van de sessie hebben we de blik verbreed en gekeken naar de gevolgen voor de samenleving van de technologie in het algemeen. Daarin bespraken we wat volgens deelnemers de grootste kansen en risico's voor publieke waarden zijn. De uitkomsten van dat gesprek zijn weergegeven op de volgende pagina's.

AI-gestuurde robots AI | Verkenning publieke waarden



Welvaart

AI-gestuurde robotica biedt de mogelijkheid om de **productiviteit** van werknemers te verhogen, door werk van ze over te nemen en met ze samen te werken. In sectoren met een groot **arbeidstekort**, zoals de zorg, horeca, bouw of retail, waar robots nu nog niet of nauwelijks worden ingezet maar in de toekomst mogelijk wel, kan dit helpen de tekorten te verminderen. Dat heeft een positief effect op **economische groei** en **welvaart**.

Een risico is dat dit vooral ten goede komt aan een selecte groep **kapitaalkrachtigen** (zoals grote bedrijven) die robots kunnen aanschaffen en inzetten om werknemers te vervangen. Deze bedrijven kunnen dan beslissen over welke werknemer worden vervangen door robot. Dat heeft effect op de **inkomsten** van de werknemers die worden vervangen. Maar mogelijk ook op de inkomsten van de staat, die geen **inkomsten-belasting** kan innen van robots.



Menselijke waardigheid

AI-gestuurde robots kunnen, **zware, gevaarlijke of vieze taken** overnemen, zoals het tillen van veel servies, werken in hitte of kou, opruimen of schoonmaken van (giftig) afval of uitwerpselen. Bovendien hebben robots geen moeite met **saaië, repetitieve taken** of lange werkdagen zonder pauze. Als robots dit soort taken overnemen kan het werk voor mensen **aangenamer** en **waardiger** worden. Maar robots kunnen ook worden ingezet om voor mensen **betekenisvol** werk over te nemen, zoals het werk van een leraar of een politieman. Als een robot deze taken kan uitvoeren kan dat als gevolg hebben dat die taken, en de mensen die ze uitvoeren, als **minder waardevol** worden gezien. Omdat een robot hetzelfde werk dan 'gratis' uitvoert. Daarnaast bestaat dan het risico dat er **eentonige banen** in de plaats komen voor hoogwaardige, betekenisvolle banen. Zoals het (op afstand) monitoren van de acties en beslissingen van zelfstandig opererende robots.



Toegankelijkheid en inclusiviteit

Het gebruik van AI in robotica biedt mogelijkheden om mensen te benaderen op een manier die beter bij zijn of haar vaardigheden of karakter passen. Bijvoorbeeld als robots worden gebruikt om te communiceren met klanten van supermarkten, horeca, of in sectoren als het onderwijs, de politie of de overheid. Dit biedt kansen om de **toegankelijkheid** en **inclusiviteit** van diensten te vergroten.

Tegelijkertijd is het mogelijk dat mensen die niet binnen de (voor de robot) bekende profielen passen (zoals neurodiverse personen) buiten de boot vallen.

AI-gestuurde robots AI | Verkenning publieke waarden



Privacy

Robots verzamelen in een sociale context via **sensoren** grote hoeveelheden (**gevoelige**) **data** over de mensen met wie zij in contact komen, en hun interactie met anderen. Vanwege de hoeveelheid gegevens, en de context waarin deze wordt verzameld, zijn deze **moeilijk te anonimiseren**. Bovendien bestaat het risico dat deze data wordt misbruikt als deze in verkeerde handen komen.



Verantwoordelijkheid en aansprakelijkheid

Het gebruik van robots voor het uitvoeren van taken in sociale contexten roepen vragen op over de verdeling van verantwoordelijkheden en aansprakelijkheden, zeker als ze samenwerken met mensen. Bijvoorbeeld, wie is er verantwoordelijk voor het voorkomen van fouten, en wie is aansprakelijk als er fouten worden gemaakt?

Het gevaar van onduidelijkheid rondom verantwoordelijkheid en aansprakelijkheid is dat in de praktijk niemand aansprakelijk kan worden gehouden als er iets fout gaat, of dat er niemand de verantwoordelijkheid pakt om ongewenste gebeurtenissen te voorkomen.



Duurzaamheid

De fabricage van robots heeft een grote **ecologische impact**. Denk bijvoorbeeld aan het delven van de materialen voor de mechanische onderdelen, de batterijen en de elektronica, en de (soms giftige) afval die ontstaat als deze worden weggegooid.

Bovendien vereist het **trainen** van de benodigde AI enorme rekenkracht waarvoor grote **datacentra** nodig zijn die veel **elektriciteit** en **water** verbruiken. Dit gebeurt buiten het zicht van gebruikers, waardoor die geen direct inzicht hebben in het verbruik.

Aandachtspunten

De waarde van werk

Uit de verkenning van de gevolgen van AI-gestuurde robotica voor publieke waarden blijkt dat veel gevolgen zich toespitsen op de context van werk. De inzet van robotica vraagt om een gesprek over de waarde werk voor individuen en de samenleving. Robotica biedt kansen om saaie, gevaarlijke, vieze of fysiek zware taken over te nemen van mensen, maar het risico bestaat dat ze ook waardevol of betekenisvol werk zullen overnemen. Dit roept vragen op over wat fysieke en mentale arbeid waard is en wat het betekent om dit te automatiseren. Voor wie is dat gunstig en voor wie juist niet? Wat brengt werk naast financiële genoegdoening?

Impact op sociale domeinen

De inzet van (sociale) robotica in publieke domeinen zoals het onderwijs, de politie of de zorg roept ook fundamentele vragen op. Het menselijke contact met burgers en patiënten, dat deze beroepen waardevol maakt, kan in zekere mate worden overgenomen door robots. Hierdoor zullen deze sectoren mogelijk meer gericht zijn op meetbare procedures, toetsbare resultaten en surveillance in plaats van menselijk contact. Dat kan op een fundamenteel niveau de manier waarop wij denken beïnvloeden over wat het betekent om te werken als onderwijzer, zorgverlener of politie. Tegelijkertijd vraagt het gebruik van

robotica in deze sectoren om ingrijpende veranderingen en een herverdeling van taken, bijvoorbeeld om het werk van de robot te controleren.

Wie ontwikkelt, beheert en bezit de technologie?

Uit de verkenning van de gevolgen van AI-gestuurde robots voor publieke waarden wordt duidelijk dat robots zowel positieve als negatieve gevolgen kan hebben. Of het gebruik van AI-gestuurde robots positief of negatief uitpakt voor een bepaalde waarde (zoals menselijke waardigheid, welvaart en toegankelijkheid) is sterk afhankelijk van hoe en waarvoor de technologie wordt ontwikkeld en hoe die wordt toegepast.

De maatschappelijke gevolgen van AI-gestuurde robots zijn dus sterk afhankelijk van de precieze toepassing, de toepassingscontext en de manier waarop de toepassingen worden geïmplementeerd. Het benutten van de kansen die AI-gestuurde robotica biedt vraagt om het maken van heldere keuzes over de ontwikkelingsrichting en de toepassingsgebieden.

Experts gaven aan dat de ontwikkeling, fabricage en het onderhoud van robots, en het trainen van de AI die deze aanstuurt vraagt om grote investeringen. Vooral grote, buitenlandse technologiebedrijven in staat om zulke investeringen te doen, wat veel risico's voor publieke

waarden met zich meebrengt. Dit verkleint de mogelijkheid om controle uit te oefenen op de ontwikkelingen, en de keuzes die worden gemaakt te operationaliseren. Onze experts gaven aan dat meer grip op de ontwikkeling van AI-gestuurde robotica ook de mogelijkheid biedt om andere soorten robots te voor te stellen. Bijvoorbeeld robots die lokaal door een scholenkoepel wordt ontwikkeld om te voldoen aan de behoefte van de eigen kinderen en leraren.

Conclusies

- AI-gestuurde robots bieden kansen om saai, gevaarlijk of vies werk over te nemen van mensen.
- Risico's ontstaan wanneer AI-gestuurde robots autonoom, zonder directe menselijke begeleiding, worden ingezet in sociale omgevingen.
- Vooral nog zijn vooral grote, niet-Europese technologiebedrijven in staat om investeringen te doen in de ontwikkeling van robots, wat risico's voor publieke waarden met zich meebrengt.
- Afhankelijk van de grip die we krijgen op de ontwikkeling kunnen AI-gestuurde robots in meer of mindere mate effect hebben op publieke waarden zoals welvaart, menselijke waardigheid, toegankelijkheid, privacy, aansprakelijkheid en duurzaamheid.

4.3 Reflectie – Overkoepelende lessen

De twee expertworkshops boden inzichten in de kansen en risico's voor publieke waarden van agentic AI en AI gestuurde robotica afzonderlijk. Maar we zagen ook terugkerende thema's in de discussies die in beide expertworkshops werden gevoerd. De terugkerende thema's wijzen op bredere lessen die gelden voor beide technologieën, en wellicht ook voor andere opkomende digitale technologieën. In deze reflectie trekken we overkoepelende lessen uit de thema's die in beide workshops terugkwamen.

De impact ligt nog niet vast

Agentic AI en AI-gestuurde robots zijn als technologie op dit moment nog niet uitontwikkeld of op grote schaal geïmplementeerd. Hoe de technologie er precies uit zal zien, en hoe deze zal worden gebruikt ligt dus ook nog niet vast. Experts in de workshops gaven aan dat er niet één manier is om agentic AI of AI-gestuurde robots te ontwikkelen. Zo kan een agentic AI systeem vooral vanuit commerciële belangen worden ontwikkeld en zich richten op langdurig gebruiksgemak en –geluk voor welvarende en digitaal geletterde personen. Maar experts konden zich ook goed voorstellen dat agentic AI wordt gebruikt om een systeem te ontwikkelen dat zich richt op de inclusie van groepen achtergestelde

personen. Hoe een AI systeem of een robot wordt ontwikkeld, en hoe en door wie deze vervolgens wordt gebruikt, bepaalt in grote mate welke impact de technologie zal hebben op de samenleving.

Omdat de doorontwikkeling en implementatie van de technologieën nog niet vast liggen, benoemden experts in beide workshop zowel kansen als risico's voor publieke waarden voor beide technologieën. Sommige experts waren meer optimistisch over de (toekomstige) maatschappelijke impact, terwijl anderen pessimistischer waren. Maar vrijwel allemaal waren ze het erover eens dat de uiteindelijke impact zal afhangen van de ontwikkeling, de toepassingscontext, de manier waarop de toepassingen worden geïmplementeerd en de controle die desamenleving hierop kan uitoefenen.

In deze constatering van de experts zat ook waarschuwing besloten. Op dit moment wordt de wereldwijde technologiemarkt gedomineerd door een klein aantal, buitenlandse technologie-reuzen met hoofdzakelijk commerciële belangen. Omdat de ontwikkeling van agentic AI en AI-gestuurde robots vragen om veel investeringen en grote hoeveelheden data, ligt het voor de hand dat zulke bedrijven ook in deze markt een grote rol zullen spelen. In die situatie zal er weinig grip zijn op de ontwikkeling en het gebruik, en zullen commerciële belangen belangrijker zijn dan publieke waarden.

Andere mogelijkheden voor toepassing en ontwikkeling

Experts wezen er tijdens de workshops op dat het niet vast hoeft te liggen dat agentic AI en AI-gestuurde robots vooral door grote buitenlandse bedrijven worden ontwikkeld. Het is ook mogelijk om andere – meer lokale en controleerbare – manieren van ontwikkelen en toepassingen te organiseren. Een voorbeeld dat experts gaven was het (laten) ontwikkelen van een eigen onderwijsrobot door een regionale scholenkoepel.

Op die manier kunnen scholen zelf kunnen bepalen welke taken de robot moet kunnen uitvoeren, en welke waarden (zoals privacy, transparantie en professionele autonomie van leraren) gewaarborgd dienen te worden. Zulke niet direct voor de hand liggende innovatiestrategieën bieden de mogelijkheid om technologie te ontwikkelen die is toegerust voor een specifieke context en rekening houdt met publieke waarden van de Nederlandse samenleving. Het vraagt om regie en investeringen om die ontwikkelingen te stimuleren die voor onze samenleving waardevol zijn. Het risico is dat commerciële big tech bedrijven, vanwege hun huidige voorsprong op de markt en innovatiekracht de markt zullen domineren als er géén keuzes worden gemaakt.

Keuzes en maatschappelijk gesprek

Agentic AI en (in grotere mate) AI-gestuurde robotica staan nog in de kinderschoenen, maar de ontwikkelingen gaan

snel. Het is nu nog mogelijk om risico's voor publieke waarden te verkleinen en kansen te vergroten. Maar dat vraagt om heldere (beleids)keuzes en een maatschappelijk gesprek. Voor welke toepassingen en in welke contexten willen we deze technologieën inzetten, en op welke manier? In welke gevallen kunnen we gebruik maken van systemen die worden aangeboden door buitenlandse technologie-reuzen en onder welke voorwaarden? Wanneer willen we de ontwikkeling en het beheer in eigen hand houden? Opkomende digitale technologieën stellen ons voor belangrijke keuzes. Om de ontwikkeling van agentic AI en AI-gestuurde robotica nog te kunnen sturen, is het belangrijk om deze keuzes op tijd te maken.

Bijlage I: Maatschappelijke impactdomeinen

Defensie en nationale veiligheid

verwijst naar het vermogen om territoriale integriteit en burgers te beschermen tegen externe bedreigingen en agressie.

Digitale infrastructuur

verwijst naar de kritieke technologie-systemen en -diensten die de werking en de connectiviteit van digitale omgevingen mogelijk maken.

Economie en ondernemingsklimaat

verwijst naar het vermogen om een gezonde economie en een economische en professionele omgeving rond een industrie of groep bedrijven in stand te houden.

Onderwijs (incl. wetenschap)

verwijst naar burgers die onderwijs volgen en wetenschappelijk onderzoek bevorderen.

Natuur en klimaat

verwijst naar het behoud van de natuur, omgaan met klimaatverandering.

Energie

verwijst naar het vermogen om te zorgen voor een betrouwbare, duurzame en betaalbare energievoorziening, zoals gas, elektriciteit, water en hernieuwbare energie.

Financiële infrastructuur

verwijst naar kritieke infrastructuur-systemen, instellingen en kaders die de effectieve werking van retail betalings-diensten en financiële markten mogelijk maken.

Voedselzekerheid

verwijst naar voedselproductie en toegang tot veilig en veilig geproduceerd voedsel in voldoende hoeveelheden

Justitie en rechtsstaat

verwijst naar de processen, instellingen, praktijken of normen die de gelijkheid van alle burgers voor de wet en democratie ondersteunen en het willekeurige gebruik van macht voorkomen.

Overheidsdiensten

verwijst naar effectieve levering van overheidsdiensten aan burgers.

Gezondheidszorg

verwijst naar het georganiseerd verlenen van medische zorg aan burgers.

Openbare ruimte en ruimtelijke ordening

verwijst naar het verzorgen van een voldoende aanbod van woningen op plaatsen waar het prettig wonen is en voor een goed gebruik van de beperkte ruimte van het land.

Sociale zaken en werkgelegenheid

verwijst naar de sociale zekerheid (uitkeringen, pensioenen, arbeidsomstandigheden, enz.) van de burgers, het beheren van de relaties tussen werkgevers en werknemers en de arbeidsmarkt

Telecommunicatie (met inbegrip van de ruimte)

verwijst naar de uitwisseling van informatie over aanzienlijke afstanden langs elektrische weg.

Waterbeheer, infrastructuur en transport

verwijst naar het creëren van een efficiënt netwerk van wegen, spoorwegen, waterwegen en luchtwegen, effectief waterbeheer tbv voorkomen van overstromingen en betere lucht- en waterkwaliteit

Bijlage II: Impact technologie per domein

- Om een beeld te krijgen van de reikwijdte van het maatschappelijke effect zijn 15 maatschappelijke domeinen gedefinieerd. Op verzoek van het ministerie zijn minderjarigen toegevoegd als doelgroep. Via een enquête is aan experts gevraagd om per technologie aan te geven op welke domeinen ze effect zullen hebben: “Selecteer de domeinen waarin de technologie(ën) waarschijnlijk een aanzienlijke maatschappelijke impact zullen hebben.” Technologieën werden geacht effect te hebben op een domein als het domein werd geselecteerd door ten minste 50% van de respondenten.
- Deze tabel geeft weer welke technologieën naar verwachting effect hebben op welke domeinen. De resultaten zijn daarbij omgezet naar een binaire weergave; Een 1 geeft aan dat ten minste 50% van de respondenten heeft aangegeven dat het domein effect zal ondervinden van de technologie. Bij een 0 waren dat minder dan 50% van de respondenten.

		Defensie en Nationale Veiligheid	Digitale infrastructuur	Economie en ondernemingsklimaat	Onderwijs (inclusief wetenschap)	Energie	Financiële infrastructuur	Voedselzekerheid	Justitie en rechtsstaat	Natuur en klimaat	Dienstverlening overheid	Gezondheidszorg	Openbare ruimte en ruimtelijke ordening	Sociale Zaken en Arbeid	Telecommunicatie (inclusief ruimte)	Waterbeheer, infrastructuur en transport	Minderjarigen	Totaal aantal domeinen
DT1	AI Agents	1	1	1	1	1	1	0	1	0	1	1	0	0	0	0	0	9
DT2	Bioinformatics	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
DT3	Cryptocurrency	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	2
DT4	AI-gestuurde robotica	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	4
DT5	Digital Twins	1	1	1	0	1	0	0	0	1	0	0	1	0	0	1	0	7
DT6	Digital Wallet	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	4
DT7	Neurological enhancement	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2
DT8	Multiparty computing	1	1	1	0	0	1	0	0	0	1	1	0	0	0	0	0	6
DT9	Neuromorphic Computing	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
DT10	Quantum Safe cryptography	1	1	0	0	0	1	0	1	0	1	0	0	0	1	0	0	6
DT11	6G	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2
DT12	Quantum algoritmen	1	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	6
DT13	Spatial ccomputing	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	4
DT14	Deep Fakes	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	3
Totaal technologieën per domein		9	9	7	2	3	6	0	3	2	5	5	2	0	2	2	0	

Bijlage III: Expertworkshop en deelnemende experts

De expertworkshop bestond in grote lijnen uit twee onderdelen: een bespreking van een use case in twee break-out groepen en een plenaire discussie.

Na een korte introductie van het project, de definitie en de scope van de technologie splitsen we de groep in tweeën, met een gelijke verdeling van TNO en niet-TNO experts en technisch-wetenschappelijke en sociaal-maatschappelijke experts.

1. In deel één bespreken de groepen elk een andere casus in de volgende stappen (zie de verschillende use cases in bijlage 8.2 van het methodedocument)
 - a. Individueel doorlezen en opschrijven van initiële reactie
 - b. Initiële reactie bespreken in de groep
 - c. Individueel (minimaal) drie negatieve en drie positieve impacts opschrijven
 - d. Impacts delen en bespreken
 - e. Impacts categoriseren en prioriteren

Na een korte pauze gaven de moderatoren van de groepen elk een samenvatting van de discussies. Daarna volgt een plenaire discussie waarin we de volgende vragen bespreken:

- a. Wat betekent dit voor de technologie als geheel?
- b. Welk effect het heeft (positief of negatief) op welke publieke waarden?

- c. Wat vindt je positief, waar hoop je op voor de toekomst?
- d. Waar maak je je zorgen over voor de toekomst?
- e. Zijn er groepen of personen die hier in het bijzonder door worden beïnvloed (positief of negatief)?
- f. Zijn er acties die ondernomen kunnen worden om je voor te bereiden – negatieve effecten te beperken of positieve effecten te versterken?

Agentic AI

Fictieve use cases

1. Agentic AI als vriend
2030: de 'BuddyBot', is een AI-agent die is gebouwd als een platonische vriend. Op basis van het profiel van de gebruiker kan de buddybot inchecken als deze bijvoorbeeld druk overkomt of die pro-actief afspraken kan inplannen.
2. Agentic AI bij de belastingdienst
In 2030 zet de Belastingdienst 'Sentinel' in, een geavanceerd AI-agentplatform, om proactieve interventies te leiden die burgers beschermen tegen een spiraal van financiële problemen. Menselijke maatschappelijk werkers ondersteunen de operaties van Sentinel, maar de besluitvorming en het zware werk berusten bij de AI.

Geraadpleegde experts

- Mark Neerincx | TNO | TU Delft
- Claartje Barkhof | TNO
- Damian Domela Nieuwenhuis Nyegaard | TNO
- Jesse van Oort | TNO
- Raquel Fernández | Universiteit van Amsterdam
- Carlos Zednik | TU Eindhoven
- Giulio Mecacci | Radboud Universiteit
- Sanaa Abrahams | Radboud Universiteit
- Pradeep Murukannaiah | TU Delft
- John Walker | SURF

AI gestuurde robotica

Fictieve use cases

1. Onderwijsrobot
 In 2030 vervangt De Zon School in Utrecht de leraar van groep 4 volledig door "Astra", een door AI ondersteunde humanoïde robot die autonoom lesgeeft, begeleidt en een klas van twintig 9- en 10-jarigen beheert.
2. Politierobot
 In 2030 introduceert de Rotterdamse Politiedienst "Guardian", een door AI ondersteunde robotische Gemeenschaps-wacht, om de openbare veiligheid te verbeteren en menselijke agenten vrij te maken voor complexe onderzoeken.

Geraadpleegde experts AI-gestuurde robots

- Cosimo Della Santina | TU Delft | German Aerospace Center
- Micah Prendergast | TU Delft
- Arnoud Visser | Universiteit van Amsterdam
- Jelle van Mil | TNO
- Serge Thill | Radboud Universiteit
- Ties van Bruinessen | TU Delft | RoboHouse
- Madelaine Ley | TU Delft
- John Walker | SURF
- Merel van Steenis | SURF | Universiteit van Amsterdam

Auteurs

Marianne Schoenmakers, Ron Oren, Sophie van Baalen,
Tom Barbereau, Vera Irmak

TNOvector
Centre for Societal Innovation and Strategy

5 September 2025

TNO 2025 R11919

tnovector.nl