

Artificial intelligence

Auteurs

Frank Benneker (Universiteit van Amsterdam), **Heleen van der Laan** (ROC van Amsterdam-Flevoland), **Natasha Alechina** (Open Universiteit), **Rob van de Star** (Windesheim), **Lars Veeffkind** (SURF), **Simone van Bruggen** (SURF), **Julius Ceasar Aguma** (SURF)

1. Meer gediversifieerde toegang tot grote AI-modellen

2. Veranderende dynamiek in responsible AI

3. Toename van co-evolutie tussen hardware en AI

4. Samenwerking tussen mensen en AI

5. Van grote tot kleine taalmodellen



Inleiding

De afgelopen jaren heeft de volledige realisatie van de zogenaamde transformer-architectuur (een digitaal neurale netwerk dat enorme hoeveelheden gegevens kan verwerken) het landschap van artificial intelligence (AI) volledig veranderd. De meeste AI-R&D richt zich tegenwoordig op Generatieve AI (GenAI), een subgebied dat computationele modellen gebruikt om tekst, afbeeldingen, video en audio te genereren.

In veel landen ontwikkelen door de overheid gefinancierde organisaties grote AI-modellen voor toepassingen in onderwijs, onderzoek en gezondheidszorg. Tegelijkertijd integreren steeds meer organisaties GenAI in hun producten en diensten voor commerciële doeleinden. Bedrijven zoeken naar real-world AI-toepassingen die onder andere efficiëntie bevorderen, terwijl ze tegelijkertijd streven naar een verantwoorde inzet van AI in die toepassingen. Zo

is er bijvoorbeeld een toename in het gebruik van AI-Concierge, waarbij AI gebruikers, zoals hotelgasten, 24 uur per dag kan assisteren met vragen via populaire messaging-apps.

Een cruciale belemmering voor de inzet van GenAI is dat zowel consumenten als veel medewerkers onvoldoende kennis hebben van het opstellen en optimaliseren van prompts.

Anders gezegd: een gebruiker moet weten welke instructies aan een GenAI-model moeten worden gegeven om het resultaat te verkrijgen. Agentic AI is hierdoor steeds prominenter geworden als oplossing voor gebruikers met beperkte AI-ervaring. Hierbij voert het AI-systeem autonoom vooraf bepaalde taken uit en is minimale menselijke supervisie nodig.

Grootschalige AI-modellen (die worden getraind met behulp van enorme datasets) worden steeds breder beschikbaar en toegan-

kelijk, en de toepassingen rond deze modellen vinden steeds vaker hun weg naar het dagelijks gebruik. Ook andere nieuwe vormen van mens-computerinteractie (human-computer interaction; HCI) op basis van AI en concepten zoals AI Love (waarbij mensen persoonlijke en emotionele banden met AI-systemen aangaan) krijgen steeds meer aandacht.

Een belangrijk gevolg van de vooruitgang op het gebied van AI is dat het nu een belangrijk onderwerp van geopolitieke belangstelling en controversen is. In wezen is AI een belangrijk instrument in de politieke wereld en vertegenwoordigt het nu de grootste asset in het geopolitieke spel om AI-suprematie. Overheden investeren meer in de ontwikkeling van AI-modellen en ondersteunende infrastructuur, omdat ze zich willen profileren als AI-supermachten om het volledige potentieel van AI voor de bredere samenleving en de nati-

onale veiligheid te realiseren. De race naar dominantie in AI lijkt helaas ten koste te gaan van verantwoorde en ethische AI-praktijken, waarbij het gaat om principes zoals trustworthiness (betrouwbaarheid), explainability (uitlegbaarheid) en human-centric AI (mensgericht ontwerp).

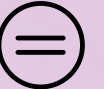
Bijdragers

Bas Smit (Erasmus Universiteit Rotterdam), **Duuk Baten** (SURF), **Bertine van Deyzen** (SURF), **Erna Sattler** (Universiteit Leiden), **Damian Podareanu** (SURF)

TREND #1

Meer gediversifieerde toegang tot grote AI-modellen

Publieke waarden

	Autonomie	Privacy Onafhankelijkheid
	Rechtvaardigheid	Transparantie Betrouwbaarheid Duurzaamheid
	Menselijkheid	Welzijn

Volwassenheid

WATCH

PLAN

ACT

Drivers

Individualisering en empowerment; Geestelijke gezondheid en welzijn; Belang van kennis en vaardigheden; Automatisering en AI; Connectiviteit en interactie

Sinds ChatGPT in 2022 beschikbaar en toegankelijk werd voor grootschalig gebruik, hebben grote taalmodellen (LLM's) wereldwijd enorm aan populariteit gewonnen. Zo bereikte ChatGPT binnen enkele dagen na de lancering al 1 miljoen gebruikers, wat leidde tot een meer gediversifieerde toegang tot LLM's en als katalysator fungeerde voor de toenemende integratie van AI in veel dagelijkse workflows. Toegang tot AI wordt vergemakkelijkt door browsergebaseerde tools (ChatGPT, Claude en DeepSeek) en door integraties in diensten en producten van leveranciers zoals Google (bijv. Docs) en Microsoft (bijv. Copilot, Outlook).

Bedrijfsmodellen variëren van freemium-niveaus tot pay-per-use API's. Gebruikers hebben nu de keuze tussen cloudgebaseerde AI-oplossingen of kleinere, lichtgewicht modellen die lokaal worden uitgevoerd. Tegelijkertijd hebben open-source-initiatieven bijgedragen aan de gevarieerde toegang tot en democratisering van LLM's



en tools. Deze ontwikkeling verandert de manier waarop we schrijven, coderen, onderzoek doen en omgaan met AI, en zelfs met elkaar.

SIGNALEN

Aantal GPT-gebruikers

De meeste studenten gebruiken ten minste één van de beschikbare platforms om toegang te krijgen tot een foundation-model.

- hepi.ac.uk 
- campustechnology.com 
- insidehighered.com 
- bibliotheek.educause.edu 

OpenAI is momenteel het populairst. Andere opties zijn Claude van Anthropic, Meta, Google-opties zoals LM Notebook & Gemini (datastudios.org) 

Eind 2024-begin 2025 heeft DeepSeek ondanks veiligheidsrisico's behoorlijk wat aandacht gekregen (dig.watch) 

Volgens Semrush was OpenAI in april 2025 de nummer vijf website (semrush.com) 

ChatGPT is niet de enige chatbot die gebruikers wint (techcrunch.com) 

Gartner: waarom taakspecifieke AI-modellen LLM's zullen vervangen.

(aimagazine.com) 

Snelle opkomst van generatieve AI: de toonaangevende bedrijven. (iot-analytics.com)

Groeiend aantal beschikbare modellen (& aanbieders) op Huggingface. (huggingface.co)

Vaticaan: nieuw document van het Vaticaan onderzoekt mogelijkheden en risico's van AI. (vaticannews.va)

Europese AI in opkomst

- openeurollm.eu 
- swiss-ai.org 
- gpt-nl.nl 
- galaxus.nl 
- swiss-ai.org 

Verliest de zoekmachine terrein ten opzichte van AI-chatbots?

- economictimes.indiatimes.com 
- mashable.com 

IMPACT



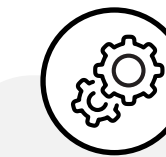
Onderwijs

- De gediversifieerde toegang tot AI biedt zowel kansen als uitdagingen voor docenten door routinematige taken te automatiseren, waardoor zij zich kunnen concentreren op lesgeven en persoonlijke begeleiding, maar tegelijkertijd moeten zij zich aanpassen aan een door AI gedreven onderwijslandschap.
- Naarmate studenten steeds vaker gebruikmaken van AI-chatbots voor leren, productiviteit en algemene cognitieve ontlasting, kan dit aanzienlijke gevolgen hebben voor traditionele begeleidingsrollen en menselijke vaardigheden zoals kritisch denken. Onder studenten kan de digitale kloof groter worden door verschillen in kennis, toegang en toestemming om digitale middelen te gebruiken.



Onderzoek

- In wetenschappelijk onderzoek kan AI de analyse van gegevens versnellen, de interpretatie van teksten verbeteren en het genereren van hypothesen ondersteunen. Dit draagt bij aan een hogere snelheid en grotere diepgang van wetenschappelijk onderzoek.
- Uitdagingen zijn onder meer zorgen over auteursrechten en intellectueel eigendom, het risico van desinformatie in gegenereerde content en het gevaar dat onderzoekers te veel gaan vertrouwen op AI-tools. Er worden meer voorstellen en papers geproduceerd, wat de beoordelingslast vergroot.
- LLM's worden steeds vaker ingezet als onderzoeks-partner (research agents/mede-wetenschappers).



Operations

- Institutionele processen worden vereenvoudigd door routinematige of repetitieve taken te automatiseren, informatiebeheer te verbeteren en slimmere workflows mogelijk te maken.
- Instellingen moeten omgaan met uitdagingen zoals mogelijke problemen met auteursrechten en eigendom van gegevens, het risico van verspreiding van desinformatie, de vermindering van menselijk toezicht en bezorgdheid over de ecologische impact van grootschalige AI-implementaties.
- Diensten en producten die AI mogelijk maken, kunnen leiden tot onbevoegde gegevensverwerking, contractbreuk en vereiste data protection impact assessments (DPIA's).

TREND #2

Veranderende dynamiek in responsible AI

Publieke waarden

	Autonomie	Privacy
	Rechtvaardigheid	Integriteit Gelijkwaardigheid Inclusie
	Menselijkheid	Veiligheid

Volwassenheid

WATCH

PLAN

ACT

Drivers

Geopolitiek en (digitale) soevereiniteit; Concentratie van rijkdom en economische ongelijkheid; Regelgeving en naleving (compliance)

De EU wordt gezien als de wereldwijde leider op het gebied van regulering in vergelijking met de VS. De AI Act, het eerste regelgevingskader ter wereld dat is gebaseerd op risico's, transparantie en rechten, toont aan dat de EU de verdediger is van responsible AI.

Ondanks deze baanbrekende wetgeving hebben recente geopolitieke gebeurtenissen – voornamelijk in de VS – de aandacht verschoven van ethiek en mensgerichte AI naar nationale veiligheid en economische concurrentie. Grote technologiebedrijven schroeven hun toezeggingen op het gebied van responsible AI terug door hun ethische commissies te ontbinden en concurrentievermogen voorrang te geven boven transparantie en verantwoorde implementatie.

Er zijn echter signalen dat veel bedrijven over de hele wereld wel degelijk verantwoordelijke AI willen implementeren, maar dat ze daarvoor ondersteuning of mechanismen missen. Hoewel de schaal over het algemeen klein is, wordt er vooruitgang

geboekt bij het opzetten van programma's voor responsible AI. Toch blijft er een kloof bestaan tussen intentie, woorden en daden op het gebied van responsible AI.



SIGNALEN

Bedrijven zoals Google en anderen ontbinden hun ethische commissies en ontslaan werknemers die betrokken zijn bij ethische vraagstukken.

- [wired.com](#) 
- [startupdaily.net](#) 
- [wired.com](#) 

Denk twee keer na voordat u DeepSeek gebruikt: uitleg over beveiligings- en vertrouwenskwesties.

([carleton.edu](#)) 

Het Verenigd Koninkrijk heeft plannen om AI te reguleren uitgesteld, omdat ministers zich willen aansluiten bij de regering-Trump.

([theguardian.com](#)) 

Het voorstel van de Europese Commissie voor de AI Liability Act, waarmee het gebruik van auteursrechtelijk beschermd materiaal voor AI-trainingen onder de auteurswet zou vallen, is door de Commissie ingetrokken.

- [openai.com](#) 
- [mayerbrown.com](#) 
- [musicbusinessworldwide.com](#) 

Onderzoek van Harvard Business Review (HBR): hoe responsible AI de winst beschermt. ([hbr.org](#)) 

Iets meer dan de helft (52%) van de bedrijven heeft daadwerkelijk een programma voor responsible AI, blijkt uit gegevens van BCG. ([bcg.com](#)) 

87% van de managers erkent het belang van responsible AI (RAI) - MIT Technology Review. ([technologiereview.com](#)) 

“Er zijn steeds meer aanwijzingen dat veel grote technologiebedrijven terugkomen op hun toezeggingen om verantwoord om te gaan met AI.”

- Virginia Dignum, hoogleraar AI en directeur van het AI Policy Lab, Universiteit van Umeå

IMPACT



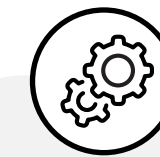
Onderwijs

- De auteursrechtelijke bescherming van lesmateriaal wordt uitgehold en er is behoefte aan een effectievere manier om studenten te leren hoe ze op een verantwoorde en veilige manier met GenAI kunnen omgaan.
- Studenten moeten leren om niet blindelings te vertrouwen op GenAI-modellen, en zoals Virginia Dignum zegt, moeten ze worden opgeleid om zich te bewegen in een bedrijfscultuur die steeds ambivalenter staat tegenover responsible AI.



Onderzoek

- Auteursrecht kwesties hebben invloed op onderzoek met GenAI-modellen. Onderzoekers hebben toegang nodig tot onafhankelijk ontwikkelde modellen, in plaats van alleen modellen die zijn afgeleid van grote technologiebedrijven, om juridische complicaties te voorkomen en verantwoorde innovatie te bevorderen.
- Industriële samenwerkingsverbanden kunnen worden beïnvloed door uiteenlopend beleid op het gebied van gegevensbeheer en -verwerking.
- Onafhankelijk onderzoek naar de impact van AI op de samenleving moet prioriteit krijgen.



Operations

- De beoordelingsmethoden moeten opnieuw worden bekeken omdat OpenAI geen maatregelen heeft genomen om fraude bij cursussen te voorkomen.
- Er zijn duidelijke richtlijnen nodig voor studenten en medewerkers over het gebruik van GenAI.
- Offline modellen voor educatief gebruik moeten aan eindgebruikers worden verstrekt.
- Voorzichtigheid is geboden bij het gebruik van GenAI voor administratieve taken, met name wat betreft gegevensbescherming en ethische kwesties.

TREND #3

Toename van co-evolutie tussen hardware en AI

Publieke waarden

 Autonomie Onafhankelijkheid

 Rechtvaardigheid Gelijkwaardigheid

 Menselijkheid

Volwassenheid

WATCH

PLAN

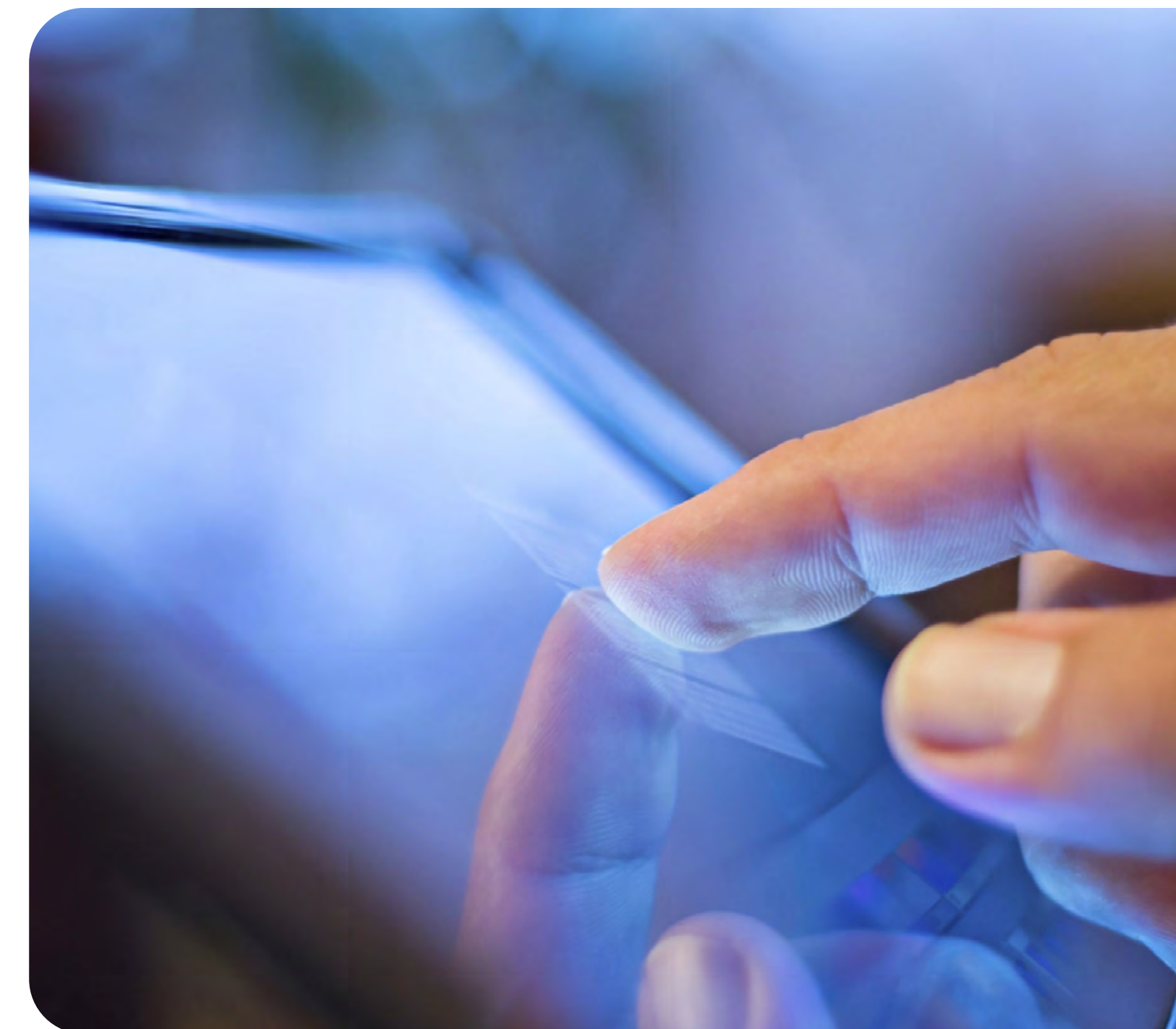
ACT

Drivers

Automatisering en AI; Technologische vooruitgang en computation; Energievoorziening en energievraag; Schaarste aan grondstoffen

Hoewel hardware aanvankelijk de berekeningen voor AI vergemakkelijkte, wordt het nu een beperkende factor, waardoor AI-ontwikkelingen moeten worden aangepast aan de computerinfrastructuur. De groeiende populariteit en toegankelijkheid van AI-modellen zorgen voor een aanzienlijke verschuiving in hardware-innovaties. AI-modellen worden sneller groter en complexer dan de vooruitgang in de in te zetten general-purpose computertechnologieën.

In het vorige SURF Tech Trends-rapport 2023 werd de opkomst van gespecialiseerde AI-hardware belicht. Naast deze trend worden er steeds meer apparaten (telefoons, laptops, enz.) geproduceerd om dagelijkse AI-workflows te ondersteunen. Een andere ontwikkeling, voortkomend uit het verschil tussen AI-complexiteit en beschikbare rekenkracht, is de ontwikkeling van AI-modellen die optimaal gebruikmaken van de beschikbare computationele middelen. Met innovaties zoals modellen met



lage precisie en AI-bewuste hardware-implementaties wordt AI opnieuw vormgegeven, gedreven door de noodzaak om 'meer met minder' te bereiken op het gebied van systeemcapaciteiten.

SIGNALEN

De aankondigingen van Nvidia op CES 2025 beloven een persoonlijke supercomputer.

(nvidianews.nvidia.com) [↗](#)

Verwachte groei van de markt voor AI-chips. (gcom.pdo.aws.gartner.com) [↗](#)

Er komen meer bedrijven op de markt die zich richten op AI-hardware, bijvoorbeeld Cerebras.

- forbes.com [↗](#)
- axelera.ai [↗](#)

Veel bestaande bedrijven richten zich nu op het maken van AI-hardware.

- digitaltrends.com [↗](#)
- gartner.com [↗](#)

Er is een toename in het ontwerpen van hardwarebewuste modellen. (pytorch.org) [↗](#)

In de marketing van de meeste hardware ligt de focus steeds sterker op AI-gebaseerde oplossingen. (gartner.com) [↗](#)

De AI-revolutie zal hardwarebronnen met grote energiebehoeften vereisen. Dit maakt enterprise-infrastructuur opnieuw tot een strategische onderscheidende factor.

(deloitte.com) [↗](#)

“De verbeteringen in AI-trainingsprestaties sinds de MLPerf-benchmarks hebben de wet van Moore ruimschoots overtroffen.”

- David Kanter, uitvoerend directeur van MLCommons
(deloitte.com)

IMPACT



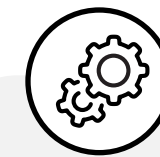
Onderwijs

- Persoonlijke AI-werkstations en meer aandacht voor betrouwbare computerarchitecturen kunnen helpen om de huidige privacyproblemen met grote technologische LLM's te verminderen en het dataverkeer over netwerken te verminderen.
- High-performance computing (HPC) voor AI biedt studenten de mogelijkheid om complexere data taken uit te voeren, maar hiervoor kan training nodig zijn.



Onderzoek

- Toegankelijke computerhardware met betere prestaties (persoonlijke supercomputers) betekent dat de onderzoeksgemeenschap over betere onderzoeksmogelijkheden zal beschikken.
- De afhankelijkheid van een handvol leveranciers voor AI-hardware brengt het risico met zich mee dat er een tekort aan AI-apparaten ontstaat als de wereldwijde toeleveringsketens onder druk komen te staan.



Operations

- Persoonlijke AI-werkstations en efficiëntere AI-modellen tegen lagere kosten betekenen dat studenten en onderzoekers minder snel behoefte hebben aan grootschalige high-performance computing (HPC)-systemen om te experimenteren met geavanceerde AI. Dit in tegenstelling tot de computerwetenschap, datawetenschap en AI-gemeenschappen, die nog steeds internationaal concurrerend onderzoek moeten uitvoeren.
- De wereldwijde vraag naar grondstoffen voor AI-specifieke chips zou de beschikbaarheid van hardware voor edge-apparaten die slimme campusautomatisering, onderzoek en onderwijs ondersteunen, kunnen beperken.

TREND #4

Samenwerking tussen mensen en AI

Publieke waarden

	Autonomie	Keuzevrijheid Onafhankelijkheid Privacy Diversiteit
	Rechtvaardigheid	Inclusie
	Menselijkheid	Veiligheid Sociale samenhang

Volwassenheid

WATCH

PLAN

ACT

Drivers

Individualisering en empowerment; Belang van kennis en vaardigheden; Automatisering en AI; Connectiviteit en interactie; Digitale transformatie; Concentratie van rijkdom en economische ongelijkheid

In plaats van banen over te nemen, zoals AI vaak wordt voorgesteld, wordt AI steeds meer een partner op het werk en in ons dagelijks leven. In bepaalde opzichten neemt AI nu al routinetaken over, zodat mensen zich kunnen concentreren op complex en creatief werk.

AI wordt steeds vaker ingezet naast routinetaken, als ondersteuner bij therapie, persoonlijke zingeving en het aanpakken van levensproblemen. De opkomst van multi-agent systemen zal deze rol verder versterken.

Agentic AI, dat sinds eind jaren negentig in ontwikkeling is, belooft de interactie tussen mens en computer aanzienlijk te verbeteren door middel van natuurlijke taalinterfaces. Daarnaast ontstaan er andere interfaces in collaboratieve robots (cobots) en humanoïde robots, waardoor meer mensachtige interacties mogelijk worden.

Bovendien wordt AI als omgevingsintelligentie geïntegreerd in verschillende tools en apparaten (ambient intelligence). Denk bijvoorbeeld aan wearables met camera's en microfoons die de



omgeving van de drager kunnen analyseren en ondersteuning bieden op een vergelijkbare manier als een digitale assistent. Binnen deze ontwikkeling is er zichtbare bezorgdheid over de afhankelijkheid van LLM's van grote technologiebedrijven die de onderliggende toepassingen mogelijk maken.

SIGNALEN

Gezondheidszorg

AI-assistenten zoals PathAI helpen artsen bij het analyseren van medische beelden (pathai.com) [↗](#)

Humanoïde robots

Zullen de komende drie jaar in de VS en China in massa worden geproduceerd.

- tesla.com [↗](#)
- unitree.com [↗](#)

Rapport: De volgende grote gok van OpenAI zal geen wearable zijn (techcrunch.com) [↗](#)

Softwareontwikkeling

Tools zoals Github copilot in Character.AI zijn bedoeld om de productiviteit en kwaliteit in de levenscyclus van softwareontwikkeling te verhogen.

- github.com [↗](#)
- character.ai [↗](#)

Je volgende AI-wearable zal altijd naar alles luisteren (wired.com) [↗](#)

Project Astra | De toekomst van leren verkennen met een prototype voor AI-tutoren (youtube.com) [↗](#)

Er komt krachtigere AI aan. De academische wereld en het bedrijfsleven moeten hier samen toezicht op houden (doi.org) [↗](#)

“Zodra AI eenmaal deel uitmaakt van de achtergrond, is het niet langer optioneel. Het is onzichtbaar, maar onomkeerbaar.”

- Pieter Loman, Universiteit Utrecht

IMPACT



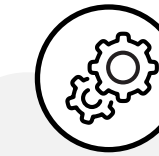
Onderwijs

- Er zijn veranderingen nodig in het curriculum om tegemoet te komen aan de nieuwe manieren van samenwerken met AI in beroepen, zonder daarbij afbreuk te doen aan publieke waarden.
- Agentic AI kan de werkdruk van docenten verminderen door beoordelingen, feedback en lesplanning te automatiseren, maar roept vragen op over de zeggenschap van studenten, vooringenomenheid in feedback en zorgen over toezicht.
- Aangezien samenwerking tussen mensen en AI op de werkplek een grote impact heeft op banen, zullen studenten aanpassingsvermogen en een andere mindset nodig hebben als cruciale vaardigheden voor de toekomst.
- Studenten en docenten worden een bron van gegevens voor grote technologiebedrijven, nu AI-tools steeds breder beschikbaar en toegankelijker worden.



Onderzoek

- Onderzoekers kunnen profiteren van AI-agenten voor het uitvoeren van repetitieve taken (bijvoorbeeld opmaak, samenvatting en literatuuronderzoek).
- Het waarborgen van reproduceerbaarheid, gegevensprivacy en auteurschapstoekenning is voor AI-agenten complexer om te behandelen.
- Door gegevensanalyse te stroomlijnen, experimenteel ontwerp te ondersteunen en laboratoriumtaken te beheren met behulp van robotica en voorspellende systemen, belooft AI meer automatisering en efficiëntie.
- Het eigendom van onderzoeksgegevens en daarmee samenhangende vragen worden blootgelegd door het gebruik van openbaar beschikbare LLM's.



Operations

- AI kan de efficiëntie en productiviteit verhogen, maar er moet voorzichtig worden omgegaan met ethische afwegingen. Belangrijk is dat het hoofddoel van onderwijs niet kostenverlaging is, maar het creëren van een zinvolle leeromgeving voor studenten en een veilige en inspirerende werkplek voor onderwijsprofessionals.
- Instellingen lopen het risico van vendor lock-in en stijgende IT-kosten vanwege hun afhankelijkheid van ingebouwde AI-tools.
- AI-gestuurde diensten kunnen de planning, studentenondersteuning en administratie optimaliseren, waardoor de operationele kosten dalen.
- Mens-AI-samenwerkingssystemen zullen worden gevoed met enorme hoeveelheden fysieke en persoonlijke gegevens, waardoor strenge privacyregels en ethische overwegingen nodig zijn om gebruikers te beschermen en verantwoord AI-gebruik te bevorderen.

TREND #5

Van grote tot kleine taalmodellen

Publieke waarden

	Autonomie	Privacy Diversiteit
	Rechtvaardigheid	Transparantie Inclusie Verantwoordelijkheid
	Menselijkheid	Veiligheid Sociale samenhang

Volwassenheid

WATCH

PLAN

ACT

Drivers

Individualisering en empowerment; Belang van kennis en vaardigheden; Connectiviteit en interactie; Cybersecurity en vertrouwen; Concentratie van rijkdom en economische ongelijkheid; Energievoorziening en energievraag; Klimaatverandering en opwarming van de aarde

Na de revolutionaire introductie van LLM's is er een groeiende belangstelling voor small language models (SLM's). Dit zijn modellen met maximaal 10 miljard parameters, in tegenstelling tot LLM's, die honderden miljarden of zelfs triljoenen parameters kunnen hebben.

Het trainen en gebruiken van LLM's vereist enorme hoeveelheden rekenkracht. SLM's zijn daarentegen aanzienlijk kleiner. Daardoor zijn ze minder veeleisend wat betreft dataverwerking, hardware en trainingstijd. SLM's verbruiken ook minder energie, waardoor ze geschikter zijn voor toepassingen op kleinere apparaten.

SLM's zijn toegankelijker voor gebruikers die deze modellen willen trainen en uitvoeren op consumentenhardware aan de rand van een netwerk, met name voor apparaten met één doel (bijvoorbeeld sensoren). Bovendien zijn SLM's vooral nuttig voor specifieke taken en minder geschikt als algemene hulpmiddelen.



SIGNALEN

Implementatie op het apparaat

Gemini Nano op Android: bouwen met GenAI op het apparaat (io.google) [↗](#)

Qualcomm maakt het ook mogelijk om modellen zoals Meta's Llama 2 op smartphones te draaien met behulp van geoptimaliseerde, gekwantiseerde versies (qualcomm.com) [↗](#)

Gespecialiseerde hardware

Europese tech-ecosystemen richten zich op de ontwikkeling van geavanceerde AI-chips die zijn ontworpen om AI-kracht van serverkwaliteit te leveren op compacte hardware.

- axelera.ai [↗](#)
- edge-ai-tech.eu [↗](#)

Industriële integratie

Door de EU gefinancierde onderzoeksprogramma's (zoals AI-PRISM) onderzoeken de integratie van AI in de productie- en industriële sectoren om realtime analyses mogelijk te maken (aiprism.eu) [↗](#)

- AI-aangedreven mensgerichte robotinteracties voor slimme productie (cordis.europa.eu) [↗](#)
 - Sociale industriële samenwerkingsomgevingen waarin AI, big data en robotica worden geïntegreerd voor slimme productie (cordis.europa.eu) [↗](#)
-

Oproep tot milieuvriendelijkere alternatieven (nos.nl) [↗](#)

SLM's staan bekend om hun lagere energieverbruik en bieden een emissiearme optie voor instellingen die hun CO₂-voetafdruk willen minimaliseren (unesco.org) [↗](#)

Waarom zijn onderzoekers geïnteresseerd in kleine taalmodellen? (quantamagazine.org) [↗](#)

Het transformatieve potentieel van AI hangt af van energie (iea.org) [↗](#)

Toepassingen die de privacy beschermen

AIPI-proefproject gestart in vijf Nederlandse ziekenhuizen (radiobotics.com) [↗](#)

IMPACT



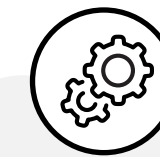
Onderwijs

- SLM's en edge AI kunnen gepersonaliseerde leerervaringen rechtstreeks op de apparaten van studenten verbeteren, waarbij privacy en toegankelijkheid worden gewaarborgd. Instellingen moeten echter infrastructuurupgrades beheren en gelijke toegang handhaven om technologische achterstanden te voorkomen. Er zullen nieuwe vaardigheden op het gebied van modelcompressie en hardware-optimalisatie nodig zijn.
- De betaalbaarheid van SLM-apparaten en ondersteunende technologie kan de digitale kloof verkleinen.



Onderzoek

- Onderzoekers kunnen profiteren van lokale, realtime analyses, zonder dat ze daarvoor uitgebreide rekenkracht nodig hebben. Dit vergemakkelijkt onderzoek in omgevingen met beperkte middelen.
- Lokale apparaten die als sensoren fungeren, zullen gegevens direct kunnen verwerken, waardoor er meer mogelijkheden ontstaan voor locatieonafhankelijk onderzoek.



Operations

- Instellingen kunnen SLM's en edge AI inzetten om de operationele efficiëntie te verbeteren, bijvoorbeeld door administratieve taken te automatiseren, terwijl ze minder afhankelijk worden van externe clouddiensten en mogelijk kosten kunnen besparen. Er zijn echter investeringen nodig in infrastructuur op apparaatniveau en in de ontwikkeling van vaardigheden om vendor lock-in te voorkomen en flexibiliteit op lange termijn te behouden.
- Instellingen moeten proactief plannen maken voor noodzakelijke infrastructuurupgrades en trainingsprogramma's om effectief gebruik te kunnen maken van de voordelen van edge AI en SLM's.

SURF Utrecht

Kantoren Hoog Overborch
Hoog Catharijne
Moreelsepark 48
3511 EP Utrecht
+31 88 787 30 00

SURF Amsterdam

Science Park 140
1098 XG Amsterdam
+31 88 787 30 00

futuring@surf.nl
www.surf.nl