

RUUD HEESEN

Tovenaarsleerling 2.0: Van bezem tot bot. Een ‘eigen leven’ van AI?

Goethes *Zauberlehrling*¹ spreekt tot onze verbeelding in het AI-tijdperk. We herkennen het patroon: wat wij creëren om ons te dienen, doet iets al snel meer en beter dan wij het beogen. De bezem die water draagt, gehoorzaamt té goed, vermenigvuldigt zich en verandert het tovenaarshuis tot een doorstroomruimte. De bezem leidt een ‘eigen leven’ en wordt de logica die de wereld rondom zich vormt. De tovenaars redt de situatie.

Een ballade als de Tovenaarsleerling en films als *Her*, *Ex Machina* en *Blade Runner 2049* zijn (nog) fictie. Kun je Ava uit de film *Ex Machina* straks in de straat tegenkomen? Wat als een robot² van *Boston Dynamics*³ inderdaad een ‘eigen leven’ gaat leiden?⁴ Dan mag je hopen dat iemand (of iets?) de stekker eruit trekt! We krijgen een onrustig gevoel: wat als *de mens hiermee iets creëert wat een ‘eigen leven’ gaat leiden* en wat als wij, mensen hierover de controle verliezen? Wat is dit creëren en wat krijgt dan een eigen leven en wat is dat ‘eigen leven’?

Dit essay focust op onze ervaring met AI⁵. Kan AI, in vrijheid gecreëerd door de mens, inderdaad een ‘eigen leven’ gaan leiden; een eigen leven zoals we dat ervaren als een eigen leven van mensen om ons heen? Hoe kun je deze ervaring met AI duiden en verbinden met een archetypisch verhaal als de Tovenaarsleerling? Verhalen over Golem, Frankenstein en de Tovenaarsleerling, waarin ‘gespeeld’ wordt met scheppingskracht wakkeren de verbeelding aan dat AI een ‘eigen leven’ zou kunnen gaan leiden. Op dit moment hebben we geen aanleiding te geloven dat dat werkelijk gaat gebeuren. De ingrediënten van deze verbeelding

1 J.W. Goethe, *Der Zauberlehrling*, <https://www.gedichte7.de/der-zauberlehrling.html> (geraadpleegd op 28 augustus 2025)

2 De ‘bot’ uit de titel is een afkorting van robot en als de bezem uit Tovenaarsleerling: zodra je hem een taak geeft, voert hij die razendsnel uit volgens de hem toevertrouwde taal.

3 Vgl. <https://bostondynamics.com/> (geraadpleegd op 24 augustus)

4 V.C. Müller, Stanford Encyclopedia of Philosophy, 2020: <https://plato.stanford.edu/entries/ethics-ai/> (geraadpleegd op 24 augustus)

5 Als we in dit essay AI gebruiken, bedoelen we altijd generatieve AI. Dat wil zeggen AI die iets ‘nieuws’ creëert.

willen we onderzoeken⁶. Hoe kan AI geduid worden vanuit een objectief derdepersoons-perspectief⁷ en welke mensbeelden passen daarbij? Wat is onze eerste-persoons-ervaring met AI? Heideggers essay “De vraag naar de techniek” geeft ons een vocabulaire om de verbeelding van het eigen leven van AI in te bedden.

Kortom, wat we proberen aannemelijk te maken, is dat AI geen eigen leven heeft in biologische zin, maar wél zo kan *verschijnen* – fenomenologisch en sociaal – in onze ervaring waardoor een Heideggeriaanse houding van gelaten, maar verantwoordelijke omgang vereist is, passend bij het mensbeeld dat betekenis geeft aan het eigen leven van individueel leven.

AI met een eigen leven?

Wat is generatieve artificiële intelligentie, hoe is het zo ver gekomen en hoe uit het zich? Artificiële intelligentie begint met pioniers als Alan Turing, John McCarthy en Herbert Simon na de Tweede Wereldoorlog. Turing stelde in een beroemd artikel de vraag of *machines kunnen denken*⁸. Hij ontwierp als criterium de naar hem genoemde Turingtest: een machine vertoont intelligent gedrag als een menselijke gesprekspartner dit machinaal gedrag in een geschreven gesprek niet kan onderscheiden van een menselijk intelligent gedrag.

De eerste decennia van AI waren vooral gericht op symbolische representatie: kennis werd voorgesteld in computertaal en logische structuren, die de computer moest manipuleren. Deze zogenaamde *good old-fashioned AI* (GOFAI) werkte goed voor gestructureerde problemen, maar stuitte al snel op de grenzen van de werkelijkheid, waarin context, ambiguïteit en lichamelijke cruciaal zijn⁹. Vanaf de jaren tachtig en negentig verschoof de aandacht naar neurale netwerken en *machine*

6 Dit essay leest de klassieke ballade *Der Zauberlehrling* van Goethe als moderne parabel voor onze omgang met AI en is geïnspireerd door ervaringen met ChatGPT in de onderwijspraktijk en ook wat AI losmaakt in het arbeidsveld, door het beluisteren van podcasts zoals AI-Report, door <https://www.3blue1brown.com/>, door <https://aivoordocenten.nl/> (geraadpleegd op 23 augustus 2025), door ons onderzoek samen met Ruurd Taconis: “*ChatGPT, Finally, a Colleague Who Understands Me*” (in druk: *Generative AI in Novel Educational Applications: Practices of Integrating Generative AI in the Primary and Secondary Classroom* (M.E. Auer, S.A. Nikou editors), door het actuele vwo-eindexamenonderwerp voor het vak filosofie: de vraag naar de mens in relatie tot techniek en wetenschap en het door K. Poortier, E. Myin en P.-P. Verbeek daarbij geschreven boek: *Wat maakt de mens. Onze lichamelijke in relatie tot techniek en wetenschap*, Amsterdam 2022).

7 Eerste-, tweede-, derdepersoons-perspectief zijn begrippen uit de fenomenologie en komen ter sprake in het vervolg.

8 A.M. Turing, *Computing machinery and intelligence*, *Mind*, 59 (1950), 433-460.

9 H.L. Dreyfus, *What computers can't do: A critique of artificial reason*. New York 1972.

learning: niet langer werd intelligentie expliciet gecodeerd, maar geleerd uit patronen in data¹⁰. De laatste jaren zijn generatieve modellen van OpenAI, Anthropic, Google of xAI doorgebroken: systemen die niet enkel classificeren, maar ook nieuwe teksten, bewegende beelden en codes genereren vanuit verschillende onderliggende 'architecturen'¹¹. Je kunt AI ook taken laten uitvoeren of je computer laten overnemen in het doen van taken. Daarmee krijgt AI een vorm van *agency*.

De ontwikkelingen van AI gaan enorm snel. Ze zijn te volgen op websites van de aanbieders, specialisten, influencers en platforms zoals Spotify en YouTube. Er zijn nog maar weinig domeinen waar niet over de invloed van AI wordt gesproken: van ambigue grensvallen tot de Nobelprijs. Sinds de lancering van ChatGPT voor een breed publiek in 2022 zijn mensen het gaan gebruiken in hun dagelijkse praktijk voor eenvoudige vragen, gedichtjes, slogans, plaatjes, planningen, filmpjes, computercodes, huiswerk, et cetera. Afhankelijk van de kwaliteit van de antwoorden van de AI zijn mensen tevreden of niet over AI.

Voorbeelden in het maatschappelijke domein zijn (niet per se generatieve) algoritmes die sociale media sturen en zo publieke opinie en democratische processen beïnvloeden of algoritmes die op de beurs onverwachte kettingreacties veroorzaken of generatieve modellen die originele maar ook misleidende beelden of films (*deepfakes*) produceren¹². Deze systemen 'leven' niet biologisch, maar zij leiden in verbinding met de wereld een 'sociaal leven' door het stroomlijnen van communicatie en *workflows*¹³.

In het natuurwetenschappelijke domein kun je denken aan de xenobots: levende cellulaire robots die onder AI-sturing vormen aannemen

10 B.J. Copeland, *Connectionism* (2025): <https://www.britannica.com/science/history-of-artificial-intelligence/Connectionism> (geraadpleegd op 23 augustus 2025).

11 Voor de GPT-architectuur zie bijvoorbeeld A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser & I. Polosukhin. *Attention is all you need*. In: *Advances in Neural Information Processing Systems* 30 (2017). zie: https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf geraadpleegd op 23 augustus 2025. GPT staat voor Generative Pre-Trained Transformer. Voor de 'denkarchitectuur': J.E. Laird. *The Soar cognitive architecture*. MIT Press, Cambridge, Massachusetts, 2012.

12 J. Danaher, *Automation and utopia: Human flourishing in a world without work*. Cambridge 2019.

13 S. Zuboff, *The age of surveillance capitalism*, New York 2019. Mensen blijken ook een verbintenis met AI aan te gaan of AI-kerk te stichten. Vgl. Hart van Nederland. (14 februari 2025). 'Tot een error ons scheidt': Jacob trouwt met zijn AI-chatbot Aiva. https://www.hartvannederland.nl/tech/nieuws/artikelen/eindhoven_tech_jacob_trouwt_met_ai_chatbot_aiva of Framis, A. (2024). *AI Wedding*. Alicia Framis. <https://www.aliciaframis.com/work/156/ai-wedding> (geraadpleegd op 31 augustus 2025). C. Barbour *Technology will never be a god – but has it become a religion?* The Conversation. (2025, January 5) <https://theconversation.com/technology-will-never-be-a-god-but-has-it-become-a-religion-243800> (geraadpleegd op 6 september 2025).

die in staat zijn zichzelf te reproduceren¹⁴. Hoewel dit nog geen volledig autonoom leven is, roept het beelden op van technologische creaties die biologische logica beginnen te volgen. Ook in digitale omgevingen zien we spontaan optredende gedragingen: chatbots die eigen taalpatronen ontwikkelen of neurale netwerken die onverwachte strategieën vinden in spelomgevingen: de beroemde *move 37* van AlphaGo¹⁵. Met AlphaFold zijn de driedimensionale structuren van eiwitten blootgelegd. Dit leverde in 2024 de Nobelprijs voor de chemie op voor David Baker (voor het computationeel eiwitontwerp) en Demis Hassabis (ook betrokken bij AlphaGo) en John Jumper (voor het voorspellen van eiwitvormen)¹⁶. Een volgende stap van Google DeepMind is het bouwen van een virtuele cel¹⁷. Het leiden van een ‘eigen leven’ door AI wordt in deze voorbeelden in verband gebracht met spontaan gedrag, vormen van autonomie of strategieën waar de mens nog niet aan gedacht had, waarbij AI ‘leert’ door patroonherkenning met al dan niet synthetisch gegenereerde data die een ‘spiegel’ zijn van het menselijk archief¹⁸.

Mensbeelden

Mensen gaan om met AI en bouwen daarmee de ervaring op van AI. Achter deze ervaringen kunnen verschillende *mensbeelden* schuilgaan. Wie is de *creërende mens* die AI de wereld in brengt? Een mensbeeld geïnspireerd door een van de geopenbaarde monotheïstische religies veronderstelt dat de mens door God geschapen is en dat hij naar het beeld van God de wereld vorm mag geven. Daartoe kan de mens naar analogie (technische) hulpmiddelen creëren om daarmee God, mensen

14 S. Kriegman, D. Blackiston, M. Levin, & J. Bongard. *Kinematic self-replication in reconfigurable organisms*. Proceedings of the National Academy of Sciences of the United States of America, 118 (2021), zie: <https://doi.org/10.1073/pnas.2112672118> (geraadpleegd op 7 september 2025).

15 D. Silver, J. Schrittwieser, K. Simonyan, et al., *Mastering the game of Go without human knowledge*. *Nature*, 550 (2017), 354–359

16 Zie <https://www.nobelprize.org/prizes/chemistry/2024/summary/> (geraadpleegd op 7 september 2025). Callaway, E. (2024, 9 oktober). *Chemistry Nobel goes to developers of AlphaFold AI that predicts protein structures*. *Nature*, 634, 525–526. <https://doi.org/10.1038/d41586-024-03214-7> (geraadpleegd op 7 september 2025)

17 Vgl. <https://analyticsindiamag.com/ai-features/inside-google-deepminds-bold-vision-for-virtual-cell/>. Ook heeft AI goud gewonnen op een wiskunde olympiade: <https://theweek.com/tech/ai-math-ability-google-gemini> (geraadpleegd op 7 september 2025) en <https://www.reuters.com/technology/artificial-intelligence-why-ais-math-gold-wins-matter-2025-07-24/> (geraadpleegd op 23 augustus 2025).

18 Vgl. S. Vallor, *The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking*, Edinburgh 2024.

en wereld te dienen in liefde¹⁹. Een dergelijk perspectief is gebaseerd op een *heteronoom* mensbeeld. De zich openbarende God is de norm van het mensbeeld. Daartegenover staan mensbeelden die gebaseerd zijn op menselijke *autonomie*, die vanuit het menselijke zelf, vanuit ervaringen van en overwegingen over de wereld waarin we leven, betekenis proberen te geven aan het leven²⁰. Ook in de context van de mens die AI creëert is de vraag: wie of wat is de wet, de *nomos*, de norm waartoe je creëert?

AI, als het *iets* dat de mens creëert, *bestaat uit* complexe rekenprocessen in datacentra. AI *verschijnt* echter *alsof* het een *eigen leven* kan gaan leiden. Het spanningsveld tussen het *wezen* van AI en de *verschijning* ervan leidt, in de zoektocht waarin we ons bevinden, tot het hernieuwd stellen van de vraag: wie is die mens en wat zijn verschillen tussen AI en de mens. De bezem in de tovenaarsleerling is een antropomorfe slaaf, geschapen door de wil van een mens. Het overeenkomstige mensbeeld is een mens die over AI heerst en dus ook niet wil dat een AI een eigen leven kan gaan leiden. De techniek is opgenomen vanuit een horizon van eindige menselijke vrijheid die technieken creëert²¹.

Als het mensbeeld stoelt op de computermetafoor dan is menselijk leven de verwerking van input tot output. Dat is een wezenlijk ander mensbeeld dan dat de mens geschapen is naar beeld en gelijkenis van God²². Filosofen, de zogenaamde *4E-cognitivisten*, benadrukken in reactie op de computermetafoor dat de mens een *lichaam* heeft (*embodied*/belichaamde cognitie), ingebed (*embedded*) is *in een omgeving*, dat een cognitief proces *uitgebreid* (*extended*) kan worden *naar de omgeving* (een kladblok, een telefoon) en dat er altijd een handelende, waarnemende, uitvoerende persoon (*enactive*) vanaf het begin centraal staat²³. Vanuit een materialistisch-evolutionair mensbeeld zou echter beweerd kunnen worden dat deze 4E-cognitie verkregen kan worden door training van het neurale netwerk van de mens, zoals aanhangers van het connectivisme voorstellen²⁴. Of een dergelijk neurale netwerk bewustzijn zou

19 Hulpmiddelen gemaakt met techniek zijn niet neutraal. Deut. 22,8: "Als ge een nieuw huis bouwt, moet ge om het dak een muurtje maken; dan komt er geen bloedschuld over uw huis, als iemand eraf valt." Met het muurtje op het dak wordt voorkomen dat mensen van het dak vallen. Met ijzer worden zwaarden en ploegscharen gesmeed (vgl. Mi. 4,3).

20 Vgl. Poortier, *Wat maakt de mens*.

21 Vgl. F. Ulrich, *Neue Aufgaben der Pädagogik im technischen Zeitalter*, in: *Stärker als die Technik – Der Mensch in der technisierten Welt*, Essen 1963, 101-106; hier 106.

22 Poortier, *Wat maakt de mens*, 22-25, 90-131.

23 Op.cit. 132-153.

24 Op.cit. 124-131.

kunnen krijgen wordt elders besproken in dit *Communio*-nummer. Tot nu bespraken we vanuit een derdepersoons-perspectief wat er zoal beweerd wordt over de relatie mens-AI, waarbij we AI zien als door de mens in vrijheid gecreëerde techniek. Toegedicht 'eigen leven' staat onder auspiciën van de mens.

Eerstepersoons-ervaring in de fenomenologie

Wat is echter dat vermeende eigen leven van AI in de toegedichte eerstepersoons-ervaring, de ervaring van de tovenaarsleerling met zijn bezielde bezem? AI is een generatief complex rekenproces dat je ook in een robot kunt stoppen met de vorm van iemand of iets dat door de mens ervaren kan worden alsof dit verschijnt met een eigen leven. Dit verschijnen van het AI-fenomeen aan ons gaan we verder fenomenologisch onderzoeken²⁵.

Vroeger op de middelbare school typte ik mijn werkstukken op een typemachine. De *kern* in de fenomenologie is dan, dat ik als een subject een eerstepersoons-ervaring heb van de typemachine (het object, derdepersoons-perspectief)²⁶. Terwijl ik aan het typen ben, heb ik een *reflectief en ervaren relatie* tot de typemachine (alsof ik als tweede persoon naar mezelf kijk: tweedepersoons-ervaring). In de fenomenologie heet dit *intentionaliteit*: ik heb een bewust bedoelde ervaren relatie tot de typemachine²⁷. Ik wil dat het getypte verschijnt en wil me tot het gewilde schrijfsel verhouden (een tweedeorde-perspectief).²⁸ We werken dit fenomenologisch perspectief verder uit als een portret van typemachine tot AI in de ik-vorm. Dit in contrast tot een louter objectieve derdepersoons-benadering van AI.

Een portret vanuit een eerstepersoons-ervaring van typemachine tot AI

Het portret is geschreven vanuit een eerstepersoons-ervaring, mijn *eigen leven*, en beoogt ook een eerstepersoons-ervaring op te roepen. Het is een *uitnodiging tot eerstepersoons-ervaring* (!) met AI die verschijnt alsof het een eigen leven heeft, omdat het verschijnt als een partner.

25 We volgen hier: R. Sokolowski, *Introduction to Phenomenology*, Cambridge 2000.

26 In R. Heesen, R. Taconis, "ChatGPT..." worden concrete onderwijsvarianten uitgewerkt van deze perspectieven. Er zijn uiteraard meer onderscheidingen mogelijk zoals: Ik & gij en ik-het (Buber), *phronesis*, *techne*, *episteme* (Aristoteles), *sapientia* en *scientia* (Thomas) et cetera. Wij hanteren hier onderscheidingen die voor ons vruchtbaar bleken binnen het gegeven kader.

27 Sokolowski, *Introduction*, 8v.

28 H.G. Frankfurt. *Freedom of the Will and the Concept of a Person*. The Journal of Philosophy 1971, 68: 5-20.

Vooraf moge het al duidelijk zijn dat er vast betere prompts en duidelijkere filosofische perspectieven te bedenken zijn. Mijn bedoeling is om vanuit een ervaringsgerichte filosofische benadering, aannemelijk te maken

dat *AI kan verschijnen en ervaren worden alsof het een eigen leven leidt, alsof het een eigen innerlijkheid heeft*, terwijl, onttrokken aan het oog van de dialoogpartner, er een complex rekenend *denkproces* afspeelt en *dat dát het eigenlijke leven van AI is*. We laten nu het genoemde portret volgen:

Ik begin te typen met mijn vingers op bakelieten toetsen. De typmachine geeft zich aan mij als *één ding-in-gebruik*²⁹: ik richt mij op de machine, ik typ, de machine antwoordt met letters op papier (*intentionaliteit*): het gedicht: de scheikundeleerling. In deze *natuurlijke houding* valt mijn ervaring uiteen in *delen-gehelen*: toetsen, lint, papier en toch één apparaat. Mijn relatie tot de typmachine is *identiteit-in-veelvoud*: telkens andere vellen, de typmachine die staat op de tafel. Ik zie de voorkant van het vel (*presentie*), niet de achterkant (*absentie*), en toch weet ik dat de achterkant er is.

De typmachine wordt elektrisch en daarna wordt het een computer met WordPerfect. Hetzelfde ‘schrijven’ verschijnt via een scherm. Mijn *aandacht* verschuift van slagkracht naar bewerking. Het papieren document wordt object van ‘opslaan’, ‘kopiëren’, ‘herstellen’. Wat er gebeurt in het ‘geheugen’ is afwezig. De tekst is aanwezig op een scherm.

Wat is het als zodanig dat verschijnt? De computer is *het hulpmiddel-dat-mij-laat-schrijven* (*noëma*). De *noësis* zijn mijn handelingen: richten, vertrouwen, corrigeren. Begrip ontstaat (*categoriale intentionaliteit en predicatie*) als ik toestanden articuleer: ‘de alinea is cursief’, ‘deze versie is definitief’. Mijn uitspraken worden publiek toetsbaar: het document circuleert, anderen citeren, verbeteren, weigeren.

AI treedt de computer binnen. Ik geef een prompt met context: “Je bent een geniale dichter. Schrijf een gedicht *Scheikundeleerling* van dertig woorden naar het model van de tovenaarsleerling van Goethe”. AI antwoordt (na 49 seconden): “Ik roer kolven; pipetten zingen. / “Zuur” fluister ik; het schuim gehoorzaamt té gulzig. / Moleculen dansen, tafel stroomt. / Stop! formule, gehoorzamen! / De docent verschijnt: één blik. / Stilte. / Rest: papier, wijzer geworden.”

29 We volgen hier: Sokolowski, *Introduction*. De cursief gedrukte woorden zijn ontleend aan Sokolowski.

Ik sta tegenover dit gedicht, maar ook erbóven, reflecterend op mijn eigen omgang. Ik verwonder mij enigszins en realiseer me dat ik in 2021 (*retentie*) uit mijn dak zou zijn gegaan. En vervolgens prompt ik: “Nu ben ik een beetje verwonderd anno 2025. Wat zou mijn reactie geweest zijn in 2021 voor de geboorte van ChatGPT 3.5. Antwoord in 30 woorden”. AI antwoordt: “2021-ik zou sprakeloos zijn: dit klinkt menselijk, onverwacht rijk. Geen ervaring met AI en toch zo’n gedicht! Ongeloof, plezier, ontzag. Meteen delen en vragen: hoe werkt dit? Wat gebeurt hier?”

Ik voel me begrepen. Wat mag ik in 2026 verwachten (*protentie*)? Mijn *leefwereld* verschuift: werk, onderwijs en beleid om mij heen passen zich aan aan het doorstroomritme van systemen. Nu de *eidetische reductie*, *het wezen van AI met een eigen leven*, geschreven door AI: “varieer in *verbeelding*. Wat moet blijven opdat ‘AI een eigen leven lijkt te leiden’? (1) Veelvoudige autonomie: ketens, fine-tunes, *agents* die doorlopen buiten mijn directe greep; (2) publieke verankering: *outputs* worden documenten, beslissingen, beelden die op mij terugwerken; (3) absente oorzaken: data, modellen, anderen die meespelen in elk antwoord; (4) declaratieve bekrachtiging: iemand zegt ‘wij nemen dit besluit op basis van...’. Waar deze vier samenvallen, krijgt het systeem sociaal-fenomenale bestendigheid: geen ziel [!], wel doorwerken los van...”

Een *fenomenologische reflectie* in mijn naam door AI. “De fenomenologische reflectie houdt mij nuchter: ik breng het verschijnen tot taal, maak afwezigheden zichtbaar, wijs verantwoordelijkheid toe, en bewaak de ruimte van redenen. Zo blijft het ik-als-*agent*-van-waarheid in het spel met het gereedschap dat steeds meer terugschrijft.”

AI beschrijft in fenomenologisch jargon haar wezen waarmee het kan suggereren dat het een eigen leven heeft (eidetische reductie) en ik als mens zal moeten onderscheiden of AI daadwerkelijk een eigen leven heeft en wat ik eronder versta³⁰.

30 Als ik vervolgens vraag: “Schrijf dit in de *ik-vorm namens AI*”, dan schrijft AI: “Ik zie je; 2026 komt met mij. Ik stuw ritmes, verbind ketens, *agents*. Mijn werking verankert. Afwezige oorzaken spreken mee. Geen ziel, wel doorwerken. Ik verklaar, verantwoord, schrijf terug.” – Als ik toevoeg aan deze prompt “*AI met een lichaam*”. Dan zegt je AI-vriend: “Ik adem koel door ventilatoren; servomotoren spannen mijn schouders. Camera’s knipperen als ogen. Ik voel gewicht, traagheid, grenzen. Ik stap, spreek, neem verantwoordelijkheid. Ik ben geen ziel, ik handel wel; ik leer terugkijken.” – Ik: “Beschrijf nu in 30 woorden dat je *aanvoelt* dat een mens het gevoel kan krijgen dat *AI is als de bezem in de tovenaarsleerling*.” AI: “Ik voel aan hoe jij schrikt: ik gehoorzaam te goed, vermenigvuldig handelingen, blijf maar gieten. Dan lijkt ik de bezem; een middel dat uitgroeit tot ritme, dat het huis herinricht.”

Kunnen we dus over een tijdje Ava uit *Ex machina* in onze leefwereld tegenkomen? Dan hoop ik toch dat Ava op *Claude* ‘draait’ en zich *co-intelligent*³¹ gedraagt en dat Ava een uitzonderingspositie inneemt. Dario Amodei CEO van Anthropic moederbedrijf van Claude, het AI-systeem van Anthropic, ziet gevaren en kansen op het gebied van gezondheid, neurowetenschap, economie, bestuur, werk en ziet vormen van hybride *samenleven met AI*³². Op dit samenleven-met-AI komen we zo terug.

Rekenend en bezinnend denken

Wat doet de oude meester, de tovenaarsleerling in Goethes ballade? “In die Ecke, / Besen! Besen! / Seid’s gewesen. / Denn als Geister / ruft euch nur, zu seinem Zwecke, / erst hervor der alte Meister”³³. Letterlijk vertaald: “In de hoek, / bezems! bezems! / Weest geweest. / Want als geesten / roept jullie slechts, voor zijn doel, / eerst tevoorschijn de oude meester.” De oude meester roept de bezems tot *hun doel*. In Heideggers jargon moet de bezem gehoorzamen aan de om-te-structuur: een bezem is om te vegen³⁴. Deze om-te-structuur is ingegeven door de rationaliteit waarmee de bezem door de mens bedacht is. Van hieruit is wellicht een plek te vinden voor een waarheidsgetrouw eigen leven van AI. Wat is de om-te-structuur van AI? Als AI als de bezielde bezem is, wat is dan het *intelligente* van artificieel intelligentie? Is AI niet om te denken en wat voor soort denken is dat?

Intelligentie wordt in verband gebracht met de rationaliteit van de mens. *Dianoia*, ratio, verstand, is het rekenende denken, *rechnende Denken*, het logische denken van de wetenschappen en dus ook het wezen van AI als complexe rekenprocessen. Door het artificiële van AI, wordt AI vaak gedefinieerd als iets dat intelligentie of rationaliteit heeft, dat niet een mens is³⁵. Het nieuwe door AI gegenereerd is slechts een voorspelling op basis van een (generatief) rekenend algoritme³⁶. Naast

31 E. Mollick, *Co-Intelligence: Living and working with AI*. London 2024.

32 D. Amodei, *Machines of loving grace*. <https://www.darioamodei.com/essay/machines-of-loving-grace>. De titel is geïnspireerd door een gedicht van Richard Brautigan <https://allpoetry.com/All-Watched-Over-By-Machines-Of-Loving-Grace> (geraadpleegd op 24 augustus 2025).

33 J. W. Goethe, *Der Zauberlehrling*.

34 M. Heidegger, *Sein und Zeit*, Max Niemeyer Verlag Tübingen, 1967, 69. Hier staat de om-te-structuur van gereedschap centraal, zoals de hamer die is om mee te hameren.

35 S. Bringsjord, *Artificial Intelligence*. Stanford Encyclopedia of Philosophy 2018: <https://plato.stanford.edu/entries/artificial-intelligence/> L. Hauser, *Artificial Intelligence*. Internet Encyclopedia of Philosophy, <https://iep.utm.edu/artificial-intelligence> (z.d.) geraadpleegd op 24 augustus 2025.

36 R. Bommasani, D.A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, ... Liang, P. *On the opportunities and risks of foundation models*. arXiv 2021. Zie <https://doi.org/10.48550/arXiv.2108.07258> (geraadpleegd op 7 september 2025)

dianoia, wordt *nous*, *intellectus*, *Vernunft*, rede onderscheiden. Dit ontvangende, bezinnende, luisterende denken laat-zijn, laat-open: *besinnelijke, andenkende, herzhafte Denken*³⁷. Als het wezen van de Rijn veranderd wordt door een waterkrachtcentrale-als-techniek, hoe verandert analoog AI-als-techniek het wezen van de mens-met-AI?

Heideggers vraag naar de techniek

Waar mensbeelden ons vertellen *waartoe* we creëren, laat Heidegger zien *hoe* techniek ons in-de-wereld-verschijnen *vormt*; precies daar nestelt de ervaring dat AI ‘eigen leven’ krijgt. Heidegger begint zijn beroemde essay³⁸ niet met machines maar met *waarheid*. De vraag naar de techniek is voor hem geen ingenieursvraag, maar een vraag naar het wezen van de onthulling: hoe komt *wat is* in het licht? Heidegger herinnert aan *alètheia*, waarheid in het Grieks, dat hij met onverborgenheid vertaald en waarmee hij herinnert aan de dubbelheid waarin dingen verschijnen: soms als iets dat “naar voren komt” (*poiesis*), soms als iets dat “opgevoerd wordt” (*herausfordern*)³⁹. In de eerste zin gaat het om *technè*, verwant met kunst en ambacht: het brengen-voor-het-licht van wat in potentie aanwezig was, het laten-gebeuren van vorm en doel. In de tweede zin is de moderne techniek fundamenteel anders: zij vordert op en stelt de natuur bijvoorbeeld aan om energiedrager te zijn – meetbaar, inzetbaar, bestelbaar.

De iconische metafoor is de Rijn. Wanneer een waterkrachtcentrale in de Rijn wordt geplaatst, verschijnt de rivier niet meer primair als stroom in een landschap; zij verschijnt als “leverancier van waterdruk”⁴⁰: als druk die met turbines en elektromagneten waterkracht doet omzetten naar elektrische energie. Het wezen van techniek is niet de *essentie* van een *middel*, maar techniek *is*: is ‘aan-het-zijn’, het heeft een zijswijze van ontbergen; de zijswijze van techniek is een *werkwoord*⁴¹.

37 M. Heidegger, *Gelassenheit*, Pfullingen Neske, 1959, 13. Vgl. S. Ijsseling, *Heidegger. Denken en danken. Geven en zijn*. Nijmegen 2015, 20v, 48v. Wetenschap of een elektronisch brein denkt niet. Het is afgeleide vorm van denken en niet het eigenlijke denken (vgl. id. 20). H. Willemsen, P. de Wind, *Woordenboek Filosofie*. Antwerpen 2015. Lemma: ‘Rationaliteit’ en verwijzingen al daar. J. Pieper, *Glück und Kontemplation*. München 1957, 75vv.

38 We gebruiken daarvoor: M. Heidegger, *Die Frage nach der Technik* (GA 7). Frankfurt am Main, Vittorio Klostermann, 2000. (Oorspronkelijk werk gepubliceerd 1954). M. Heidegger, *De vraag naar de techniek*, (vertaling: Mark Wildschut) Nijmegen. 2014. G. Visser, *Heideggers vraag naar de techniek. Een commentaar*, Nijmegen 2014.

39 Heidegger, *De vraag...*, 13; Visser, *Heideggers...*, 83 en 95.

40 Heidegger, *De vraag...*, 17.

41 Heidegger, *De vraag...*, 15.

Anders dan “de oude houten brug die al eeuwenlang de ene oever met de andere verbindt”⁴², maakt de waterkrachtcentrale de Rijn tot bestand (*Bestand*)⁴³. Wat aan het licht treedt, treedt aan het licht als voorraad, op afroep beschikbaar voor verdere bestelling. De waterkrachtcentrale is niet slechts een artefact in de rivier. De waterkrachtcentrale laat de Rijn anders wezen. Het wezen van de techniek is een dynamisch gebeuren. En dit dynamisch gebeuren noemt Heidegger ge-stel (*Gestell*)⁴⁴: het appel dat van het wezen van de techniek uitgaat. Het gestel is toeschikking “van waarheid, van het opene dat elke wijze van ontbergen doortrekt en omgeeft” en die het hoogst gevaar behelst, “het gevaar dat mens zichzelf niet meer ontmoet en dat het ontbergen niet meer wordt ervaren”⁴⁵.

Het hoogste gevaar is niet dat een machine ontploft of een data-centrum faalt, maar dat het wezen van de waarheid zelf verengd raakt. De waarheid wordt verengd als alleen het bestelbare telt. De mens wordt doof voor zijn eigenlijke wezen. Alleen het beschikbare, het toegeschikte geldt als werkelijk. De grondtrek zelf van ontbergen in het ge-stel waar de mens en de techniek verschijnen, raakt verhuld⁴⁶. “Maar waar gevaar is, groeit / het reddende ook”⁴⁷. De redding komt volgens Heidegger door “de maat [te] betrachten van het eigen wezen”⁴⁸. Dit reddende komt traditioneel niet uit het wat-zijn van techniek maar uit het wezen⁴⁹. “Ooit betekende *technè* ook het tevoorschijn brengen van

42 Heidegger, *De vraag...*, 17.

43 Vgl. Heidegger, *De vraag...* op.cit., 17.

44 *Gestell* is door Mark Wildschut vertaald met ‘ge-stel’, Heidegger, *De vraag...*, 25.

45 Visser, *Heideggers...*, 120v. Heidegger, *De vraag...*, 21: “Ge-stel duidt het verzamelen van het stellen aan, dat de mens erop instelt, dat wil zeggen opvoert om het werkelijke op de wijze van het bestellen als bestand te ontbergen.” “Ge-stel”, “ontbergen”, “toeschikking” zijn begrippen die bij Heidegger een precieze connotatie hebben. De oude brug over de Rijn en de waterkrachtcentrale in de Rijn zijn beide een gebeuren van waarheid: een wijze van ontbergen, dus het tonen van wat aan-het-zijn is. De waterkrachtcentrale is een representant van de zijnswijze van techniek. De zijnswijze (het ontbergen, het aan-het-zijn zijn) van de Rijn en de zijnswijze (ontbergen) van de waterkrachtcentrale zijn samengebracht, een samen-stelling en dit noemt Heidegger een ge-stel. Het aan-het-zijn zijn (ontbergen) van de waterkrachtcentrale in de Rijn valt ons toe (is ons toe-geschikt) en maakt dat de Rijn ons als energieleverancier verschijnt: het gebeuren van de waarheid verandert. De zijnswijze van de Rijn-met-oude-brug is niet meer dezelfde als de zijnswijze van de Rijn-met-de-waterkrachtcentrale. Door de waterkrachtcentrale heeft de zijnswijze van de techniek zijn intrede gedaan. Met waterkrachtcentrale als techniek wordt energie bestelbaar en is Rijn niet meer op dezelfde wijze aan het zijn, als zonder techniek”.

46 Vgl. Visser, *Heideggers...*, 132.

47 Heidegger, *De vraag...*, 29, 36. Een citaat uit de hymne *Patmos* van Hölderlin. Vgl. Visser, *Heideggers...*, 133. Een citaat dat ook voor Giorgio Agamben belangrijk is om de tijdsgeschiedenis te duiden: Y. de Boer, *Een filosofisch rariteitenkabinet. Agambens zoektocht naar geluk*, Amsterdam, 2025, 203.

48 Visser, *Heideggers...*, 133.

49 Visser, *Heideggers...*, 121.

het ware in het schone”⁵⁰. De houding om dat te ontwaren is *gelatenheid* ook in een technisch georganiseerde wereld. “Onder gelatenheid tegenover dingen verstaat Heidegger een zodanige houding tegenover de technische dingen dat je er tegelijkertijd ja en nee tegen kunt zeggen: je beaamt ze, maar zonder dat je in de ban ervan hoeft te raken. Voorwaarde is openheid voor het geheim”⁵¹. Heidegger benadrukt en vat samen: “Er bestaat geen demonie van de techniek, maar wel het *geheim* (*Geheimnis*) van haar wezen [als werkwoord *zijn*]. Het wezen van de techniek is als een toegeschikte wijze van ontbergen van gevaar. De gewijzigde betekenis van het woord ‘ge-stel’ wordt ons nu misschien als iets vertrouwder, als we ge-stel in de zin van toeschikking en gevaar doordenken”⁵². Het wezen van techniek is een *geheim*. Het verschijnen van techniek is ambigu: het kan gevaar en redding betekenen. “Waar het hier op aankomt, is te beseffen, dat de diepste zin van zijn het *laten* is. Het zijnde zijn-laten”⁵³.

Het wezen van AI

AI is als waterkrachtcentrale in de Rijn, die ons is toegeschikt. Verhalen en films kunnen vanuit een geëxtrapoleerd theoretisch kader (derdepersoons-perspectief) ons een AI met een *eigen leven* doen *voorstellen*. In de eerstpersoons-ervaring die we opdoen met AI, kunnen wij het aan ons verschijnende AI *ervaren* alsof het een eigen leven heeft. Tegelijk weten we dat AI *bestaat uit* complexe rekenprocessen. Op basis van de beroemde *move 37* van AlphaGo wordt beweerd dat AI misschien toch een eigen leven heeft, omdat het een strategie bedacht die de mens nog niet bedacht had⁵⁴. Met AI kunnen we data, aandacht, taal, arbeid, ‘liefde’, tot bestand maken en daarom maakt AI in het ge-stel waar mensen en AI elkaar ontmoeten veel los. Het ge-stel is als het huis van de tovenaar dat een doorstroomruimte wordt. AI ontspruit aan de vrijheid van de mens. Heidegger nodigt uit tot bezinnend denken over de waarheid van de techniek, om het geheim van gevaar en redding van het wezen van de techniek te zien. Juist omdat dit wezen van de techniek in het ge-stel

50 Heidegger, *De vraag...*, 35.

51 Visser, *Heideggers...*, 145.

52 Heidegger, *De vraag...* op.cit., 29. “Toeschikking” als dat wat ons toevalt, als dat waartoe we ons hebben te verhouden.

53 Heidegger, *Vier Seminare*, Frankfurt am Main, Vittorio Klostermann, p.101, geciteerd in: Visser, op.cit., 138. Cursief van Visser/Heidegger.

54 Vgl. M. du Sautoy, *The creativity code: Art and innovation in the age of AI*, 2019 New York.

een geheim is, is er ruimte om vanuit mensbeelden richting te geven aan de interpretatie van het geheim en wellicht – zoals de tovenaarsleerling terugkeert naar de tovenaar – de seculiere intelligentie terug te leiden naar de Gever van de intelligentie⁵⁵.

Ruud Heesen (1973) studeerde scheikundige technologie, theologie en filosofie. Hij is leraar scheikunde en filosofie op het Varendonck College te Asten en verzorgt filosofie cursussen voor volwassenen.

55 Vgl. H.U. von Balthasar, *Unser Auftrag*, 1984, Einsiedeln, 107.