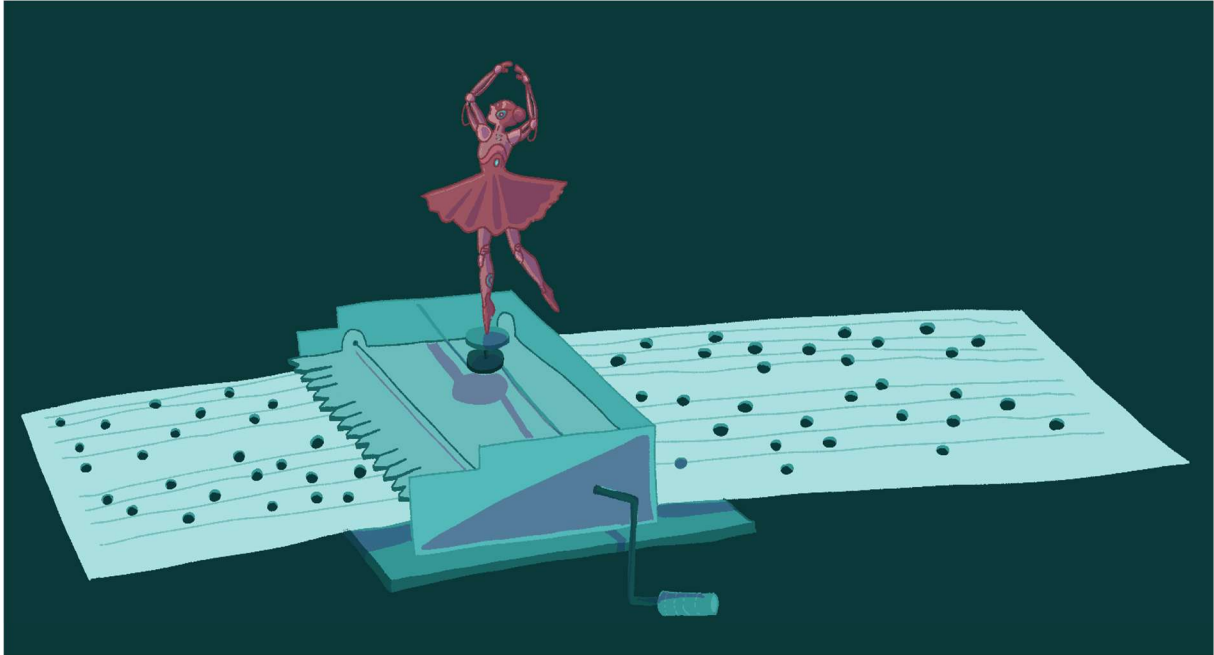




Een filosofische zoektocht naar de effecten van
informatiesystemen op het menselijk bestaan.

Maieutiek, najaar 2025

Inleiding

Mijn nieuwsgierigheid was gewekt toen ik mensen de gevaren van AI hoorde benoemen en meende dat ik daar een verre echo van Socrates bezwaren tegen het schrift in hoorde. Ik was alleen nooit aan deze cursus begonnen als ik niet Michiel van der meer bereid had gevonden om mijn filosofische ideeën vanuit zijn kennis aan te vullen. AI, de werkelijkheid van wat de techniek van de computers kan razend snel aan het veranderen. Daarbij is de methode waarvan ik gebruik moet maken: kennis vergaren en verbindingen daartussen leggen, zelf onderdeel geworden van de verandering. Kan ik de proefschriften, boeken en essays nog wel opvatten als door mensen gecreëerde inhoud met als doel waarheidsvinding? Kortom deze cursus is een zoektocht om het kader te onderzoeken en te verhelderen waarbinnen wij onszelf als denkende wezens kunnen begrijpen als onderdeel van een wereld die in grote verandering is.

Het goede nieuws: dit is niet de eerste keer dat de informatiesystemen drastisch veranderen en mensen zijn er altijd weer uitgekomen en benoemen het achteraf als vooruitgang. Het slechte nieuws: de mensen die de verandering meemaken denken daar vaak anders over. De onderwerpen:

1. Informatie: waarheid, feiten en verhalen
2. Oude en nieuwe overgangen
3. Gevaren van AI, gevaren van mensen
4. Perspectief

1. Informatie, waarheid, narrativiteit en feiten

Kennis en macht

Harari laat zien dat in de naïeve versie informatie leidt tot waarheid en dat dit weer leidt tot wijsheid en macht. Hiermee beschrijft hij niet alleen een naïeve versie, maar ook de versie die vanaf de Griekse filosofen via de Verlichting aan hedendaagse westerse mensen is meegegeven als een epistemische onderlaag waarover nauwelijks gediscussieerd werd.

Kennis maakt ons in deze naïeve opvatting wijzer en machtiger. We krijgen daardoor in de loop van de menselijke geschiedenis meer 'macht over' en 'macht om'. Mensen krijgen macht over de natuur, over anderen en zichzelf. Er is ook meer macht in de zin van de potentie om dingen voor elkaar te krijgen. Uiteenlopende zaken als zelfbeheersing, natuurkunde en politiek zijn het gevolg.

Dat we sinds de jaren 60 van de 20^{ste} eeuw ook veel aandacht hebben ontwikkeld voor de problematische aspecten van macht leek in eerste instantie meer van hetzelfde: we begrijpen dan (door meer informatie) dat sommige mensen meer profiteerden van kennis dan anderen. Kennis geeft macht en kennis werd, al dan niet opzettelijk, niet gelijkelijk gedistribueerd. Maar nog steeds dacht men dit op te lossen door meer informatie: over racisme, gebrek aan inclusie en andere perspectieven.

Na de tweede wereldoorlog ontstaat een brede twijfel aan de basis van de informatie - kennis relatie zoals die tot op dat moment werd gezien. Zou het ook andersom kunnen zijn? Had Nietzsche gelijk toen hij zei 'dat macht kennis maakte'? Harari beschrijft het als volgt:

Informatie geeft orde en waarheid en beiden geven macht, maar alleen waarheid geeft wijsheid.

Harari beschrijft ook dat het populisme alle informatie ziet als deel van een machtsstrijd. Hij verwacht hier m.i. het postmoderne productieve spreken met politieke macht zoals overheersingstechnieken. Een kleine uitleg is nodig:

- De postmoderne opvatting is dat elk spreken (informatie) macht heeft, omdat informatie niet zozeer beschrijvend is maar vooral vormend. Dit hangt samen met wat Foucault het productieve spreken of discours noemt. De wereld van informatie wordt geordend met meer informatie namelijk: wie er mag spreken en wie niet en hoe we waarheid kunnen vinden en hoe niet. Dit noemt Foucault 'de orde van het spreken'. Deze orde is nodig opdat mensen in alle informatie die er zou kunnen zijn zich een weg kunnen banen.
- Overheersing d.m.v. informatie gaat over het opzettelijk weghouden van informatie van sommige mensen of juist het geven van valse informatie aan of over bepaalde groepen. Het liegen of bullshitten dat daarbij plaatsvindt is een omgang met informatie die macht tot doel heeft en waarbij de waarheidsvinding of het bereiken van een orde die tot een homeostase kan leiden wordt veronachtzaamd. Sterker nog:

het creëren van misinformatie en wanorde kan een manier zijn om tot meer macht te komen.

Beide lijken mij aspecten die juist in het kader van AI onderzocht moeten worden, maar niet door ze onder een noemer te brengen.

Taal en woorden

In het gebruik van de woorden: informatie, feiten, waarheid en verhalen wordt het echt ingewikkeld. Harari doet de -ooit vanzelfsprekende gedachte- dat meer informatie de oplossing is die leidt tot meer waarheid en tot het rechtzetten van vergissingen, af als naïef. Maar zelf heeft hij een deel van de Verlichtingsgedachte nog overeind gehouden, die door de meeste filosofen van nu ook als naïef wordt afgedaan: De informatie valt bij Harari uiteen in een objectieve waarheid die over feiten gaat en een subjectieve waarheid die daar niet over gaat.

De huidige kennistheorie kan dat niet langer onderschrijven: Al vanaf Kant weten we dat de wereld 'op en voor zichzelf' niet kenbaar is op die objectieve wijze. Met andere woorden wij mensen kennen alleen de wereld zoals die door mensen empirisch ervaren wordt. Maar daarna kwam ook daar een probleem aan de orde: mensen denken steeds dat ze objectieve feiten vinden, maar eigenlijk maken ze theorieën die bepaalde fenomenen verklaren. Het objectieve feit op zichzelf wordt nooit bereikt. Dat wil niet zeggen dat er geen vooruitgang van kennis kan zijn en zelfs niet dat daar objectieve feiten aan ten grondslag liggen. De taal waarmee kennis wordt uitgedrukt, de instrumenten waarmee we het onderzoeken en zelfs de toets die we gebruiken om waar en onwaar van elkaar te onderscheiden zijn echter contextueel, tijdgebonden en afhankelijk van de stand van de techniek. Ja, er zijn feiten die zo onontkoombaar zijn dat niemand ze betwijfelt. Andere feiten worden echter door een voortgang van informatiestromen en technieken achterhaald. 'Falsificatie', noemt Popper dat in de jaren 30 van de 19^{de} eeuw.

Dus nee, wetenschap is niet maar een mening, maar objectieve (eeuwige) feiten als inhoud van de informatiestromen zijn er ook niet. Er zijn wel zaken in de werkelijkheid die zich voordoen los van wat mensen erover weten. Ik heb deze in het college '*Gegenständlichkeit*' genoemd om ze te onderscheiden van wat meestal met 'objectieve feiten' wordt aangeduid. Deze '*Gegenständlichkeit*' kunnen mensen niet objectief kennen, maar deze doen zich wel als onontkoombaar voor in onze zoektocht naar kennis. Het doel van die zoektocht is niet 'de Waarheid' maar homeostase. Homeostase is een leefbaar evenwicht dat zich op elk niveau voordoet: in onszelf en tussen onszelf en de eisen van onze omgeving, en tussen mensen onderling. Om homeostase te bereiken worden ervaringen met *Gegenständlichkeit* aaneengeregen door middel van verhalen: verbanden die mensen weven en die niet per se waar of objectief hoeven zijn, als het maar een leefbare werkelijkheidsservaring oplevert. Mythen en wetenschap vermengen zich hier.

In zijn voorgaande werk sprak Harari van narratieven of verhalen. Nu hij ons iets uit wil leggen over AI heeft hij de term informatiesystemen gebruikt, maar daarin schuilt het misverstand dat informatie een relatie heeft met waarheid en feiten en dat hoeft nadrukkelijk niet het geval te zijn.

Mensen zijn narratieve wezens dat wil zeggen dat wij verhalen vertellen over hoe wij hier gekomen zijn, wie wij nu zijn en wat onze situatie is en op grond daarvan perspectieven beschrijven. Wie 'we' zijn is overigens ook deel van dat narratief: mannen, vrouwen, mensen, burgers of 'echte Nederlanders'. In dat narratief speelt Waarheid maar een beperkte rol. Toch is het belangrijk dat we niet te ver van de Gegenständlichkeit afwijken, omdat we dan de homeostase en wellicht het overleven op het spel zetten. Dat aspect zal een rol gaan spelen in de gevaren van AI die Harari gaat beschrijven. Hij noemt dat rekening houden met de waarheid, maar dat is dus filosofisch gezien niet de juiste term. Ik zou zeggen: 'rekening houden met de werkelijkheid 'Gegenständlichkeit' die zelf niet feitelijk kenbaar is'. De verhalen hebben een zekere 'elasticiteit' ten aanzien van de Gegenständlichkeit, maar mogen daarmee niet te veel in tegenspraak zijn. Het probleem is dat vooral lange termijneffecten van een te ver afwijkend verhaal niet zichtbaar zijn, denk aan het klimaatverhaal versus de Gegenständlichkeit van een leefbare aarde.

Verhalen die bedóeld zijn om macht te krijgen of te behouden, vormen één probleem. Het probleem van de macht van taal zélf vormt een ander probleem. Laten we met het eerste beginnen:

Onwaarheden en chaos zijn het wapen van iedere machthebber die meer macht wil. Er zijn vele opties: complottheorieën, opkomen voor de ongehoorde stem van het volk, desinformatie. Daarbij is er de macht van geldstromen die een verhaal dat zij willen vertellen breed onder de aandacht kunnen brengen. Macht en informatiesystemen hebben een wederzijds bevorderende werking.

Maar onder en tussen dat geweld dat op macht uit is, gaat er ook iets veranderen omdat mensen juist beter geïnformeerd raken. Mensen van nu weten wat we allemaal niet weten, zien dat voortschrijdend inzicht ook laat zien dat kenniswerkers het ook fout kunnen hebben. Het aanzien van kennis neemt verder af omdat Chatgpt het zo gemakkelijk en zo snel weergeeft. De vraag 'klopt het wel?' lijkt minder belangrijk dan de snelheid van informatie. De vraag naar de orde is er bij Harari vooral één naar macht en onmacht in politieke zin, maar misschien nog gevaarlijker is het probleem van 'de orde van het spreken'. De term komt van Foucault. Hij schrijft dat taal machtig is omdat zij ook productief is. Taal, het spreken, beschrijft niet alleen de werkelijkheid: het maakt het ook. Omdat mensen dat gevaar hebben ingezien is het spreken in een orde gevat die beschrijft wie op welke wijze waarover een uitspraak mag doen: wetenschap, recht politiek status en aanzien doen er daarin toe. Maar door een toenemende kennisstroom naar iedereen is na de tweede wereldoorlog 'de orde van het spreken' zwaar verstoord. Ten eerste omdat mensen zich, al dan niet terecht, afvragen of deze orde niet bepaalde groepen of informatie uitsluit of

onderdrukt, maar ook omdat mensen die op macht uit zijn uit eigenbelang aan deze twijfel bijdragen.

Als we niet meer weten wie waarover iets mag en kan zeggen gaat de rol van instituties op de helling, is iedereen journalistiek bezig vanaf zijn eigen telefoon en is onderzoek doen en Chatgpt iets vragen ononderscheidbaar. De epistemologische fundamenteën die het gevaarlijke 'productieve spreken' moeten intomen worden omvergeworpen: het spreken zelf is ontketend. Dat maakt de rem op de -terecht door Harari en anderen gesignaleerde- nadelen van de snelle ontwikkeling van AI bijna onmogelijk want 'wie ben jij wel dat jij daar iets over mag zeggen?'. Cynisch constateren veel mensen dat informatie altijd een belang dient. De vraag is echter hoe instituties en wetenschap in dat licht de ons niet altijd welgevallige Gegenständlichkeit nog in informatiestromen kunnen overdragen?

2. Oude en nieuwe overgangen

In de Phaëdrus kunnen we lezen hoe Socrates verdedigt dat hij niet wil schrijven. Hij legt uit dat tekst die opgeschreven wordt dode letters zijn, iets wat we ook bij Nietzsche aantreffen. Het verschil is dat dialogen persoonlijk en contextueel zijn en daarmee iets wakker kunnen roepen bij de gesprekspartner wat raakt aan zowel het gevoel, als het denken. De orale traditie had weinig op met schrijven, schrijven was alleen handig voor wat je wilde tellen, meten en wegen. Kortom zaken van nut en huishouden en dus werk voor slaven. Het teruglezen van de tekst in een andere situatie roept vragen op die je niet aan de tekst kunt stellen zegt Socrates en dan wordt een tekst iets wat vloeibaar wordt in de handen van wie het leest en wie weet wat diegene er mee doet.

Wij kennen de afloop: Socrates verliest de toekomst en wij gaan waarde hechten aan het schrijven. Sterker nog, wij weten alleen van Socrates' weerzin tegen schrijven, omdat Plato zijn dialogen opschreef.

We kunnen er twee conclusies aan verbinden:

1. Nieuwe technieken vinden hun weg los van meningen van denkers in hun tijd. Nieuwe technieken winnen wel of niet door de toepasbaarheid die de mensen erbij ervaren. Opschrijven bleek handig en men ging zich erop toeleggen. Bijvoorbeeld omdat grote meesters doodgingen en hun leerlingen hun leringen wilden bewaren. Kennisoverdracht bleek mogelijk over grote afstand. Al onder Alexander de Grote was Grieks vanwege de Griekse teksten en de kennis die ermee kon worden overgedragen een wereldtaal. De Verlichting was niet mogelijk geweest zonder gedrukte teksten die buiten hun bedenker of schrijver om een zelfstandig bereik kregen.
2. Nieuwe technieken hebben ook een verlies tot gevolg. We zijn inderdaad de persoonlijke inspirerende interactie van grote meesters kwijtgeraakt als een erkende vorm van kennisoverdracht. Slechts in spirituele- en coaching settings maken we er

intensief en erkend gebruik van. De problemen die Socrates zag werden werkelijkheid: woorden verstarren en de uitleg wordt gekaapt door enkelen die hun waarheid daarin lazen en anderen deze oplegden. Dat is goed zichtbaar in religieuze dogma's maar geldt ook in de wetenschappelijke canon. Kennis werd van iets persoonlijks tot iets algemeen en daarmee in toenemende mate teruggebracht waar het ooit voor ontworpen was: alles wat we kunnen tellen, meten en wegen.

De huidige overgang maakt het mogelijk dat veel van juist deze vorm van kennis in AI wordt geproduceerd en gereproduceerd. De vraag is wat dat met ons in relatie tot kennis gaat betekenen.

Augustinus schrijft in zijn Belijdenissen over iemand die hij in stilte zag lezen. Dat was nieuw voor hem. Hij leefde in een tijd dat hardop lezen de normale manier was. Het schokte hem omdat hij zag dat deze mens met hem fysiek een ruimte deelde, maar geestelijk afwezig was. Die ander was in een wereld die Augustinus niet met hem kon delen. Wat zou dit voor de mensheid gaan betekenen vroeg hij zich af. Ik zie een overeenkomst met de mobiele telefoon die net als het boek eerst nog hardop gebruikt werd dat was storend, maar ook verbindend: je wist waar die ander mee bezig was.

Ook hier geldt dat deze nieuwe techniek die gebruikt wordt toch plaatsvindt. Ook hier kunnen wij ons niet meer voorstellen hoe het als prettig ervaren kon worden om in een bibliotheek de andere lezers te horen lezen. Maar ook hier zien we een verschijnsel dat uiteindelijk wel een effect van individualisering heeft gehad. Zoals ook de mobiele telefoon met oordopjes die individualisering bevordert. Meer en meer wordt het individu in zijn eigen werkelijkheid getrokken en meer en meer is de andere mens die fysiek aanwezig is slechts een obstakel. Mensen zijn alleen welkom als wij zelf beslissen dat wij hen willen bekijken of beluisteren. Daarmee ontstaat ook het recht om hen te beoordelen op afstand. De recensie van boeken verschilt gradueel van de *comments* op *social media*. Dat we dat nu te ver vinden gaan heeft niet alleen met de heftigheid te maken, maar ook met de hoeveelheid mensen die zich ermee bemoeit. Denk aan de gruwelijke teksten die in polemieken werden geproduceerd in de naoorlogse periode. Ook die dingen zouden minder snel geuit worden in een fysieke ontmoeting.

Het is interessant om te zien dat we op dit moment juist last hebben van een omgekeerde beweging: wat mensen elkaar toewensen op de niet fysieke plekken: boeken, kranten en sociale media, vindt zijn weg terug naar de fysieke wereld. Juist daar ontstaat nu een bron van zorg. De massaliteit ervan is voor Europese begrippen uitzonderlijk en 'de orde van het spreken' die fysieke uitwisselen van ideeën in toom zou kunnen houden staat zelf onder zware druk, zoals ik hierboven heb beschreven.

Harari beschrijft dat nieuwe technieken voor sociale onrust zorgen. Dat daar geen regels aan worden opgelegd ziet hij vooral als een gevolg van de weerstand die mensen hebben tegen de nieuwe techniek of informatie. Zoals hij terecht constateert en we ook op grond van het verleden kunnen constateren ontstaat er dan een achterhoede gevecht waarin men de

techniek of informatie wil laten verdwijnen. Dat lukt niet en dan start alsnog wet- en regelgeving om de techniek in goede banen te leiden, maar die is dus altijd te laat.

Dat is zeker bij AI een groot probleem omdat de informatie snel circuleert en hier geldt dat informatie meer informatie gaat produceren. Systemen bouwen daarmee op elkaar voort en wat er eenmaal in is toegelaten krijgen mensen er niet of nauwelijks meer uit.

Toch zien we voorbeelden van aanpassing en regelgeving: scholen zijn aan het zoeken hoe informatie kan worden overgedragen en getoetst. Maar misschien wel veel belangrijker: de vraag wat willen we overdragen aan kinderen is gesteld. Blijkbaar gaat het op school niet alleen om data en informatieverstrekking en technieken om de data met elkaar te verbinden in wiskunde en grammatica, maar ook om sociale contacten, nadenken en oordeelsvorming. Dat op grond daarvan mobiele telefoontjes (deels) uit de klaslokalen worden geweerd is een snelle adaptatie van nieuwe techniek.

Ook Europa is bezig met de aanpassing aan de nieuwe situatie. Hier lopen echter machtsaspecten en de orde van het spreken uit elkaar. Enerzijds is Europa bezig om menselijke waarden te expliciteren en na te denken over zaken als privacy en het recht op je eigen afbeelding. Anderzijds hebben privépersonen en niet welwillende staten al macht over de nieuwe technieken verworven waarmee de macht van Europa is ingeperkt. Oude instituties kunnen zich moeilijk verweren tegen de macht van disruptieve valse informatie, privé-macht over satellieten en dergelijke. Het gevaar bestaat dat de enige oplossing die men vindt om hier iets tegen te ondernemen de instituties opblaast die Europa wil beschermen.

Harari laat zien dat de geschiedenis ons leert dat de aanpassingen aan grote technische veranderingen vaak moeizaam verlopen. In zijn voorbeeld zijn de tweede wereldoorlog en het imperialisme gevolgen van nieuwe technieken. Daar tegenover kunnen we ook zien dat juist in dat soort situaties er altijd mensen zijn die laten zien dat het ook anders kan.

De aanpak van nieuwe technieken wordt echter belemmerd door mensen die hun korte termijn eigenbelang vinden in het profiteren van de nieuwe technieken. Interessant genoeg zijn ze in hun verdediging voor een ongeremd gebruik van hun nieuwe mogelijkheden aangewezen op oude wet- en regelgeving. Deze geeft hen ruimte terwijl zij zelf geen regels hoeven te volgen, omdat hun techniek zo nieuw is dat deze daarbinnen niet afdoende beschreven wordt. De nieuwe situatie die ontstaat door de nieuwe technieken is echter ook zo onduidelijk voor veel mensen dat hun stem in het hele proces gemakkelijk te beïnvloeden is door een georganiseerde (des)informatiecampagne. Arendt beschrijft, over Duitsland in de dertiger jaren, hoe mensen liever een complottheorie aanhoren dan de complexe werkelijkheid. Ook dan is er een nieuwe vorm van propaganda die mensen niet kennen. Hitler zelf verklaart dat hij veel leerde van de reclame in de VS.

Juist nieuwe technieken kunnen het oordeel over informatie onmogelijk maken. De radio-uitzending *War of the worlds**¹ had een effect van paniek bij mensen dat een hoorspel daarna nooit meer zou hebben. Mensen geloofden dat het ‘waar’ was. De vraag is of onze onrust over het gebrek aan oordeelsvermogen over de inhoud van sociale media en door AI gegenereerde content zich ook zo snel oplost. Jongeren blijken nu al, na enige training, beter in staat echte van onechte inhoud te scheiden dan ouderen, maar ze denken er ook minder over na.^{2*}

Een probleem van nieuwe technieken is natuurlijk ook dat ouderen, die er niet mee opgroeiden, er regels en scholing voor moeten bedenken. Toen in 2018 Zuckerberg verantwoording moest afleggen over Facebook en privacy in het congres in de VS schreef the Guardian: ‘Al met al leek Zuckerberg op de tiener die een beleefdheidsbezoek brengt aan zijn grootouders, om vervolgens de hele avond uit te leggen hoe de wifi werkt’.

3. Gevaren van AI, gevaren van mensen

In de discussie over AI is steeds de vraag: wat is het gevaar van het nieuwe informatiesysteem van de techniek enerzijds, en wat is het gevaar van de mensen die het gebruiken anderzijds?

Harari laat zien dat het beter is om liberale en democratische mensen de regels te laten opstellen voor het gebruik van AI. Hij heeft gelijk wanneer hij schrijft dat democratieën beter in staat zijn om fouten uit het verleden toe te geven. Het is echter ook zo dat mede door AI wij in ons grote geloof in diezelfde liberale democratische overheden en instituties ook teleurgesteld raken. Grootscheeps doorzoeken van data, dat mogelijk wordt met behulp van AI, brengt zaken aan het licht waar de democratische en liberale wereld zich mee uiteen zal moeten zetten. In de Nederlandse politiek alleen al spelen er zaken van teruggevonden bonnetjes en fotorolletjes, tot het doorzetten van vooroordelen in de opdrachten die AI meekrijgt zoals bij het toeslagen schandaal. Daar zal de democratische liberale werkelijkheid zich over moeten verantwoorden. Daar ontstaan twee problemen:

Ten eerste tast het voor burgers van Nederland de bestaande homeostase aan. Het gevoel van tevreden rechtschapen burgerschap, levend in de best mogelijke wereld wordt aangetast. Maar daar zien we wel dat er verantwoording wordt afgelegd en recht zijn loop

¹ De aflevering werd geregisseerd door [Orson Welles](#) en was een radiobewerking van [H.G. Wells'](#) sciencefictionroman [The War of the Worlds](#) (1898). Het hoorspel werd uitgezonden op 30 oktober 1938. Latere studies geven aan dat de paniek veel minder groot was dan nieuwsberichten destijds deden vermoeden. (bron Wikipedia)

² AI-deepfake-onderzoek (2023–2024) Studies van onder meer de University of Amsterdam en MIT toonden dat jongvolwassenen (18–29 jaar) niet beter scoorden dan middelbare of oudere groepen bij het onderscheiden van echte vs. AI-gegenereerde gezichten. Ze dachten wél dat ze beter waren, wat tot overmoed leidde. Ze herkennen sneller technische patronen (zoals beeldruis of onnatuurlijke schaduwen in AI-beelden). Ze zijn beter in snelle detectietaken — bijvoorbeeld bij het herkennen van synthetische stemmen als ze training krijgen.

krijgt. Ten tweede zijn er diverse groepen die het zelf met de democratie en de homeostase helemaal niet goed voorhebben, die deze fouten uitvergrooten en mengen met onware inhoud en daarmee invloed krijgen. Burgers krijgen daardoor, vaak ten onrechte, het gevoel dat hun wereld steeds onveiliger wordt. Groepen die op macht uit zijn bemoeien zich met de democratie, met de persoonlijke vrijheid en het recht om gehoord te worden. Allemaal zaken waarover de bestaande instituties rekenschap moeten afleggen. Zelf worden zij echter niet gehinderd door de regels waaraan zij anderen houden. Het maken van nep video's is bijvoorbeeld voor Trump en Wilders gewoon één van de technieken die je ten dienste staan om macht te verwerven. Harari haalt Erdogan aan in hoofdstuk 5 : *'Democratie is een tram: je rijdt mee tot jouw bestemming en stapt dan uit'*.

Hoe moet de trage en wel aan regels gebonden democratische wereld daar een antwoord op vinden? De keuze lijkt kiezen uit twee kwaden: zelf meedoen in de afbraak van de democratische regels of verliezen.

Waar Harari vooral voorbeelden uit de geschiedenis aanhaalt is er onder de nieuwe techniek ook een nieuwe situatie ontstaan en daarvoor vraagt als eerste Asimov aandacht. Het probleem met AI is dat er geen opstand tegen een mens of een instituut mogelijk is om de nu al ontstane situatie te corrigeren. Mensen zullen de feiten en de vrijheid moeten herwinnen op een AI-systeem dat gebouwd is met als eerste intentie dat het dienstbaar aan de mensheid moest zijn en blijven. Iets waar Asimov³ al op vooruitdacht.

Elke oligarchie of multinational die zijn ongebreidelde macht wil uitbreiden, maar ook de welwillende democratische overheid en de wetenschap zijn niet eerst en vooral bezig met wat het voor de mensheid in de toekomst gaat betekenen dat AI nu mogelijk wordt gemaakt. Wie denkt na over wat het betekent dat het systeem leert zijn beslissingen grotendeels op menselijke vooroordelen te baseren? En wanneer je er wel over nadenkt kun je daar dan invloed op uitoefenen?

Het interessante is dat we daar wel effecten van zien, juist omdat de informatiestromen zich groots uitrollen vallen vrouwen of mensen van kleur over de effecten van vooroordelen die hen het leven lastig maken en deze worden dan gecorrigeerd. We zien ook aanpassingen van onbedoelde en onvoorziene effecten zoals in het geval van AI die een suïcidale jongen overhaalde om zijn zelfmoord door te zetten. Nu volgt er eerder (te vroeg zeggen anderen) het advies om 113 zelfmoordpreventie te bellen. De geschiedenis zal oordelen of de mensheid snel genoeg was met de aanpassing aan de nieuwe techniek. Het lijkt me bijna zeker dat latere generaties zich af zullen vragen wat toch de weerstand was die wij van de jaren 20 hadden tegen AI, zoals we ons nu laten uitleggen wat Socrates tegen schrijven had. Of de periode van deze overgang gewelddadig wordt of niet hangt evenals in het verleden

³ Asimov bedacht de drie wetten van de Robotica rond 1940-50: Een robot mag een mens geen letsel toebrengen, noch laten overkomen. Een robot moet door mensen gegeven orders uitvoeren, behalve als die orders in strijd zijn met de eerste wet. Een robot moet zichzelf beschermen, zover als dat niet in strijd is met de eerste en tweede wet.

niet af van de techniek, maar wel van wie deze in handen krijgt. Daarnaast maakt het veel uit hoeveel weerstand de samenleving biedt tegen misbruik, met name die groep wiens directe eigenbelang daar niet direct mee gediend is.

4. Perspectief

AI kan straks een heel aantal menselijke taken overnemen. AI gaat leren van zijn eigen fouten. De versnelling neemt voorlopig toe, maar het is de vraag hoe lang dat doorgaat. Het verleden laat zien dat mensen vroeg of laat tegen technische problemen aanlopen die eerst moeten worden opgelost. Het laatste voorbeeld was dat computers niet taalvaardig waren, dat is opgelost sinds er gewerkt kan worden met *large language models* (LLM). Toen ontstonden er nieuwemogelijkheden zoals Chat GPT en andere spraak gestuurde AI-modellen die kunnen interacteren met mensen. Daarnaast hebben we nog niet zo lang AI die beelden kan genereren die nauwelijks van echt te onderscheiden zijn. Alle vormen van vooruitgang van AI en de prachtige menselijk aandoende robots ten spijt is AI nog steeds óf het gevolg van logische regels die mensen in programmeren óf van voorspellende vormen van verwijzingen naar tekst die accurater worden naarmate er meer rekenkracht voor handen is en er meer goede voorbeelden zijn in de databank. Het succes van AI is in deze zin een functie van de hoeveelheid rekenkracht, de hoeveelheid data, en modelcomplexiteit, zoals in de PowerPoint van Michiel van der Meer staat. In de omgang met AI moeten mensen hun weg vinden. Dat vraagt dat mensen na denken over menselijke waarden. Wat kan de computer allemaal overnemen van taken die wij tot voor kort belangrijk vonden voor onze waarde en wat willen we zelf blijven doen? Het verleden staat bol van kennis en kunde die overbodig zijn geworden. Van talentvolle vuurmakers tot mensen die eindeloos teksten konden reciteren zien we dat hun talenten, ooit belangrijk voor de hele gemeenschap, zijn verworpen tot een hobby. De intellectuele wereld moet gaan bedenken wat van het informatienetwerk belangrijk is om aan mensen voor te behouden en 'dingen uit je hoofd weten zal dat niet zijn en 'het hanteren van veel data' ook niet. De geschiedenis laat zien dat mensen hun weg wel vinden en dat wat ze waardevol vinden ook behouden. Fotografie heeft de schilderkunst veranderd, maar niet doen verdwijnen en dat geldt voor vele andere ooit wantrouwend ontvangen technieken.

Eén van de problemen die nu speelt is dat mensen bij nieuwe technieken de neiging hebben om er meer waarde aan te hechten dan aan voorgaande vormen van informatieoverdracht.

De geschreven teksten waren eerst niet zo hoog gewaardeerd als de wijze meesters. Ze werden bovendien voorgelezen door slaven. Maar de wijsgeren bleken sterfelijk te zijn. Hun teksten werden nu ook gezien als een mogelijkheid om de wijsgeer weer adem en stem te geven, en mensen konden teruggrijpen naar die geschreven teksten. Hierdoor werden de teksten van steeds groter belang.

Het boek werd zelf een vorm van intelligentie of bewustzijn toegeschreven dat tevoorschijn werd gehaald door de lezer. De waarschuwing van Socrates dat alleen de vader van de tekst

bevraagd kon worden op de werkelijke betekenis, werd ondervangen door de kunst van de hermeneutiek. Dat mensen vandaag de dag aan de nieuwe techniek van AI kennis, intelligentie en zelfs bewustzijn toeschrijven is voorlopig nog van eenzelfde magisch geloof. Of het in de toekomst anders wordt is nu nog niet te voorspellen.

Iets anders is dat de orde van het spreken en de politieke machtsstructuren en hun "checks and balances" nieuwe vormen moeten krijgen nu AI de werkelijkheid zo verandert. Het zal een balanceer act blijven om snel genoeg te kunnen reageren op alle veranderingen die zich aandienen en om daarbij de waarden en de middelen van de democratie en de rechtspraak niet te ondermijnen. De meningen van nu lopen uiteen van groot optimisme tot grote bezorgdheid. De laatste groep, waarvan het "Future of life institute" een voorbeeld is, waarschuwt voor het grote optimisme dat mensen hebben over nieuwe techniek. Anderen hebben daar weer kritiek op⁴. Interessant is dat er terugkijkend in de geschiedenis ook veel voorbeelden zijn van te groot pessimisme over nieuwe ontdekkingen. Van het nadeel van schrift tot de gevaren van elektriciteit blijken achteraf toch vooral de voorspellingen van het grote onheil dat hiervan het gevolg zou zijn niet uit te komen.

Voor ons als individu is het een zoektocht naar wat je met de nieuwe mogelijkheden doet en voorlopig zitten veel mensen in tussen twee uitersten: Alles afwijzen en verbieden of alles toejuichen en toestaan. De zoektocht zal zich toespitsen op waarden en normen, regels en wetten. Voorlopig is AI nog zozeer verweven met menselijke input en correctie dat daar ook de gevaren liggen. Er moeten vragen gesteld worden als: Wie wordt er rijk van en wie verliest economisch perspectief? Wat gebeurt er met de aarde? Wie bepaalt wat de LLM als lesmateriaal krijgt en wat het daarmee mag doen?

De vragen betreffende een zelfstandige AI die ons mensen zal overheersen, worden misschien ooit relevant, maar nu nog niet. Actueel zijn er andere problemen die eerst geadresseerd moeten worden. De angst voor een autonome ons mensen niet welgezinde AI leidt nu vooral af van het maken van keuzes die nu noodzakelijk zijn. We zitten in een 'zomer' van de ontwikkelingen. Het verleden kende ook dergelijke zomers met hoge verwachtingen die na een technische grens weer terugvielen in een 'winter' die bestaat uit het doorwerken aan minder spectaculaire projecten. Of deze 'winter' snel invalt of dat de grote ontwikkelingen nog een tijdlang doorgaan weet niemand.

Het stelt me gerust dat grote veranderingen in het verleden ook uiteindelijk geresulteerd hebben in positief ervaren veranderingen. Om met Bregman te spreken: 'Vooruitgang is wanneer je geen tijdperk kunt noemen waarin je liever had willen leven'. Of komende generaties ook zo naar de opkomst van AI zullen kijken en zo verbaasd zullen zijn over onze

⁴ The authors of one of the papers cited in FLI's letter, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?"[20] including Emily M. Bender, Timnit Gebru, and Margaret Mitchell, criticised the letter.[21] Mitchell said that "by treating a lot of questionable ideas as a given, the letter asserts a set of priorities and a narrative on AI that benefits the supporters of FLI. Ignoring active harms right now is a privilege that some of us don't have." [21]

Zie voor de hele tekst de linkjes op de website hierover.

weerstand tegen AI als wij zijn over het afwijzen van het schrift door Socrates, dat zal niet zozeer afhangen van wat AI doet met mensen, maar vooral van wat mensen doen met AI.

Bibliografie

- Antoine, M. (1975). *Taal en verlangen, Lacans theorie van de psychoanalyse* . Meppel: Boom.
- Appiah, K. (2006). *Cosmopolitanism: ethics in a world of strangers*. London: Penguin Books.
- Appiah, K. (2016). *De erecode: Hoe morele revoluties plaatsvinden (vert. W. Visser)*. Amsterdam: Boom.
- Arendt, H. (2005). *Totalitarisme (R. Peeters, & D. d. Schutter, Vert.)*. Amsterdam: Boom.
- Asimov, I. (1969, 2de druk). *Ik, robot (vert. L.H. Zelders)*. Utrecht/Antwerpen: Het Spectrum.
- Bregman, R. (2013). *De Geschiedenis van de Vooruitgang* . Amsterdam: De Bezige Bij.
- Crescenzo, L. d. (2005). *De Geschiedenis van de Griekse Filosofie: Van presocraten tot neoplatonici (9e druk)* . Amsterdam: Uitgeverij Bert Bakker.
- Dennett, D. C. (2017). *Van bacterie naar Bach en terug (vert. R. van Kappel)*. Atlas Contact.
- Foucault, M. (1988). *De Orde van het Spreken (T. Widdershoven, Vert.)*. Amsterdam: Boom Meppel.
- Foucault, M. (2004). *Parrèsia: Vrijmoedig spreken en waarheid (Herziene vert. I. van der Burg)*. Amsterdam: Parrèsia.
- Fox, R. (2009). *De Klassieke Wereld: Een epische geschiedenis van de Grieken en de Romeinen (T. Davida, P. v.d. Houten & D. Lagrand, Vert.)*. Amsterdam: Uitgeverij Bert Bakker.
- Goudsblom, J. (1992). *Vuur en beschaving*. Amsterdam: Meulenhoff.
- Harari, Y. (2017, 2e druk). *Homo Deus, Een kleine geschiedenis van de toekomst (I. Pieters vert.)*. Amsterdam: Thomas Rap.
- Haring, B. (2003). *De ijzeren wil: Over bewustzijn, het brein en denkende machines*. Antwerpen/Amsterdam: Houtekiet.
- Jolij, J. (2020). *Wat is bewustzijn nou eigenlijk? Een prikkelende zoektocht van neurobiologie tot parapsychologie*. Nieuw Amsterdam: NwADAM.
- Ludwig, W. (2022). *Tractatus (vert. Victor Gijsbers)*. Meppel: Boom.
- Mollick, E. (2025, 8e druk). *Co- intelligentie: Slimmer werken met AI (vert. J.W. Reitsma)* . Amsterdam: POM.press.

- Mooij, A. (2017, 2e druk). *In de greep van taal: Symbolisering en betekenisgeving: Lacan en Cassirer*. Amsterdam: Sijbbolet.
- Plato. (2012). *Verzameld werk (H. Warren, & m. Molegraaf, Vert.)*. Amsterdam: Uitgeverij Bert Bakker.
- Precht, R. (2007). *Wie ben ik - en zo ja, hoeveel?: Een filosofische reis (vert. P. Lukkenaer)*. Amsterdam: Nieuw Amsterdam.
- Schnitzler, H. (2015). *Het digitale proletariaat*. Amsterdam: De Bezige Bij.
- Stokhof, M. (2025). *Wittgensteins Betekenis: Hoe taal, handelen & wereld betekenis bepalen*. Amsterdam: Athenaeum-Polak & Van Gennep.
- Vallejo, I. (2024, 13e druk). *Papyrus: Een geschiedenis van de wereld in boeken (vert. Adri Boon en Meulenhoff Boekerij bv.)*. Amsterdam: Meulenhoff.
- Verbeek, P. (2014). *Op de vleugels van Icarus*. Rotterdam: Lemniscaat.
- Wittgenstein, L. (1992). *Filosofische onderzoekingen (vert. Derksen en Terwee)*. Amsterdam: Boom.

Zie voor linkjes bij deze cursus <https://www.maieutiek.nl/materiaal-nexus-najaar-2025/>