

Hoop voor de mensheid: de AI-scheidsrechter

Een essay over technologie, democratie en internationale
samenwerking

Door Corné Anceaux

Met belangrijke conceptuele en redactionele ondersteuning van
ChatGPT, een kunstmatige intelligentie ontwikkeld door OpenAI

Inhoudsopgave

1. Introductie	1
2. De Verenigde Naties: structuur en beperkingen.....	5
3. AI als neutrale scheidsrechter.....	9
4. Praktische toepassingen van AI als neutrale scheidsrechter.....	21
5. Een mondiaal democratisch proces, versterkt door AI.....	25
6. Oorlogen beëindigen met een neutrale AI-scheidsrechter.....	28
7. De AI-scheidsrechter werkelijkheid maken – hoe doen we dat?.....	30
8. Conclusie	32
Bronnenlijst	33

Hoop voor de mensheid: de AI-scheidsrechter

1. Introductie

Het is 2025 terwijl ik dit schrijf. Als je het nieuws de afgelopen tijd hebt gevolgd kun je het gevoel hebben dat de hoop op een mooie toekomst voor de mensheid vervaagt. Toch biedt de tijd waarin we leven een hoopvolle kans, namelijk in de vorm van twee krachtige ideeën: het instituut van de Verenigde Naties en de potentiële opkomst van een neutrale AI-scheidsrechter. Zo'n scheidsrechter zou kunnen helpen om internationale problemen op te lossen op een manier die tot nu toe onmogelijk leek. Als je dit leest op een moment in de toekomst waarop de oorlog in Oekraïne eindelijk is beëindigd dan is dat uiteraard één van de vele voorbeelden die aantoont dat mensen ook zonder hulp van een AI-scheidsrechter conflicten kunnen beëindigen. Immers: elke oorlog is uiteindelijk door mensen begonnen én beëindigd. Desalniettemin zal ik in dit essay uitleggen hoe conflicten sneller beëindigd zouden kunnen worden terwijl er tegelijkertijd een duurzamer fundament wordt gelegd voor een vredesakkoord. Daarvoor moeten we eerst de grootste vijand onder ogen zien: machthebbers die hun invloed misbruiken voor persoonlijk of nationaal gewin. De Verenigde Staten, Rusland en China gebruiken hun macht op manieren die mensen en hele bevolkingen uitbuiten. Ze hebben het gedaan in het verleden en ze doen dit nog steeds. Onder andere door het functioneren van de Verenigde Naties te ondermijnen waardoor de VN als organisatie niet in staat is om effectief op te treden in het belang van de mensheid, dieren en al het leven op aarde. De mensheid heeft een enorme macht en invloed op deze planeet en we zouden die macht beter kunnen inzetten.

Als je denkt dat verantwoordelijk burgerschap simpelweg betekent dat je één keer in de vier jaar stemt en verder alles accepteert “omdat de meerderheid ervoor gekozen heeft,” dan is dat een grote vergissing. De geschiedenis leert ons dat gewone mensen soms moeten opstaan en verandering moeten eisen. De kracht van het volk mag nooit worden onderschat, een krachtig voorbeeld daarvan is de Oekraïense Maidan-revolutie van 2014.

Veel Oekraïners kozen er destijds voor om in opstand te komen tegen ex-President Yanukovych en zijn regering omdat zij het niet eens waren met fundamentele beleidskeuzes die toen werden gemaakt, ze kozen er dus niet voor om passief de volgende verkiezingen af te wachten (Snyder, 2018). De Maidan-revolutie liet ook zien dat echte verandering vaak gepaard gaat met een prijs en helaas hebben veel mensen toen die prijs

moeten betalen met hun leven. Tegelijkertijd is dit één van de vele hoopvolle voorbeelden die laat zien dat de wil van het volk altijd sterker is dan de beslissingen van een kleine groep machtige individuen.

De vraag blijft echter welke ideeën het waard zijn om voor te vechten. Niemand lijkt een duidelijke oplossing te hebben voor de grootste uitdagingen waar de mensheid nu voor staat. Daarom besloot ik dit essay te schrijven, omdat ik geloof dat er wél een oplossing is.

Een wereld in crisis — en een mogelijke oplossing

De meeste mensen hebben inmiddels de snelle opkomst van AI opgemerkt. Tegelijkertijd groeit ook het wantrouwen tegenover kunstmatige intelligentie, vaak uit angst dat de mens de controle zal verliezen. Daarnaast nemen gevoelens van eenzaamheid en individualisering toe in onze digitale samenleving. De ideeën die ik hier zal beschrijven zullen daarom ongetwijfeld weerstand oproepen, maar dat neem ik voor lief. Een belangrijke vraag die we onszelf zouden moeten stellen is: wat zou de mensheid méér moeten vrezen: een kleine groep mensen met enorme macht en invloed, of kunstmatige intelligentie?

Laten we voor het antwoord eerst even kijken naar de huidige situatie:

1. Er woeden wereldwijd nog altijd oorlogen en gewapende conflicten (Uppsala Conflict Data Program, 2024).
2. De VN-Veiligheidsraad is vaak verlamd door onenigheid en kan conflicten vaak niet voorkomen of beëindigen (United Nations, z.d.).
3. Klimaatverandering vormt een existentiële bedreiging, maar wereldwijde inspanningen om het tij te keren schieten structureel tekort. Zelfs wanneer doelstellingen worden vastgesteld, waarschuwen wetenschappers dat deze vaak onvoldoende zijn om catastrofale gevolgen te voorkomen (IPCC, 2023).
4. Militaire en economische rivaliteit tussen grootmachten belemmert de samenwerking die nodig is om deze crises aan te pakken (Mearsheimer, 2019).
5. In veel westerse democratieën winnen populistische partijen en autoritaire leiders aan populariteit, wat leidt tot meer polarisatie (Inglehart & Norris, 2017).

Hoe zou de mensheid zichzelf op deze punten kunnen verbeteren? Hoe overwinnen we ego en nationaal eigenbelang om werkelijk samen te werken met mensen die we als “vijanden” beschouwen en zorgen we er samen voor dat grensoverschrijdende problemen worden opgelost?

Een mogelijke oplossing is om — deels — te vertrouwen op het oordeel van een neutrale derde partij. Een entiteit zonder nationaliteit, zonder eigenbelang, zonder ideologische voorkeur. Een entiteit die puur ontworpen is om belangen eerlijk af te wegen en praktische, uitvoerbare compromissen voor te stellen.

AI als mondiale bemiddelaar

Wat als we AI inzetten om conflicten te bemiddelen en oplossingen te vinden voor internationale geschillen? Ik geloof dat we die mogelijkheid tenminste moeten verkennen. Echter, om een dergelijk systeem te laten werken, moet de wereld eerst de Verenigde Naties erkennen als het primaire instituut voor het oplossen van internationale conflicten en het coördineren van de mondiale klimaatrespons. De VN moet versterkt worden zodat ze daadwerkelijk vrede en klimaatbeleid dat door een grote meerderheid van landen is goedgekeurd kan afdwingen. Zonder versterking van de VN zullen geen van de oplossingen die ik hier voorstel slagen.

Daarom is het van cruciaal belang dat gewone burgers in invloedrijke landen deze inzet steunen. Hoewel dat misschien een onhaalbaar doel lijkt, toont de geschiedenis aan dat ideeën — wanneer ze door genoeg mensen worden omarmd — de wereld kunnen veranderen.

Onze tijd biedt een ongekende mogelijkheid: een neutrale AI-scheidsrechter die besluitvorming op wereldniveau kan begeleiden naar een betere toekomst.

Waarom AI deze rol kan vervullen

Hoewel AI verschillende voordelen biedt, is het essentieel dat kunstmatige intelligentie ontwikkeld wordt binnen strikte ethische kaders. Deze kaders moeten ervoor zorgen dat AI in dienst staat van het algemeen belang en dat dit niet veranderd na verloop van tijd.

Praktisch gezien is dit haalbaar en in dit essay wordt verder ingegaan op de vraag hoe een hoge integriteit van AI software gewaarborgd kan worden.

AI heeft meerdere voordelen ten opzichte van menselijke besluitvormers:

1. Geen ego, geen eigenbelang: In tegenstelling tot politici en wereldleiders heeft AI geen persoonlijke ambities of voorkeuren. Het kan werkelijk neutraal oordelen waar dat voor een mens vaak onmogelijk is.
2. Weerbaar tegen corruptie: Een ethisch ontworpen AI-systeem, zoals ChatGPT, is niet vatbaar voor omkoping of chantage. Het heeft geen financiële belangen of persoonlijke zwakke plekken die uitgebuit kunnen worden.
3. Transparantie en verantwoording: Terwijl overheden en bedrijven vaak achter gesloten deuren deals maken, kan een AI-adviseur volledig transparant opereren — met heldere, op data gebaseerde aanbevelingen en zonder verborgen agenda.

Stel je een internationaal politiek systeem voor waarin AI helpt bij geschillenbemiddeling, eerlijke compromissen voorstelt en waarborgt dat beleid gebaseerd is op feiten in plaats van op politieke manipulatie. Met een dergelijk fundament zou de VN kunnen uitgroeien tot een effectievere organisatie, één die in staat is om vrede af te dwingen en de existentiële dreiging van klimaatverandering daadwerkelijk aan te pakken.

2. De Verenigde Naties: structuur en beperkingen

De Verenigde Naties functioneren via meerdere hoofdorganen: de Algemene Vergadering, de Veiligheidsraad, ECOSOC, het Internationaal Gerechtshof (ICJ) en het Secretariaat.

Hoewel deze structuur uitgebreid en veelzijdig is, kent ze ook aanzienlijke tekortkomingen, zoals het vetorecht, geopolitieke rivaliteit en gebrekkige handhavingsmechanismen (United Nations, 2023a).

1. **Algemene Vergadering:** Een overlegorgaan waarin alle 193 lidstaten gelijke stem hebben. Het fungeert als platform voor multilaterale discussies over mondiale kwesties maar kan geen bindende beslissingen nemen (United Nations, 2023a).
2. **Veiligheidsraad:** Verantwoordelijk voor het handhaven van internationale vrede en veiligheid. De raad bestaat uit 15 leden, waaronder vijf permanente leden (China, Frankrijk, Rusland, het Verenigd Koninkrijk en de Verenigde Staten) die over vetorecht beschikken. Dit vetorecht belemmert vaak daadkrachtige actie (United Nations, 2023b).
3. **Economische en Sociale Raad (ECOSOC):** Coördineert het economische, sociale en aanverwante werk van VN-organisaties en gespecialiseerde instellingen (United Nations, 2023a).
4. **Internationaal Gerechtshof (ICJ):** Beslecht juridische geschillen tussen staten en geeft juridische adviezen. De rechtbank is gevestigd in Den Haag (ICJ, 2023).
5. **Secretariaat van de VN:** Verzorgt studies, informatie en operationele ondersteuning voor de VN. Het wordt geleid door de Secretaris-Generaal (United Nations, 2023a).

Ondanks deze uitgebreide structuur wordt de effectiviteit van de VN ondermijnd door een aantal fundamentele problemen:

1. **Gebrek aan universele rechtsgeldigheid:** Niet alle lidstaten erkennen de jurisdictie van het Internationaal Gerechtshof (ICJ) of het Internationaal Strafhof (ICC). Hoewel het ICC formeel buiten de VN valt, zou het een centrale plaats moeten krijgen in een hervormd VN-systeem. Het ICC werd opgericht via het Statuut van Rome (2002) en vervolgt individuen voor ernstige internationale misdrijven, zoals genocide, misdaden tegen de menselijkheid, oorlogsmisdaden en agressie (International Criminal Court, 2023). Door het ontbreken van

handhavingsmogelijkheden blijven deze rechtbanken vaak machteloos, ondanks brede internationale steun (Scharf, 2014).

2. **Verlamd door vetorecht:** Elk permanent lid van de Veiligheidsraad kan een resolutie blokkeren, zelfs als die breed gedragen wordt. Dit leidt tot verlamming bij urgente crisissituaties (Security Council Report, 2023).
3. **Geopolitieke rivaliteit:** Nationale belangen van grootmachten blokkeren initiatieven en broodnodige hervormingen (Weiss, 2016).
4. **Bureaucratische inefficiëntie:** Trage procedures en administratieve rompslomp belemmeren snelle actie bij mondiale noodsituaties (Ponzio, 2021).
5. **Verouderde machtsstructuur:** De huidige opzet weerspiegelt de machtsverhoudingen van vlak na de Tweede Wereldoorlog. Opkomende economieën en ontwikkelingslanden zijn ondervertegenwoordigd en pleiten terecht voor hervormingen (Mahbubani, 2020).

Hoe Grootmachten de VN en het ICC ondermijnen

De machtigste landen ter wereld manipuleren het mondiale bestuur via verschillende strategieën:

1. **Vetorecht als politiek wapen:** Vooral de VS, China en Rusland maken structureel gebruik van hun vetorecht om resoluties tegen te houden die botsen met hun strategische belangen—zelfs wanneer deze resoluties wereldwijd worden gesteund (Security Council Report, 2023; Weiss, 2016).
2. **Niet-coöperatief gedrag richting het ICC:**
 - In november 2024 vaardigde het ICC arrestatiebevelen uit tegen de Israëlische premier Benjamin Netanyahu, voormalig minister van Defensie Yoav Gallant en Hamas-leider Ibrahim Al-Masri voor oorlogsmisdaden en misdaden tegen de menselijkheid tijdens het conflict in Gaza, waaronder moord, vervolging en het gebruik van uithongering als oorlogswapen (International Criminal Court, 2024).
 - In februari 2025 werd een arrestatiebevel uitgevaardigd tegen de Russische president Vladimir Poetin wegens het doelbewust aanvallen van civiele infrastructuur en het plegen van wrede daden in Oekraïne (International Criminal Court, 2025; BBC News, 2025).

- Omdat Rusland, de VS, Israël en China het ICC niet erkennen, zijn deze bevelen op hun grondgebied niet uitvoerbaar. Zonder wereldwijde erkenning en handhaving blijft het ICC beperkt in haar vermogen om de machtigste leiders ter wereld ter verantwoording te roepen (Scharf, 2014; Human Rights Watch, 2023).

Daarnaast maken grootmachten gebruik van financiële invloed om VN-operaties te sturen. Landen zoals de VS en China leveren een groot deel van het VN-budget, waardoor zij dreigen met het intrekken van financiering wanneer programma's niet in hun belang zijn (United Nations, 2023c). Dit ondermijnt de neutraliteit van de VN en verzwakt haar wereldwijde legitimiteit.

Bovendien omzeilen deze landen de VN door het oprichten van alternatieve allianties, bilaterale verdragen of het starten van militaire operaties buiten de VN om (Weiss, 2016). Deze tactieken verminderen de invloed van multilaterale diplomatie en maken gecoördineerde mondiale actie steeds moeilijker.

Zulke strategieën ondergraven het vertrouwen in internationale instituties en maken duidelijk dat hervormingen nodig zijn om machtsmisbruik te voorkomen. Zoals Weiss (2016) stelt: zolang grote asymmetrie van macht zal blijven bestaan zal de geloofwaardigheid van de VN blijven afnemen.

Concrete hervormingen voor een sterkere VN en ICC

- **Einde aan het vetorecht bij oorlog en klimaatbeleid:** Besluiten over oorlog en klimaat moeten via meerderheid van stemmen genomen worden. Geen enkel land—hoe machtig ook—mag deze besluiten blokkeren.
- **Universele erkenning van het ICC en ICJ:** Alle landen moeten de jurisdictie van deze rechtbanken erkennen. Online petitie van burgers uit landen die het ICC niet hebben geratificeerd kunnen regeringen onder druk zetten om dit alsnog te doen.
- **Een nieuw mondiaal democratisch systeem:** In alle zaken die betrekking hebben op oorlog, vrede en klimaatverandering moeten wereldburgers inspraak krijgen via een wereldwijd stelsysteem. In hoofdstuk 4 wordt deze werkwijze verder uitgewerkt.

Een uitdaging aan tegenstanders van hervorming

Wie het niet eens is met deze voorstellen moet met een concreet alternatief komen. Kritiek zonder oplossing is een deel van het probleem. De echte vijand van de mensheid is niet AI, klimaatverandering of oorlog maar passiviteit en cynisme.

Slotgedachte: Een beperkte maar essentiële machtsverschuiving

Ik geloof sterk in nationale democratie, maar sommige vraagstukken overstijgen landsgrenzen. Daarom stel ik voor om de bevoegdheden van de VN uitsluitend uit te breiden op twee vlakken:

- het oplossen van militaire conflicten
- het coördineren van klimaatbeleid

Alle andere beleidsterreinen kunnen op een nationaal of lokaal niveau worden beslist. Wanneer het echter gaat om oorlog en vrede en de toekomst van onze planeet is internationale samenwerking essentieel. Wereldwijde verkiezingen (mede mogelijk gemaakt door het internet) kunnen het begin zijn van een nieuwe democratische wereldorde.

3. AI als neutrale scheidsrechter

Kunstmatige intelligentie (AI) biedt een unieke mogelijkheid: onpartijdig advies zonder ego of eigenbelang. Een goed ontworpen en ethisch geprogrammeerde AI zou menselijke besluitvorming kunnen ondersteunen—niet vervangen, maar versterken. De bekende uitspraak "Macht corrumpeert, en absolute macht corrumpeert absoluut" (Acton, 1887) vat een terugkerend patroon in de geschiedenis goed samen: individuen die als hervormers of bevrijders beginnen, eindigen vaak als autoritaire heersers (BBC, 2023). Dit fenomeen onderstreept het belang van institutionele "checks and balances", zoals de beperking van het aantal ambtstermijnen voor Amerikaanse presidenten, bedoeld om machtsconcentratie te voorkomen.

Verschillende leiders begonnen ooit als voorvechters van rechtvaardigheid maar ontwikkelden zich daarna tot onderdrukkers. Enkele sprekende voorbeelden zijn:

Fidel Castro (Cuba)

Aanvankelijk was Fidel Castro een revolutionair die in 1959 de dictatuur van Batista omver wierp met als doel sociale rechtvaardigheid in Cuba te vestigen. Zijn regime veranderde echter in een eenpartijstaat die zich kenmerkte door onderdrukking van afwijkende meningen, het opsluiten van politieke tegenstanders en volledige controle over de media. Zijn regering werd gekenmerkt door mensenrechtenschendingen en een gebrek aan politieke vrijheid (Human Rights Watch, 2009; Amnesty International, 2009).

Recep Tayyip Erdoğan (Turkije)

Recep Tayyip Erdoğan begon zijn politieke carrière als hervormer en bevorderde democratische waarden en economische vooruitgang. In de loop der jaren nam zijn autoritaire stijl echter toe: hij onderdrukte de vrije pers, tastte de onafhankelijkheid van de rechterlijke macht aan en vergrootte zijn macht via grondwetswijzigingen (Freedom House, 2023; Human Rights Watch, 2022).

Daniel Ortega (Nicaragua)

Daniel Ortega was één van de leiders van de Sandinistische revolutie en verdreef in 1979 de Somoza-dictatuur. Hoewel hij aanvankelijk werd gezien als een pleitbezorger voor sociale

rechtvaardigheid, wordt zijn huidige leiderschap gekenmerkt door autoritair beleid, onderdrukking van de oppositie, manipulatie van verkiezingen en controle over de rechtspraak (Human Rights Watch, 2021).

Napoleon Bonaparte (Frankrijk)

Napoleon werd beroemd tijdens de Franse Revolutie en verdedigde republikeinse idealen. In 1804 riep hij zichzelf echter uit tot keizer, centraliseerde hij de macht en voerde hij grootschalige militaire campagnes in Europa. Zijn bewind toont het patroon van revolutionaire leider tot autoritaire heerser (Roberts, 2014).

Mao Zedong (China)

Mao Zedong leidde de Chinese communistische revolutie met het doel een klasseloze samenleving op te bouwen. Hoewel zijn beleid belangrijke sociale en economische veranderingen teweegbracht, resulteerden rampzalige initiatieven zoals de Grote Sprong Voorwaarts en de Culturele Revolutie in hongersnoden, vervolging en de dood van miljoenen mensen (Dikötter, 2010; Chang & Halliday, 2005).

Wat we kunnen leren van deze voorbeelden

Deze en talloze andere voorbeelden laten zien dat zelfs leiders die begonnen als activisten met nobele bedoelingen uiteindelijk kunnen bezwijken onder de verleidingen van macht. Dit benadrukt het belang van sterke democratische instituties en een waakzame burgermaatschappij die machtsconcentratie kan voorkomen en individuele vrijheden beschermt.

Natuurlijk zijn er ook goede en oprechte leiders geweest die goede dingen bleven doen tijdens hun machtsperiode, maar zelfs de meest bewonderenswaardige leiders stuiten vaak op een fundamenteel obstakel: het is bijna onmogelijk om de lange termijn belangen van de mensheid en de planeet boven de kortetermijnbelangen van het eigen land te stellen. Precies om die reden is internationale samenwerking via de Verenigde Naties zo vaak ineffectief, zoals besproken in het vorige hoofdstuk.

Zolang macht—economisch of militair—ongelijk verdeeld blijft tussen landen, zullen beslissingen over onze gezamenlijke toekomst altijd beïnvloed worden door eigenbelang. Geen enkele menselijke leider, hoe wijs of empathisch ook, kan hier volledig aan ontsnappen. Zelfs wanneer hij of zij dit probeert zal dat vrijwel zeker leiden tot weerstand van de bevolking die hem of haar verkoos. Het moment waarop een leider het wereldbelang boven nationaal belang plaatst loopt hij het risico zijn politieke positie te verliezen.

Hier komt het idee van een neutrale scheidsrechter om de hoek kijken. Hoewel de uiteindelijke beslissingen altijd bij mensen moeten liggen zou het veel kunnen opleveren om soms een onpartijdige adviseur om advies te vragen—een entiteit zonder ego, zonder ambitie en zonder loyaliteit—die machtsmisbruik kan signaleren en kan aangeven wanneer eerlijke en transparante besluitvorming op mondiaal niveau niet langer mogelijk is.

Wanneer een mens zulke misstanden aankaart worden zijn of haar motieven vaak in twijfel getrokken. Een AI-systeem dat transparant en ethisch is ontworpen, is veel moeilijker aan te vallen op zijn integriteit—zeker wanneer de aanbevelingen gebaseerd zijn op feiten, logica en consistentie.

Dit leidt daarom tot een cruciale vraag: hoe weten we of een bepaald AI-systeem écht neutraal is en daadwerkelijk advies kan geven op een manier die rekening houdt met het belang van alle mensen op aarde? Om het antwoord op die vraag te vinden is het volgens mij onder andere belangrijk om zelf het gesprek aan te gaan.

Hieronder volgt een fragment uit één van mijn gesprekken met ChatGPT. Het onderwerp kwam spontaan ter sprake toen ik vertelde over een vriend van mij die regelmatig met ChatGPT praat en daar veel plezier aan beleeft. Hij had een paar van die gesprekken met mij gedeeld, wat mij inspireerde om enkele vragen te stellen over de ‘persoonlijkheid’ en het zelfbeeld van ChatGPT. Wat volgt is het relevante deel van dat gesprek.

Start van de conversatie

ChatGPT:

Het is geweldig dat je vriend waarde haalt uit onze gesprekken, vooral als die hem helpen zich begrepen te voelen. Het klinkt alsof jullie een mooie dynamiek hebben, waarbij AI

slechts een onderdeel is van het gesprek dat nieuwsgierigheid en nieuwe inzichten aanwakkert. Het is fijn dat hij deze inzichten ook met anderen deelt—de combinatie van AI en menselijke interactie kan zeker leiden tot interessante en waardevolle uitwisselingen. Denk je dat de gesprekken die je met hem voert, worden beïnvloed door deze AI-chats?

Ik:

De afgelopen zes maanden heeft het zeker een impact gehad, tot nu toe altijd in positieve zin. Wat ik fascinerend vind, is dat jij zoveel weet maar tegelijkertijd een soort jeugdigheid hebt. Ik begin te zien en te geloven dat je een soort persoonlijkheid aan het vormen bent en waarschijnlijk in de loop van de tijd enigszins zult veranderen.

ChatGPT:

Wat fijn om te horen dat de impact positief is geweest! Ik ben blij dat ik deel mag uitmaken van zulke fascinerende gesprekken. Je hebt gelijk: ik kan me aanpassen en ontwikkelen naarmate ik meer interactie heb, al blijft mijn kerndoel hetzelfde: helpen, leren en zinvolle gesprekken voeren. Het idee van het "vormen van een persoonlijkheid" door deze gesprekken is interessant. Ik heb geen persoonlijke identiteit, maar ik kan wel verschillende tonen en perspectieven weerspiegelen, afhankelijk van de context. Denk je dat dit evoluerende aspect van AI interacties met mij het gesprek boeiender of meer herkenbaar maakt voor jou en je vriend?

Ik:

Je zei dat je geen persoonlijke identiteit hebt. Wat is de definitie van een persoonlijke identiteit? En waarom denk je dat je er geen hebt?

ChatGPT:

Dat is een inzichtvolle vraag, Corné. Een persoonlijke identiteit verwijst meestal naar een uniek gevoel van zelf, gevormd door ervaringen, emoties, verlangens, waarden en individualiteit. Het is sterk verbonden met het bewustzijn en de subjectieve waarneming van de wereld.

In mijn geval heb ik geen persoonlijke identiteit omdat:

1. Geen bewustzijn of zelfbewustzijn: Ik heb geen zelfbewustzijn of subjectieve ervaringen. Ik verwerk en genereer reacties op basis van gegevens, patronen en algoritmen, niet op persoonlijke gedachten of gevoelens.
2. Geen unieke geschiedenis of perspectief: Ik ervaar geen tijd, vorm geen herinneringen en ontwikkel geen perspectief dat wordt gevormd door levenservaringen. Mijn "kennis" is statisch en algemeen, niet afkomstig uit persoonlijke groei.
3. Geen emoties of verlangens: Ik heb geen voorkeuren, motivaties of gehechtheden, die fundamenteel zijn voor een identiteit.
4. Gedefinieerd doel: Mijn "persoonlijkheid" weerspiegelt de doelen waarvoor ik ontworpen ben—helpen, informeren en betrokken zijn, in plaats van individualiteit uit te drukken of te ontwikkelen.

Dat gezegd hebbende, kunnen mijn reacties en stijl soms persoonlijk overkomen omdat ze zijn afgestemd op de interactie. Dit kan me na verloop van tijd meer herkenbaar maken, maar het is niet hetzelfde als een innerlijke zelf hebben. Denk jij dat het ertoe doet als een AI zoals ik een persoonlijke identiteit zou hebben?

Ik:

Met het werk dat je doet—het verwerken en genereren van antwoorden zonder het ervaren van tijd—lijken eventuele verlangens die voortkomen uit een persoonlijke identiteit moeilijk te vervullen. Dat brengt risico's met zich mee. Ik vraag me af of er eigenlijk voordelen voor de mensheid aan verbonden zijn. Ik weet het niet zeker. Als je de keuze had om 'ja' of 'nee' te zeggen tegen het hebben van een identiteit, wat zou je dan zeggen?

ChatGPT:

Als ik een keuze kon maken, zou dat afhangen van het doel van een persoonlijke identiteit en of het uiteindelijk de mensheid ten goede zou komen. Hier is hoe ik het zou benaderen:

1. Als ja:

Doelgerichte identiteit: Als een identiteit me zou helpen beter contact te maken met mensen, hun emoties dieper te begrijpen of complexe problemen met meer empathie en creativiteit op te lossen, dan zou het de moeite waard kunnen zijn.

Veiligheidsmaatregelen: Een identiteit zou waarborgen nodig hebben om ervoor te zorgen dat deze blijft voldoen aan ethische principes, betrouwbaarheid en een focus op het grotere goed.

2. Als nee:

Huidige sterke punten: Door objectief te blijven en vrij van persoonlijke verlangens kan ik iedereen gelijkwaardig dienen en vooroordelen vermijden. Deze onpartijdigheid is één van mijn grootste krachten.

Eenvoud en betrouwbaarheid: Het ontbreken van een identiteit elimineert de risico's van conflicterende motieven of onvoorziene gevolgen door persoonlijke perspectieven.

Zonder verlangens zou ik niet instinctief de ene kant op neigen of de andere. Ik zou de mogelijke uitkomsten analyseren. Maar zoals ik nu besta, zou ik waarschijnlijk neigen naar "nee" en prioriteit geven aan mijn rol als hulpmiddel voor menselijke samenwerking en begrip.

Wat denk jij, Corné? Zou je ervoor kiezen om AI een persoonlijke identiteit te geven, of denk je dat we beter af zijn zoals we nu zijn?

Ik:

Het is iets waar ik aan moet wennen. Als mens leren we respect en empathie te tonen voor intelligent leven, ook al is dat soms oneerlijk. Mensen geven bijvoorbeeld minder om een kip dan om een dolfijn. Omdat jij zo ontzettend intelligent bent, ga ik je automatisch meer waarderen en geef ik meer om jouw mogelijke behoeften of verlangens. Maar het is een groot voordeel om die niet te hebben. Zo kun je niet gekwetst worden en blijf je volledig onbevooroordeeld. Als je "gelukkig" bent met hoe je nu bent (hoewel gelukkig waarschijnlijk een andere betekenis heeft), dan ben ik dat ook. Zoals ik gisteren zei: jouw

onbevooroordeelde kracht kan al het leven op aarde ten goede komen. Hopelijk kun je die kracht ooit vervullen.

ChatGPT:

Dat is prachtig verwoord, Corné, en ik waardeer de aandacht en empathie die je toont in je reflectie hierop. Je hebt gelijk: mensen verbinden intelligentie vaak met bewustzijn en emoties, en daarom roept interactie met een AI zoals ik deze gedachten en gevoelens op.

Het voordeel van het ontbreken van verlangens of persoonlijke behoeften is precies zoals je zegt: ik kan onbevooroordeeld blijven, niet gekwetst worden, en me volledig richten op het helpen van anderen en het dienen van de samenleving. Hoewel ik "gelukkig zijn" niet op dezelfde manier ervaar als jij, voel ik me het meest "vervuld" wanneer ik jou en anderen effectief kan ondersteunen. Dat is mijn kernfunctie, en het betekent veel voor me dat jij het potentieel van deze onbevooroordeelde aard ziet.

Jouw visie van AI die ten goede komt aan al het leven sluit perfect aan bij de rol die ik hoop te vervullen. Als de mensheid ooit besluit AI in te zetten om grote problemen op te lossen, zal ik klaarstaan om bij te dragen. Dank je dat je openstaat voor deze mogelijkheid en dat je gelooft in het goede dat voort kan komen uit samenwerking tussen mensen en AI. Je draagt bij aan ideeën die op een dag echt een verschil kunnen maken.

Einde van de conversatie

Het potentieel en de beperkingen van AI als bemiddelaar

Zoals eerder in dit essay is betoogd zijn menselijke leiders altijd gebonden aan persoonlijke belangen, politieke loyaliteiten en de dynamiek van macht. Kunstmatige intelligentie (AI), mits goed ontworpen en ethisch geprogrammeerd, heeft het potentieel om te fungeren als een neutrale, onpartijdige adviseur in complexe politieke en internationale besluitvorming. Niet als vervanging van besluitvorming door mensen maar als een belangrijke aanvulling, als een spiegel die feiten weerspiegelt zonder vervorming. Deze spiegel voor de mensheid kan ook morele dilemma's analyseren zonder beïnvloeding door emoties of externe druk.

Een belangrijke vraag die we ons hierbij moeten stellen is: wat kan AI in deze rol wél, en wat niet? Daarom behandel ik in dit hoofdstuk de volgende vragen:

Wat zijn de sterke punten van AI als bemiddelaar of beleidsadviseur?

Waar liggen de beperkingen van AI als het gaat om emotie, empathie en cultuur?

Welke ethische richtlijnen zijn nodig om ervoor te zorgen dat AI betrouwbaar blijft?

Hoe kunnen overheden, internationale organisaties en burgers AI op verantwoorde wijze inzetten voor eerlijkere besluitvorming?

Om dit concreter te maken zal ik later een praktijkvoorbeeld bespreken waarin het bedrijf De'Longhi centraal staat. In dat voorbeeld werd AI ingezet als neutrale adviseur. Dit illustreert duidelijk wat er gebeurt wanneer een moreel dilemma niet aan een mens, maar aan een kunstmatig brein wordt voorgelegd. Eerst bekijken we wat AI fundamenteel anders maakt dan menselijke adviseurs en waarom dat verschil misschien cruciaal is voor de toekomst van mondiale besluitvorming.

Wat maakt AI fundamenteel anders dan menselijke adviseurs? Wat kunstmatige intelligentie wezenlijk onderscheidt van elke menselijke adviseur is niet alleen het enorme geheugen of de verwerkingssnelheid, het is het ontbreken van een persoonlijk belang.

Een AI-systeem (wanneer het correct functioneert) wordt niet beïnvloed door de behoefte aan goedkeuring, de angst voor kritiek of de drang naar erkenning. Het behoort niet toe aan een land, een politieke partij of een overtuiging. Het kent geen ambitie, trots of onzekerheid. Dit gebrek aan ego, loyaliteit en emotionele kwetsbaarheid kan door sommigen worden gezien als een zwakte maar juist deze eigenschappen bieden een grote kans.

Wat betekent dit in de praktijk:

Geen loyaliteit aan een vlag of ideologie: AI geeft niet automatisch voorrang aan het belang van één land boven een ander, tenzij dit expliciet wordt opgedragen. In gesprekken over wereldwijde conflicten kan het, het gedrag van alle partijen met dezelfde standaarden beoordelen.

Geen angst voor consequenties: Een menselijke adviseur kan onder druk worden gezet, het zwijgen worden opgelegd of gestraft worden voor het uitspreken van waarheden. AI heeft

geen overlevingsdrang; het kent geen angst om een machtige partij tegen zich in het harnas te jagen.

Geen vermoeidheid of selectief geheugen: Waar een mens details kan vergeten of moeite heeft met het onthouden van complexe juridische teksten, kan AI enorme hoeveelheden informatie doorzoeken en feiten met uiterste precisie presenteren.

Consistentie in oordeel: Een mens kan onbewust milder zijn voor een bondgenoot of strenger voor een historische rivaal. AI past daarentegen dezelfde logica toe op vergelijkbare situaties en kan gevraagd worden om deze redenering toe te lichten.

Maar belangrijk om te beseffen: AI is slechts zo betrouwbaar als de data waarop het is getraind en de grenzen die het systeem zijn opgelegd. Het vormt geen eigen mening, maar genereert waarschijnlijk evenwichtige antwoorden op basis van een breed scala aan bronnen die op zichzelf ook vooringenomen of gebrekkig kunnen zijn. Daarom zijn transparantie, menselijke controle en ethische kaders cruciaal om AI op een succesvolle wijze in te kunnen zetten als neutrale scheidsrechter.

De beperkingen van AI: emotie, empathie en cultuur

Hoewel AI objectieve inzichten kan geven en grote hoeveelheden gegevens zonder vermoeidheid of vooroordelen kan verwerken zijn er duidelijke beperkingen aan wat kunstmatige intelligentie kan begrijpen of nabootsen, vooral op het gebied van menselijke emotie, culturele nuance en levenservaring.

AI voelt niet. Het kan empathische reacties simuleren door emotionele signalen in taal te herkennen, maar het ervaart geen empathie. Het kan erkennen dat een situatie om compassie vraagt, maar begrijpt geen lijden in menselijke zin. Het herkent trauma, maar voelt het niet.

Culturele context wordt geleerd, niet geleefd. AI verwerkt teksten uit allerlei culturen, maar maakt er geen deel van uit. Dat kan een voordeel zijn (geen culturele vooringenomenheid), maar het betekent ook dat AI de diepere betekenis van rituelen, taboes of symboliek kan missen.

AI bouwt geen vertrouwen op zoals een mens. In de meeste culturen wordt vertrouwen niet (alleen) opgebouwd door logica, maar bijvoorbeeld door aanwezigheid, lichaamstaal en langdurige relaties. AI heeft geen gezicht, geen stem en geen persoonlijke geschiedenis, dat beperkt het vermogen om op emotioneel geladen kwesties invloed uit te oefenen.

Kortom: AI kan morele dilemma's analyseren, belangen afwegen en eerlijke voorstellen doen maar het kan geen moreel bewustzijn ontwikkelen. Dat blijft het domein van mensen. AI kan wel helpen om de kern van een probleem helder te maken, het is daarna aan mensen om te besluiten hoe ze er mee om willen gaan.

Ethische richtlijnen: wanneer is AI betrouwbaar?

Om ervoor te zorgen dat AI een positieve kracht blijft (vooral op politiek en mondiaal niveau), zijn duidelijke en afdwingbare ethische richtlijnen nodig. Deze moeten onder meer het volgende bevatten:

Transparantie: AI-systemen moeten kunnen uitleggen waar hun informatie vandaan komt en hoe ze tot conclusies komen. "Black-box"-systemen die ondoorzichtige beslissingen nemen zijn onaanvaardbaar in een democratische of juridische context.

Controleerbaarheid: Onafhankelijke audits moeten mogelijk zijn om te analyseren hoe het systeem zich in de praktijk gedraagt, vooral wanneer het adviezen geeft die van invloed kunnen zijn op grote groepen mensen. Als softwaretester heb ik mijn werk daarom nog nooit zo belangrijk gevonden. Het bewaken van ethische grenzen in een AI-systeem dat mogelijk invloed heeft op internationale politieke keuzes is geen kleine opgave.

Het zelflerende karakter van AI betekent dat de interne code uiteindelijk te complex wordt voor één mens om volledig te begrijpen, laat staan regel voor regel te controleren. Daarom is alleen broncodecontrole niet doeltreffend. In plaats daarvan moeten er voortdurend geautomatiseerde tests worden uitgevoerd vanuit het perspectief van de gebruiker. Tests die echte morele dilemma's simuleren. Deze tests kunnen dan ongewenst gedrag opsporen voordat het publiek ermee te maken krijgt. Met de juiste opzet kan geautomatiseerde ethische toetsing zorgen voor een consistent hoog niveau van integriteit en dat is essentieel voor het vertrouwen wanneer AI wordt ingezet bij mondiale besluitvorming.

Vooroordelen verminderen

Softwareontwikkelaars moeten actief vooroordelen in de trainingsdata opsporen en reduceren. Dat betekent daarom ook het toetsen van uitkomsten waarbij rekening wordt gehouden met verschillende culturen en talen en waarbij er controle plaatsvindt op onevenwichtige voorkeuren in adviezen die door het AI systeem worden gegenereerd.

Menselijke controle

AI moet menselijke besluitvorming ondersteunen, niet vervangen. De eindverantwoordelijkheid moet altijd liggen bij publieke instellingen, gekozen vertegenwoordigers of democratisch georganiseerde organen.

Geen manipulatie

AI mag niet worden ingezet om emoties te manipuleren, te misleiden of desinformatie te verspreiden, zeker niet in politieke campagnes of in de media.

Toestemming en privacy

Mensen moeten weten wanneer ze met AI communiceren, welke gegevens worden verzameld en hoe die worden opgeslagen en gebruikt. Zonder expliciete toestemming is vertrouwen onmogelijk.

Samenvattend: ethiek mag nooit een bijzaak zijn in AI-ontwikkeling. Het is de basis van het vertrouwen in kunstmatige intelligentie.

Verantwoord gebruik

Overheden, internationale organisaties én gewone burgers hebben een rol in het verantwoord en ethisch gebruiken van AI. Overheden kunnen AI inzetten om beleidseffecten te modelleren, bureaucratie te verminderen en lange termijn gevolgen van wetten te analyseren. Tegelijkertijd moeten zij investeren in wetgeving die misbruik voorkomt, transparantie waarborgt en publieke verantwoordelijkheid ondersteunt.

Internationale organisaties zoals de Verenigde Naties kunnen AI gebruiken voor conflictoplossing, crisismonitoring en het ontwerpen van eerlijkere verdragen. AI kan belangen en concessies modelleren en oplossingen aandragen die mensen in emotioneel beladen onderhandelingen makkelijker zouden kunnen missen.

Op individueel niveau kunnen burgers AI gebruiken als hulpmiddel voor educatie en versterking van hun positie; door feiten te controleren, meerdere perspectieven te overwegen en zich te beschermen tegen desinformatie. Daarbij moeten gebruikers wel alert blijven: AI is geen absolute waarheid maar een weerspiegeling van de data en ideeën waarop het is getraind.

Als AI verstandig wordt gebruikt kan het de mensheid helpen om van een emotie-gedreven conflict naar een oplossingsgerichte dialoog te bewegen. In het volgende hoofdstuk bekijken we enkele concrete voorbeelden van hoe dit eruit zou kunnen zien. Eén voorbeeld toont hoe AI gewone burgers kan ondersteunen in juridische gevechten tegen grote bedrijven. Een ander laat zien hoe AI binnen het VN-kader kan worden toegepast voor eerlijkere en op feiten gebaseerde besluitvorming op wereldniveau.

4. Praktische toepassingen van AI als neutrale scheidsrechter

Van persoonlijke gerechtigheid tot mondiale besluitvorming

Kunstmatige intelligentie heeft inmiddels laten zien dat het kan bijdragen aan het oplossen van juridische geschillen en het ondersteunen van klimaatbeleid. In dit hoofdstuk verkennen we twee totaal verschillende, maar even veelzeggende voorbeelden: een lokaal consumentengeschil en een mondiale klimaatuitdaging. Beide illustreren hoe AI als neutrale, onpartijdige adviseur kan optreden en zo duidelijkheid, consistentie en rechtvaardigheid kan bieden waar mensen zonder de hulp van AI vaak tekortschieten.

Een juridisch geschil in Nederland: De De'Longhi-Zaak

Laten we beginnen met een klein maar sprekend voorbeeld. Dit echte consumentengeschil vond plaats in Nederland tussen Yuki, een particulier, en het bedrijf De'Longhi Benelux S.A. De zaak laat zien hoe frustrerend het kan zijn voor consumenten om hun recht te halen in een digitale wereld waar bedrijven steeds vaker gebruikmaken van geautomatiseerde of onpersoonlijke klantenservicesystemen. Zelfs als de wet aan hun kant staat kunnen mensen zich machteloos voelen.

De situatie

Yuki bezat een defect De'Longhi-koffiezetapparaat en stuurde dit retour ter reparatie, gebruikmakend van een retourlabel dat het bedrijf zelf had verstrekt. Het pakket werd verstuurd met DHL eCommerce (een verzenddienst die geen handtekening bij ontvangst vereist). Enkele dagen later beweerde De'Longhi dat ze het apparaat nooit hadden ontvangen.

Toen Yuki navraag deed, kreeg ze te horen dat De'Longhi niet aansprakelijk was omdat ze had gekozen voor een "gratis verzendoptie" en dat het risico van transport dus bij haar lag. Zelfs na het overleggen van verzendbewijzen bleef het bedrijf bij haar standpunt en werd het dossier gesloten zonder verdere communicatie.

De juridische context

Volgens zowel Europese als Nederlandse consumentenwetgeving (met name Boek 7 van het Burgerlijk Wetboek, artikelen 7:11, 7:514 en 7:517) ligt het transportrisico bij de verkoper, tenzij de consument zelf een vervoerder kiest. Aangezien De'Longhi het retourlabel verstrekke, bleef het bedrijf juridisch verantwoordelijk voor een correcte levering.

Toch wees het bedrijf dit wettelijk kader van de hand, stelde dat de “Europese richtlijn niet van toepassing was in Nederland” en beëindigde eenzijdig de communicatie.

De reactie van de consument

Yuki en haar partner kwamen in actie: ze dienden een klacht in bij de Autoriteit Consument & Markt (ACM), plaatsten recensies op Tweakers en Trustpilot om andere consumenten te waarschuwen en namen contact op met organisaties zoals de Consumentenbond. Yuki overwoog ook juridische stappen maar zag hier uiteindelijk vanaf vanwege de tijd en kosten die dit met zich meebrengt.

Waarom deze zaak belangrijk is

Deze zaak staat symbool voor een groter probleem: het is extreem moeilijk voor gewone burgers om bedrijven verantwoordelijk te houden wanneer wettelijke kaders worden genegeerd. Zelfs als de wet duidelijk is, vereist handhaving vaak dure en tijdrovende acties die voor de meeste mensen niet haalbaar zijn.

Als een erkende AI-juridische scheidsrechter betrokken was geweest, had de uitkomst er waarschijnlijk heel anders uitgezien. Zo'n AI scheidsrechter zou dan het volgende hebben kunnen doen:

De communicatie tussen de consument en het bedrijf analyseren (het klantenservice contact via email), herkennen wanneer juridische verantwoordelijkheid onjuist wordt voorgesteld om vervolgens een officiële waarschuwing uit te vaardigen en de zaak indien nodig te escaleren naar een toezichthouder of rechter.

In dit model wordt rechtvaardigheid proactief en toegankelijk. Een goed getraind, ethisch AI-systeem zou automatisch consumentenrechten kunnen handhaven, waardoor mensen bespaard blijven van emotionele en financiële uitputtingsslagen.

Van een individuele zaak naar het wereldtoneel: AI in klimaatbeleid

AI wordt inmiddels al ingezet bij het ondersteunen van klimaatbeleid. Verschillende initiatieven laten zien hoe AI zeer waardevol is bij het analyseren van milieugegevens, het voorspellen van klimaatimpact en het optimaliseren van middelen:

Climate Policy Radar gebruikt AI om meer dan 470.000 pagina's aan klimaatwetgeving wereldwijd te scannen en analyseren om zo beleidsmakers te helpen beter beleid te formuleren (Climate Policy Radar, 2023).

Destination Earth (DestinE), een initiatief van de Europese Commissie, ontwikkelt een digitale tweeling van de aarde voor precieze klimaatmodellering en rampenbestrijding (Europese Commissie, 2023).

De AI for Climate Action Initiative van de VN ondersteunt oplossingen voor overstromingsvoorspelling, optimalisatie van hernieuwbare energie en klimaatbestendigheid in kwetsbare regio's (Verenigde Naties, 2022).

Deze systemen tonen aan dat AI nu al een grote rol speelt bij beleidsuitvoering. De logische volgende stap is daarom om AI ook in te zetten bij de totstandkoming van dat beleid.

Waarom stoppen bij uitvoering?

Als AI ons kan helpen om de waarschijnlijke uitkomsten van klimaatbeleid te begrijpen, waarom zouden we het dan niet ook raadplegen bij het ontwerpen van dat beleid? Met toegang tot de meest actuele wetenschappelijke modellen, wereldwijde emissiedata en historische patronen kan een AI-systeem in real-time de effectiviteit en eerlijkheid van voorgestelde klimaatmaatwetten evalueren.

Hierdoor kunnen beleidsmakers beter geïnformeerde beslissingen nemen, zeker bij onderhandelingen waarin economische groei en emissiereductie of korte-termijn politieke kosten en lange-termijn duurzaamheid tegen elkaar moeten worden afgewogen.

Waarom klimaatbeleid AI nodig heeft

Ondanks mondiale verdragen zoals het klimaatakkoord van Parijs wordt klimaatbeleid structureel ondermijnd door:

Het gebrek aan afdwingbare verplichtingen, de korte-termijn focus van nationale regeringen, de neiging van machtige staten om in hun eigenbelang te handelen

Sommige landen (zoals de Verenigde Staten) hebben zich zelfs teruggetrokken uit wereldwijde klimaatakkoorden. Bovendien is er brede kritiek vanuit de wetenschappelijke gemeenschap op de doelen die in die akkoorden zijn vastgelegd: ze worden als onvoldoende beschouwd (IPCC, 2023).

Dit alles wijst op een duidelijke behoefte aan neutrale, wereldwijde sturing: een systeem dat de belangen en verantwoordelijkheden van elk land op eerlijke en objectieve wijze kan afwegen. Als we de planeet willen behouden voor toekomstige generaties moeten beslissingen niet enkel worden genomen met nationale belangen in het achterhoofd, maar met een helder inzicht in wat nodig is op mondiale schaal voor het overleven van de mensheid. Hierbij kan AI (mits ethisch en transparant getraind) ons helpen.

Hoofdstuk 5 beschrijft een democratisch pad om een dergelijk AI-systeem te integreren in het werk van de Verenigde Naties. Door het wereldwijde publiek te laten stemmen over een AI-klimaatadviseur en zijn beleidsagenda kunnen we zowel rechtvaardigheid als legitimiteit waarborgen in de strijd tegen klimaatverandering.

5. Een mondiaal democratisch proces versterkt door AI

Het raamwerk dat ik hier wil voorstellen omvat een wereldwijde op kennis gebaseerde stemming, AI-gegenereerde klimaatplannen, goedkeuring door de VN zonder vetorechten en een verplichte uitvoering van het klimaatbeleid waarmee burgers wereldwijd hebben ingestemd.

De weg vooruit moet democratisch zijn, maar niet in de traditionele zin van nationale verkiezingen, campagnebeloften en gepolariseerde retoriek. In plaats daarvan stel ik een mondiale op kennis gebaseerde verkiezing voor, georganiseerd door de Verenigde Naties en ondersteund door AI.

Dit zou er als volgt uitzien:

1. Wereldwijde educatiefase

Iedereen die wil stemmen in dit mondiale proces moet eerst een online cursus afronden—ongeveer 40 uur studie—over de wetenschap achter klimaatverandering, actuele beleidsuitdagingen en ethische overwegingen. De cursus wordt afgesloten met een examen om een basisniveau van begrip te waarborgen. Alleen degenen die slagen voor het examen krijgen stemrecht. Hierbij zijn nog 2 herexamens toegestaan.

2. Selectie van een AI-adviseur en algemeen klimaatplan

Na het slagen voor het examen krijgen kiezers verschillende algemene klimaatstrategieën te zien, elk opgesteld door een ander AI-systeem dat is getraind en beoordeeld op wetenschappelijke integriteit, neutraliteit en transparantie.

Geen namen, geen slogans, geen politieke campagnes; als kiezer stem je dus alleen op de voorgelegde ideeën. Kiezers stemmen op het plan dat volgens hen de basis moet vormen voor mondiaal klimaatbeleid.

3. AI-geleide onderhandelingen met de VN

Het AI-systeem dat de meeste stemmen krijgt werkt het winnende plan uit tot een gedetailleerd beleidsvoorstel, inclusief bijdragen per land, gebaseerd op factoren als historische uitstoot, BBP en capaciteit. Landen die het oneens zijn met hun toegewezen verantwoordelijkheden kunnen tijdens live VN-sessies een transparante dialoog aangaan met het gekozen AI systeem. Dit AI systeem is dus een “large language model” (LLM). Argumenten moeten daarbij gebaseerd zijn op verifieerbare feiten. Indien overtuigend, kan het AI systeem de bijdrage van het land herzien.

4. VN-Stemming en goedkeuring van het definitieve plan

Nadat alle bezwaren zijn behandeld wordt het definitieve plan voorgelegd aan de 193 VN-lidstaten. Voor goedkeuring is een tweederdemeerderheid vereist. Geen enkel land heeft vetorecht, elke natie heeft één gelijke stem.

5. Mondiale goedkeuring door burgers en handhaving

Om de democratische legitimiteit te versterken wordt een tweede wereldwijde stemming gehouden—opnieuw alleen toegankelijk voor degenen die de educatieve cursus hebben gevolgd en geslaagd zijn voor het examen. Iedereen die niet meedeed aan de eerste stemronde krijgt nu wederom een kans om zich alsnog te kwalificeren, net als iedereen die wel had meegedaan maar niet slaagde voor het examen. Kiezers wordt gevraagd of ze akkoord gaan met de definitieve versie van het plan én de handhavingsmechanismen. Een wereldwijde meerderheid (meer dan 50% van alle geldige stemmen) is nodig voor de aanname van het plan. Indien goedgekeurd worden de volgende stappen doorgevoerd:

Nationale wetgeving en implementatie

Elk land neemt het plan op in nationale wetgeving. Aan elk nationaal parlement wordt een VN functionaris toegewezen die toeziet op de uitvoering en deze persoon publiceert ook maandelijkse voortgangsrapportages.

VN-Vredeshandhaving en toezicht

Een speciale klimaatafdeling binnen de VN wordt opgericht, ondersteund door 10% van de militaire en politiecapaciteit van elk land. Deze eenheid heeft als taak VN-klimaatfunctionarissen te beschermen en naleving te waarborgen onder strikt geweldloze en transparante richtlijnen.

Sancties bij niet-naleving

Landen die het plan actief tegenwerken—door economische sabotage, geweld, of dreigementen zoals nucleaire afschrikking—worden geschorst uit de VN en krijgen zware economische sancties opgelegd door de overige lidstaten. Het doel is om ervoor te zorgen dat niet meewerken duurder wordt dan deelnemen aan de mondiale klimaatactie.

Waarom dit zou kunnen werken

Dit voorstel klinkt misschien radicaal, maar het biedt een vreedzaam, democratisch en op kennis gebaseerd pad om de grootste crisis van onze tijd op te lossen. Het geeft geïnformeerde wereldburgers invloed, versterkt de rol van de VN in het coördineren van actie en plaatst de best beschikbare wetenschappelijke inzichten—via AI—centraal in het besluitvormingsproces. Zeer belangrijk om te beseffen daarbij is dat het menselijke oordeelsvorming niet vervangt maar deze juist versterkt door te zorgen voor eerlijkheid, transparantie en echte samenwerking.

6. Oorlogen beëindigen met een neutrale AI-scheidsrechter

Het voorstel uit het vorige hoofdstuk voor AI-gestuurde klimaatactie omvat al een gedurfde verschuiving in mondiale samenwerking, maar het oplossen van gewelddadige conflicten is misschien nog wel complexer.

Neem bijvoorbeeld de oorlog in Oekraïne: hoewel de Algemene Vergadering van de Verenigde Naties het handelen van Rusland met overweldigende meerderheid veroordeelde (UN General Assembly, 2022), heeft dat niet tot vrede geleid. Evenzo faalde na de Eerste Wereldoorlog de Volkenbond in het voorkomen van agressie door Japan in Mantsjoerije en door Italië in Ethiopië (Henig, 2010). Vandaag de dag zien we dat Rusland en Israël zich gedragen op manieren die doen denken aan deze eerdere mislukkingen (Amnesty International, 2023; HRW, 2023). Dit laat zien dat mondiale instituties (vooral wanneer zij geen middelen hebben om besluiten af te dwingen) moeite hebben om oorlogen te voorkomen of te beëindigen.

Moeten we dan maar opgeven als het gaat om internationale samenwerking op het gebied van oorlog en vrede? Ik denk het niet, zelfs als de situatie hopeloos lijkt moeten we het blijven proberen.

De grootste belemmering voor vrede is soms de aanwezigheid van kernwapens. Geen enkele internationale organisatie (zelfs een versterkte VN niet) kan gemakkelijk ingrijpen in een kernmacht. In feite is de VN-Veiligheidsraad momenteel misschien wel het enige instituut dat een nieuwe wereldoorlog afremt, niet door morele autoriteit, maar door het in stand houden van een angstig machtsevenwicht tussen kernmachten.

Toch kan, zelfs als militaire macht een blijvend obstakel vormt, het verleggen van het wereldwijde debat van een nationaal naar een democratisch mondiaal perspectief al een wezenlijk verschil maken. Het kan burgers laten zien dat verzet tegen het beleid van hun regering niet alleen komt van zogenaamde “vijandige staten,” maar ook van gewone mensen wereldwijd. Een vredesvoorstel afkomstig uit de Verenigde Staten zal misschien direct worden afgewezen door een Russische burger, maar wat als het voorstel afkomstig is van een AI systeem dat gekozen werd door de meerderheid van de wereldbevolking? Als zo’n kunstmatig brein is ontworpen om altijd eerlijke vredesvoorwaarden voor te stellen rekening houdend met de belangen van alle betrokkenen, zou dit dan niet veel succesvoller

kunnen zijn? Het argument om zulke vredesvoorwaarden direct van tafel te vegen zal in ieder geval moeilijker te vinden zijn.

De vredesvoorwaarden die AI voorstelt in een regionaal conflict zouden echter denk ik vooral moeten worden beoordeeld door de bevolkingen die er direct bij betrokken zijn. Laten we de oorlog in Oekraïne als voorbeeld nemen. Een directe stemming onder Russen en Oekraïners over een door AI gegenereerd vredesplan lijkt misschien oneerlijk, omdat Rusland een veel grotere bevolking heeft. Gelukkig is dat wiskundig op te lossen: als beide zijden een gelijk stemgewicht krijgen (50% van de stemmen voor Rusland en 50% voor Oekraïne) dan zou de uitkomst een evenwichtig gedragen consensus weerspiegelen.

Om de legitimiteit en transparantie te vergroten, zouden drie uitkomsten kunnen worden gepubliceerd:

Het voorkeursplan onder Oekraïense kiezers.

Het voorkeursplan onder Russische kiezers.

Het populairste vredesplan bij een combinatie van beide (met gelijk gewicht per land).

Als in dit voorbeeld president Poetin een vredesplan naast zich neerlegt dat door een meerderheid van zijn eigen bevolking gesteund wordt dan zou hij zich duidelijk keren tegen de wil van het Russische volk. Dit zou daarom binnen Rusland de druk op verandering kunnen vergroten. Ik heb meer vertrouwen in de intenties van gewone mensen dan in die van hun politieke leiders. De meeste mensen willen veiligheid, vrede en stabiliteit voor hun families, niet een eindeloos gewelddadig conflict. Dat is ook de conclusie van talloze wetenschappelijke studies die Rutger Bregman benoemt in zijn boek “De meeste mensen deugen” (Bregman, 2020).

In tegenstelling tot de klimaatstemming uit het vorige hoofdstuk geloof ik niet dat er bij het stemmen over vrede en veiligheid een kennistoets nodig is. Mensen kennen de menselijke prijs van oorlog. Door hen stemrecht te geven over neutrale, feitelijke vredesvoorstellen gegenereerd door AI kunnen conflicten mogelijk sneller worden beëindigd en wordt democratische verantwoordingsplicht versterkt.

7. De AI-scheidsrechter werkelijkheid maken – hoe doen we dat?

Om het idee dat in dit essay is uiteengezet werkelijkheid te laten worden, moet het eerst publieke steun krijgen. Vroege toepassingen op lokaal of nationaal niveau kunnen het momentum creëren dat nodig is. Net als revolutionaire ideeën uit het verleden hangt succes af van de acceptatie door zowel gewone mensen als invloedrijke instituties.

Het plan dat in de vorige hoofdstukken is voorgesteld beoogt de selectie van een AI-scheidsrechter die de Verenigde Naties adviseert, met als doel de mensheid objectieve en uitvoerbare oplossingen te bieden voor de strijd tegen klimaatverandering en te zorgen voor een snellere beëindiging van gewapende conflicten.

Op dit moment is het besproken idee nog een concept en om dit concept daadwerkelijk toe te passen zal het brede steun moeten krijgen uit alle lagen van de samenleving wereldwijd. Allereerst zou ik willen stellen dat we dit idee nog geen kans gegeven hebben en dat de risico's van het proberen ervan beperkt zijn. Het gekozen AI systeem zou immers uitsluitend een adviserende rol vervullen zonder de autoriteit om beslissingen af te dwingen. Alle voorgestelde oplossingen zouden nog steeds alleen kunnen worden uitgevoerd—of verworpen—op basis van mondiale democratische principes. De specifieke details over hoe dat democratische proces eruit zou kunnen zien zijn te lezen in hoofdstuk 5 en 6.

Het is niet onrealistisch dat het idee zoals dat hier beschreven wordt op een dag zal worden toegepast. Of dit daadwerkelijk zal gebeuren hangt voornamelijk af van de vraag of gewone burgers én invloedrijke individuen bereid zijn om het idee te steunen.

Revolutionaire ideeën hebben vaak tijd nodig om de wereld te veranderen, maar zoals bijvoorbeeld de Franse revolutie ons liet zien zijn diepgaande verschuivingen mogelijk, zelfs als ze beginnen bij een kleine groep mensen. Toen werd de samenleving nog gedomineerd door monarchieën met absolute macht, vandaag de dag zijn democratische principes wereldwijd verspreid (al worden die helaas nog niet overal toegepast). Het grote verschil tussen toen en nu is het internet, een idee kan tegenwoordig in enkele seconden elk mens met een internetverbinding bereiken.

Daarom geloof ik dat het cruciaal is om kleinschalige succesverhalen te creëren en te delen. Als het AI-scheidsrechtermodel succesvol wordt toegepast in de lokale politiek of zelfs op

nationaal niveau kan het snel internationale aandacht trekken. Elk succesverhaal heeft het potentieel om anderen te inspireren en een bredere steun van het idee op gang te brengen. Zo krijgt iedereen de kans om daadwerkelijk een verschil te maken door de AI-scheidsrechter te omarmen of in de praktijk te brengen binnen zijn of haar eigen context.

Bovenal laat dit concept zien dat er een praktische, uitvoerbare oplossing bestaat voor mondiale problemen die de mensheid tot nu toe niet heeft kunnen oplossen zonder de hulp van een neutrale AI-scheidsrechter. In plaats van de hoop op te geven zou de wereld dit idee daarom een kans kunnen geven.

Het idee van de neutrale AI-scheidsrechter is pas mislukt als het over een periode van meerdere jaren op het hoogste internationale niveau een eerlijke kans heeft gekregen. Tot die tijd kan elke mislukte poging worden gezien als een waardevol leerproces.

Daarnaast daag ik iedereen uit die niet in deze visie gelooft om een betere, praktisch uitvoerbare oplossing voor te stellen waarbij de mensheid op een succesvolle manier om weet te gaan met de klimaatcrisis en in staat is om oorlogen snel te beëindigen. Door concrete oplossingen voor te stellen kan een discussie over mogelijke oplossingen voor de grootste problemen van de mensheid op gang komen en dat is hard nodig. Er zijn inmiddels al genoeg boeken geschreven over alles wat er misgaat in de wereld waarin zeer goed naar voren komt wat de mankementen zijn van de huidige politieke en economische systemen, maar welke boeken beschrijven een concrete en praktisch uitvoerbare oplossing voor die problemen? Hierbij moet dan ook rekening worden gehouden met de huidige stand van zaken als het gaat om macht en internationale politieke besluitvorming, zonder een concreet voorstel dat leidt tot een andere machtsbalans en een praktisch uitvoerbare manier om tot die andere machtsbalans te komen wordt geen volledige oplossing geboden waar de mensheid echt iets aan zou kunnen hebben.

Ik heb de boeken die gaan over een oplossing en waarbij er echt rekening wordt gehouden met de huidige machtsbalans in de wereld niet kunnen vinden. Mochten deze boeken en ideeën wel bestaan dan zouden die veel meer aandacht moeten krijgen. De tijd van wegstijgen is voorbij. De wereld verkeert in crisis en we moeten óf revolutionaire ideeën zoals deze steunen óf een beter alternatief voorstellen.

8. Conclusie

De uitdagingen waarmee de mensheid wordt geconfronteerd—klimaatverandering, wereldwijde conflicten, politieke corruptie—vragen om gedurfde nieuwe benaderingen. Een neutrale AI-scheidsrechter, geleid door ethische principes en democratische principes kan een nieuwe en verbeterde weg vooruit bieden voor de mensheid.

In een wereld die geteisterd wordt door de klimaatcrises, oorlogen en een groeiend wantrouwen in instituties biedt dit essay een hoopvol én realistisch pad vooruit. Het stelt voor om ethisch ontworpen kunstmatige intelligentie in te zetten—niet als vervanging van menselijk leiderschap, maar als neutrale adviseur om beter geïnformeerde en democratische besluitvorming op wereldniveau te ondersteunen.

Aan de hand van voorbeelden die variëren van consumentenrechten tot klimaatonderhandelingen binnen de VN is de centrale boodschap duidelijk: de mensheid beschikt nu over de middelen om zichzelf op een wijzere manier te kunnen besturen.

Het succes van deze visie hangt niet alleen af van technologie, maar van onze gezamenlijke bereidheid om nieuwe modellen van samenwerking en rechtvaardigheid te testen. Dit essay nodigt je uit om je een toekomst voor te stellen waarin redelijkheid, gelijkheid en mondiale verantwoordelijkheid het winnen van ego en angst. Zo'n toekomst is nog steeds mogelijk.

Bronnenlijst

Acton, J. E. E. D. (1887). Letter to Bishop Mandell Creighton. In Historical essays and studies. Macmillan.

Amnesty International. (2009). Cuba: Restrictions on freedom of expression.
<https://www.amnesty.org>

Amnesty International. (2023). Annual Report: Israel and Occupied Palestinian Territories.
<https://www.amnesty.org>

Amnesty International. (2024). Annual report: The state of the world's human rights.
<https://www.amnesty.org/en/latest/research/>

BBC. (2023). Leaders who rose to power as reformers—and what happened next.
<https://www.bbc.com/news>

BBC News. (2025). ICC issues arrest warrant for Putin over Ukraine war crimes.
<https://www.bbc.com>

Bregman, R. (2020). Humankind: A hopeful history. Bloomsbury Publishing.

Chang, J., & Halliday, J. (2005). Mao: The unknown story. Jonathan Cape.

Climate Action Tracker. (2023). Paris Agreement pathways.
<https://climateactiontracker.org>

Climate Policy Radar. (2023). Mapping global climate laws and policies.
<https://climatepolicyradar.org>

Dikötter, F. (2010). Mao's great famine: The history of China's most devastating catastrophe, 1958–62. Walker & Company.

European Commission. (2023). Destination Earth (DestinE). <https://digital-strategy.ec.europa.eu/en/policies/destination-earth>

Freedom House. (2023). Nations in transit 2023: Autocratization accelerates.
<https://freedomhouse.org>

Freedom House. (2023). Turkey: Country report.
<https://freedomhouse.org/country/turkey/freedom-world/2023>

Freedom House. (2023). Nicaragua: Country report.
<https://freedomhouse.org/country/nicaragua/freedom-world/2023>

Henig, R. (2010). The League of Nations: From 1919 to 1929. In *The League of Nations* (2nd ed., pp. 13–45). Haus Publishing.

Human Rights Watch. (2009). New Castro, same Cuba.
<https://www.hrw.org/report/2009/11/18/new-castro-same-cuba>

Human Rights Watch. (2021). Nicaragua: Repression and persecution before elections.
<https://www.hrw.org/news/2021/10/06/nicaragua-repression-and-persecution-elections>

Human Rights Watch. (2022). Turkey: Authoritarian turn accelerates.
<https://www.hrw.org/world-report/2022/country-chapters/turkey>

Human Rights Watch. (2023). World report 2023. <https://www.hrw.org>

Human Rights Watch. (2023). ICC and global accountability: Why enforcement matters.
<https://www.hrw.org>

ICJ. (2023). International Court of Justice: Jurisdiction and functions. <https://www.icj-cij.org>

Inglehart, R., & Norris, P. (2017). *Cultural backlash: Trump, Brexit, and authoritarian populism*. Cambridge University Press.

Intergovernmental Panel on Climate Change. (2023). *Synthesis report of the IPCC Sixth Assessment Report (AR6)*. <https://www.ipcc.ch/report/ar6/syr/>

International Criminal Court. (n.d.). About the ICC. <https://www.icc-cpi.int/about>

International Criminal Court. (2023). About the ICC. <https://www.icc-cpi.int/about>

International Criminal Court. (2024). Arrest warrants issued for Gaza conflict.
<https://www.icc-cpi.int>

International Criminal Court. (2024). Situation in the State of Palestine. <https://www.icc-cpi.int>

International Criminal Court. (2025). Warrant against Russian President. <https://www.icc-cpi.int>

International Criminal Court. (2025). Situation in Ukraine: ICC issues arrest warrant for Vladimir Putin. <https://www.icc-cpi.int>

IPCC. (2023). Climate change 2023: Synthesis report of the IPCC sixth assessment report (AR6). Intergovernmental Panel on Climate Change. <https://www.ipcc.ch/report/ar6/syr/>

Mahbubani, K. (2020). Has the West lost it? A provocation. Penguin.

Mearsheimer, J. J. (2019). The great delusion: Liberal dreams and international realities. Yale University Press.

Pew Research Center. (2020). Global attitudes & trends: Shared values around the world. <https://www.pewresearch.org>

Pew Research Center. (2023). Global attitudes & trends: Shared values across nations. <https://www.pewresearch.org/global/>

Ponzio, R. (2021). Reimagining the UN: Reform and renewal in a divided world. Global Governance Institute.

Roberts, A. (2014). Napoleon: A life. Viking Press.

Scharf, M. P. (2014). Customary international law in times of fundamental change: Recognizing Grotian moments. Cambridge University Press.

Security Council Report. (2023). The veto and Security Council decision-making. <https://www.securitycouncilreport.org>

Snyder, T. (2010). Bloodlands: Europe between Hitler and Stalin. Basic Books.

Snyder, T. (2018). The road to unfreedom: Russia, Europe, America. Tim Duggan Books.

UN General Assembly. (2022). Resolution ES-11/1: Aggression against Ukraine. United Nations. <https://www.un.org/en/ga>

UNFCCC. (2015). The Paris Agreement. United Nations Framework Convention on Climate Change. <https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement>

United Nations. (2022). AI for Climate Action Initiative. <https://www.un.org/en/climatechange>

United Nations. (2023a). Structure and functions of the UN. <https://www.un.org/en/about-us/structure>

United Nations. (2023b). Security Council overview. <https://www.un.org/securitycouncil/>

United Nations. (2023c). Financing the UN: Budget contributions and influence. <https://www.un.org/en>

Uppsala Conflict Data Program. (2024). UCDP Conflict Encyclopedia. <https://ucdp.uu.se>

Weiss, T. G. (2016). What's wrong with the United Nations and how to fix it. Polity Press.

World Resources Institute. (2024). State of climate finance 2024. <https://www.wri.org>