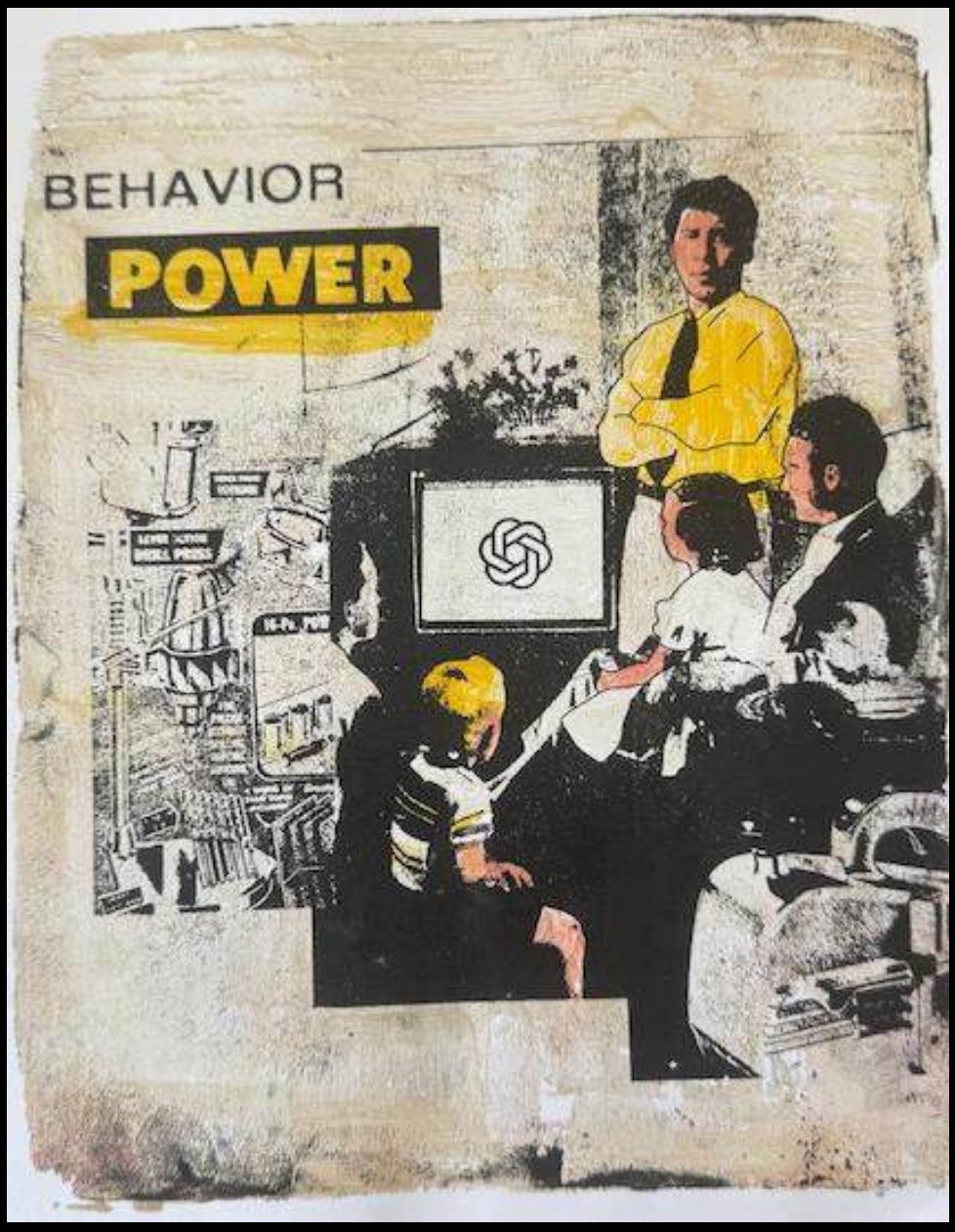
Algoritmen uitleggen

INSPIRERENDE BENADERINGEN UIT HET ALGO→LIT-PROJECT



Algo → Lit
Algorithmic literacy
for effective transparency
in the eu

Inleiding

Deze gids presenteert een reeks inspirerende voorbeelden van activiteiten en projecten die gericht zijn op het uitleggen van algoritmen. Het is een hulpmiddel voor inclusiewerkers, onderzoekers, pleitbezorgers, activisten of andere professionals die op een aantrekkelijke manier voor leken moeten uitleggen hoe algoritmen werken. De reeks voorbeelden kan ook inspiratie bieden voor opdrachtgevers van projecten op het gebied van algoritmische geletterdheid, zoals overheidsorganisaties, scholen of regelgevers op het gebied van data. Let wel: ons project is voornamelijk gericht op gemeenschappen die te maken hebben met algoritmen in Frankrijk, België en Nederland. De selectie van voorbeelden is niet volledig, maar vertegenwoordigt een evenwicht tussen de volgende criteria:

— Ons doel was het presenteren van een verscheidenheid aan formats – van serious games tot workshops, van gedrukte posters tot methodes voor collectieve besluitvorming – uit landen als Frankrijk, België, Nederland, Spanje, de VS en het VK.
— Gezien de opkomst van veel gestandaardiseerde benaderingen van AI-geletterdheid, wilden we de aandacht vestigen op voorbeelden die bekend staan om hun inhoudelijke kwaliteit en documentatie en die bovendien gericht zijn op het kritisch benaderen van algoritmes in plaats van het 'ethisch witwassen'.
— We keken voorbij de hype en publiceerden enkele voorbeelden die in de vergetelheid zijn geraakt, maar we namen ook voorbeelden op die weliswaar bekend zijn in academische kring, maar nooit de kans kregen om bekeken te worden door inclusiewerkers.
— We moesten een evenwicht vinden tussen het presenteren van projecten die makkelijk herbruikbaar zijn en projecten die misschien moeilijker te herhalen zijn, maar zeer inspirerend blijven.

Over het Algo→Lit-project

Het Algo→Lit-project heeft tot doel de algoritmische geletterdheid van medewerkers op het gebied van digitale inclusie te verbeteren, zodat zij burgers beter kunnen helpen algoritmen te begrijpen en hun recht op transparantie effectief kunnen uitoefenen. In een tijd waarin burgers steeds vaker te maken krijgen met ondoorzichtige algoritmen, krijgen medewerkers op het gebied van digitale inclusie (hoe zij ook worden genoemd: ‘helpers’, ‘bemiddelaars’, maatschappelijk werkers die zich bezighouden met de automatisering van openbare organisaties of ambtenaren in lokale districten) te maken met toenemende eisen met betrekking tot deze systemen. We zullen deze professionals de vaardigheden en instrumenten aanreiken die ze nodig hebben om mensen te helpen de systemen te begrijpen en hun rechten uit te oefenen. ALGO->LIT wordt gecoördineerd door Datactivist, de Franse coöperatie die gespecialiseerd is in het openstellen van data en algoritmen, in samenwerking met La Mednum (Parijs, FR), de Franse coöperatie die het netwerk van digitale inclusiewerkers vertegenwoordigt, en de onderzoekscentra FARI - Ai for the common good institute (Brussel, BE) en Waag Futurelab (Amsterdam, NL).

Voor meer informatie over het Algo->Lit-project kunt u terecht op onze website.


Funded by
the European Union

Categorisering van de voorbeelden in deze gids:

1. COLLABORATIEVE WORKSHOPS

Collaboratieve workshops zijn praktische bijeenkomsten van 1 tot 3 uur bedoeld om het publiek actief te betrekken bij het leren over algoritmen en de effecten van AI. Meestal worden daarbij met behulp van sjablonen de problemen in kaart gebracht.

2. MECHANISMEN VOOR COLLECTIEVE BESLUITVORMING EN TOEZICHT

Mechanismen voor collectieve besluitvorming en toezicht zijn initiatieven waarbij een groep burgers de kans krijgt om deel te nemen aan en in beperkte mate invloed uit te oefenen op het bestuur, het ontwerp of de effecten van een bepaald algoritmisch systeem. Ze nemen meestal de vorm aan van burgerjury's, burgerconventies/ vergaderingen, toezichtmechanismen, observatoria, projecten voor het doneren van gegevens of andere vormen van crowdsourcinginitiatieven.

3. STATISCHE EN PEDAGOGISCHE DIAGRAMMEN

Statische en pedagogische visualisaties of diagrammen zijn complexe afbeeldingen die bedoeld zijn om een algoritmisch mechanisme visueel weer te geven. Ze helpen om moeilijk waarneembare dimensies van algoritmen te vereenvoudigen en uit te leggen. Bovendien zijn ze soms bedoeld om een volledig beeld te geven van de vele actoren die betrokken zijn bij het maken van algoritmen of om algoritmen anders weer te geven dan hun ontwerpers, bijvoorbeeld door de nadruk te leggen op dimensies die verband houden met de sociale of ecologische impact van algoritmen en AI.

4. ALGORITMEDOCUMENTATIE EN DATASETVERHALEN

Documentatie bestaat uit geschreven documenten waarin de keuzes die tijdens het ontwerp van een bepaald algoritme zijn gemaakt, worden uitgelegd, becommentarieerd en verantwoord. Ze kunnen informatie geven over de output en effecten van het algoritme. Datasetverhalen zijn verhalende vormen van storytelling die tot doel hebben te documenteren waar de gegevens die worden gebruikt om een AI-systeem te trainen of te verfijnen vandaan komen, hoe ze zijn getransformeerd, wat hun kenmerken en valkuilen zijn.

5. SIMULATORS EN DEMONSTRATORS

Een algoritmesimulator of -demonstrator biedt een dynamische interactie met een interface die berekeningen vereenvoudigt en die door heterogene gegevens in te zetten een impressie geeft van het rechtstreeks beïnvloeden van een algoritme. In vergelijking met een geschreven algoritme of de documentatie ervan, is het voordeel van een simulator dat verschillende parameters en configuraties kunnen worden getest en dat een speelse interactie kan plaatsvinden waardoor je al experimenterend kennis kunt opdoen.

I) Collaboratieve workshops

LLM-PRAKTIJKEN BESCHRIJVEN

Projectwebsite Ecologies of LLM

Het project Ecologies of LLM is ontwikkeld door ontwerpers en sociologen van Sciences Po medialab en stelt de vraag: hoe kunnen we de rol van Large Language Models (LLM's) zoals ChatGPT in gewone werkpraktijken herdefiniëren? Volgens Sciences Po medialab is het noodzakelijk om ruimte te maken voor alternatieve kaders voor AI, om verder te kijken dan toekomstvoorspellingen over AI en juist te kijken naar het heden. Het projectteam roept op om te beginnen met het langzame en stille werk van het opmerken wat AI al doet met ons werk. Pas als we aandacht hebben besteed aan onze huidige werkwijzen, kunnen we deze technologie op onze eigen voorwaarden herdefiniëren en kiezen hoe we ermee omgaan. Belangrijker nog, is dat het observeren van hoe AI voet aan de grond krijgt in onze professionele routines inzicht kan geven in de systemen rondom arbeid die dit mogelijk hebben gemaakt. Het gaat niet alleen om hoe we met AI werken, maar ook om wie bepaalt hoe werk eruit moet zien. Het project Ecologies of LLM biedt een proces of protocol, om nauwkeurig te documenteren hoe AI je werk beïnvloedt en zou kunnen beïnvloeden. Het boek fungeert als een kunstmatige lens om onderzoek te ondersteunen. In zekere zin is het een hulpmiddel om te reflecteren op je eigen gebruik van LLM en je mening te delen met een groep professionals die met soortgelijke uitdagingen en onvoorziene gebeurtenissen te maken hebben.

Hieronder staan voorbeelden van oefeningen uit het boek:

Ex^{2b} – All the *Things* You *Could* Do

Instructions

Think about the tasks you currently do without the help of LLMs and where you could benefit from their assistance.

This list isn't about automation (delegating to the machine) but augmentation (doing more tasks or diversifying tasks). Think about activities or projects you haven't attempted yet or tasks you might hesitate to delegate.

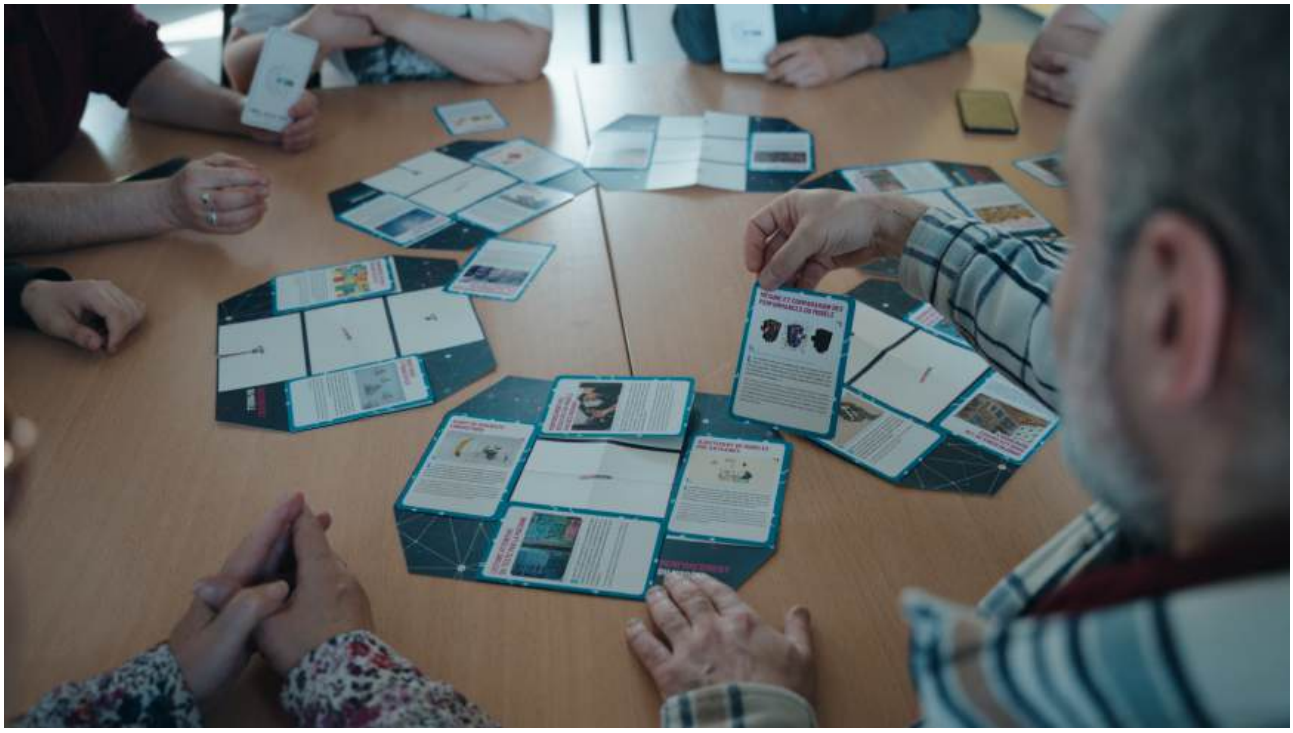
Even if you suspect that an LLM might perform poorly on a given task, include it in your list: this exercise is about exploring potential opportunities rather than evaluating performance.

Ex^{2b} – All the *Things* You *Could* Do

Task List

Task ID	K0	Task	
Task ID	M4	Task	

GEN AI-SCHRIJFTOOLS ONTDEKKEN MET THE AI BLACK BOX



Een sessie van het spel in Nantes in april 2025

Hoe moeten we ons opstellen ten opzichte van een technologie die zelfs experts moeilijk kunnen bijhouden? The AI Black Box is een educatief spel dat is ontwikkeld door de coöperatie Datactivist en het netwerk van inclusiewerkers van Nantes Métropole (Frankrijk). Het is bedoeld om het grote publiek te helpen begrijpen hoe tekstgenererende AI werkt. Spelers ontdekken de levenscyclus van een generatief AI-product zoals ChatGPT, van de winning van mineralen tot de energiekosten van datacenters. Daarbij debatteren ze ook over de maatschappelijke, ecologische en ethische gevolgen van AI. In tegenstelling tot veel vergelijkbare formats richt The AI Black Box zich op het maken van AI en niet alleen op het debatteren over de gevolgen ervan voor werk, het sociale leven of de natuur. Het eerste deel van het spel bestaat uit het reconstrueren van de belangrijkste stappen die leiden tot de creatie van een tekstgenererende AI (bijv. training, fine-tuning, enz.). In het tweede deel moeten spelers bij elke stap kaarten plaatsen die overeenkomen met een subontwerpstap (bijv. de kaart over het filteren van gegevens moet worden geplaatst bij de stap die betrekking heeft op het opschonen van gegevens).

In dit spel worden veel rekenkundige processen vereenvoudigd zonder jargon. Beschouw het als een grote duik in het hart van een AI-systeem. Het spel kan het beste worden gespeeld in een kleine groep van 7 personen. Het spel vereist geen voorkennis, het is een workshop die voor iedereen toegankelijk is. Het format is minimalistisch, open source (CC-BY-SA) en gemakkelijk herbruikbaar. De kit bevat kaarten die u gemakkelijk kunt afdrukken, advies over hoe u het format kunt faciliteren en een aantal slides om de sessie te starten.

Met dit spel kan ik AI op zijn minst minder mysterieus maken. Voor mij is het nog steeds een zegen om een spel te hebben om erover te discussiëren.

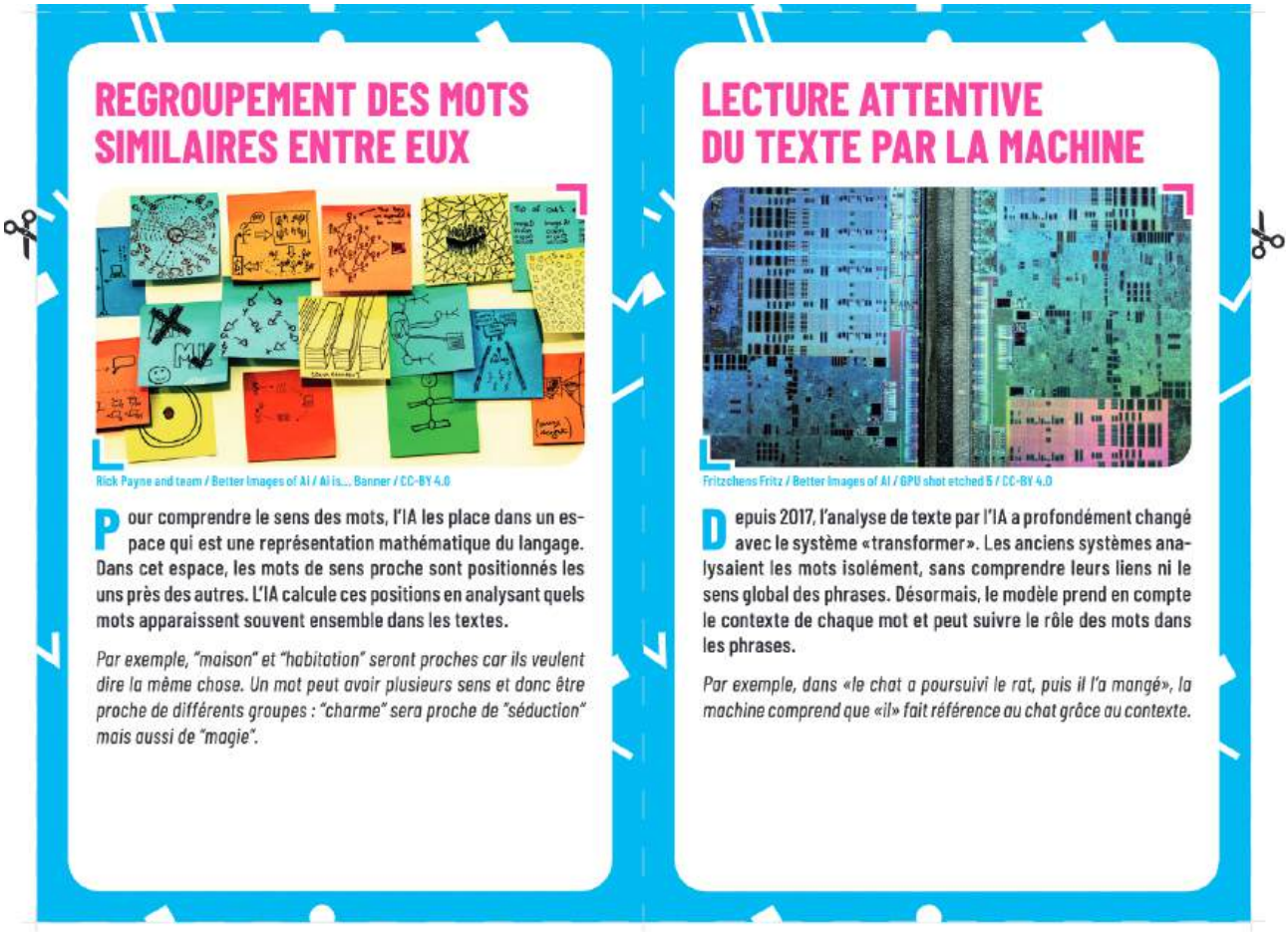
– *AI Black Box deelnemer*

Ik vond het leuk om een overzicht te krijgen van alle stappen die komen kijken bij het bouwen van AI. Het hielp me de toepassingen en milieukwesties rondom AI te begrijpen.

– *AI Black Box deelnemer*

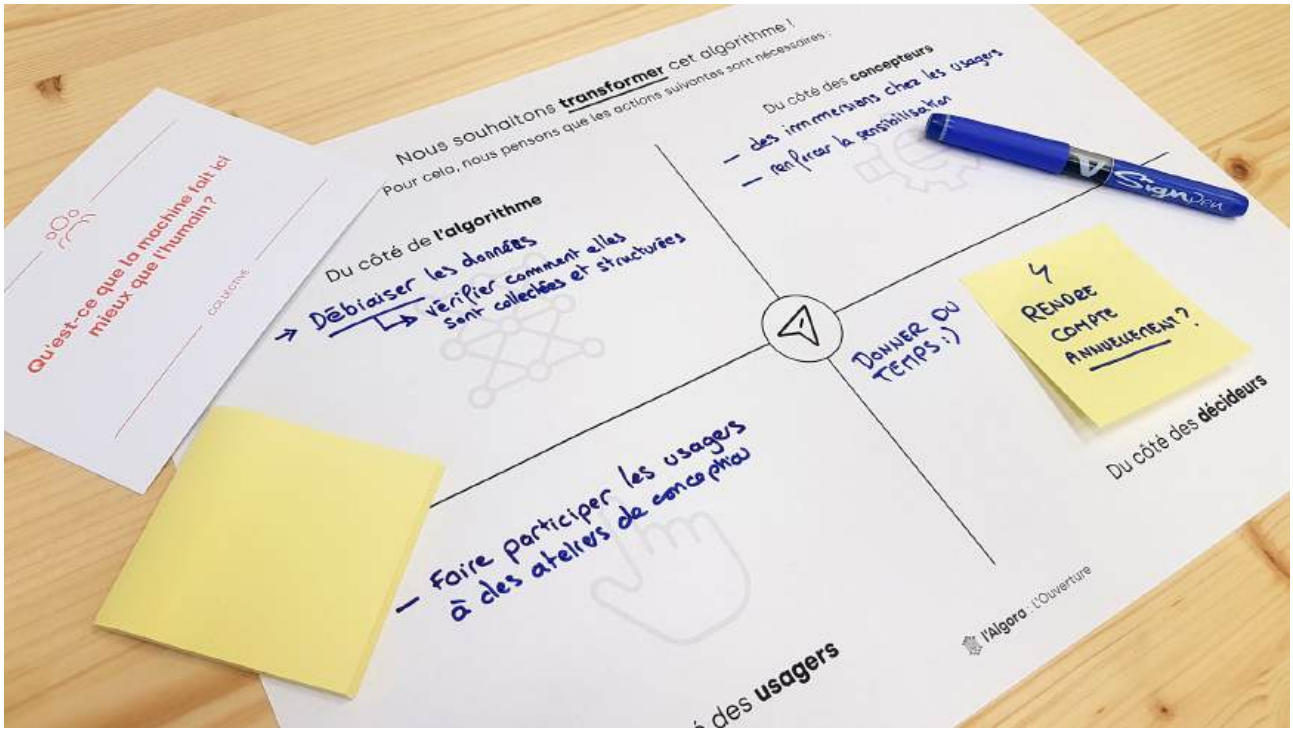
Ik vond het herindelen van de stappen leuk. Ik vond het sociale spel nog leuker, omdat AI weer een sociaal object wordt en niet alleen een technisch object.

– *AI Black Box deelnemer*



Een voorbeeld van twee actiekaarten

DISCUSSIËREN OVER DE GEWENSTE EFFECTEN VAN AI MET THE ALGORA



Een afbeelding van het postersjabloon dat in een workshop wordt gebruikt

The Algora is een discussietool om inzicht te krijgen in de gewenste, onverwachte en ongewenste effecten van een bestaand en actief algoritme, en om daarop te reageren. Het is ontwikkeld door Design Friction in opdracht van Etalab, het open data-team van Frankrijk. The Algora is ontworpen voor gebruik tijdens een collectieve workshop en vergemakkelijkt de communicatie tussen deelnemers met verschillende rollen of functies (besluitvormers, ontwerpers, ontwikkelaars, ambtenaren, gebruikers). Aan de hand van een set gesprekskaarten wordt elke deelnemer uitgenodigd om zijn of haar standpunt kenbaar te maken en dit te toetsen aan de realiteit van de anderen. Het doel van het spel is om gezamenlijk vast te stellen wat het besproken algoritme daadwerkelijk oplevert. De Algora-kit bevat: een set van 40 gesprekskaarten, syntheseborden en actiekaarten, hulpmiddelen voor facilitatie en een versie voor online workshops.

AI ONTDEKKEN MET ST[IA]MMTISCH

Het spelbord

Het spel StIAmmtisch, bedacht door de Délégation académique régionale du Grand Est (Frankrijk), is geïnspireerd op de traditie van de Stammtisch (een traditionele tafelschikking in de Elzas en andere Duitstalige regio's): het gaat vooral om samenzijn en plezier maken rondom kunstmatige intelligentie. Het is een eenvoudig te gebruiken bordspel voor 2 tot 6 spelers of teams, begeleid door een spelleider. Het spel duurt ongeveer 50 minuten en is bedoeld voor gebruik in de klas. Het spel kan echter zowel door volwassenen als door studenten worden gespeeld. Er zijn momenteel niveaus voor beginners en gevorderden beschikbaar. Een niveau voor experts is vanaf het begin gepland, maar is nog niet beschikbaar.

HET ORGANISEREN VAN EEN MENSELIJK PARLEMENT OVER DIGITALE TECHNOLOGIEËN

Een menselijk parlement in 2023 op het Sint-Jansplein bij het Franstalige parlement in Brussel

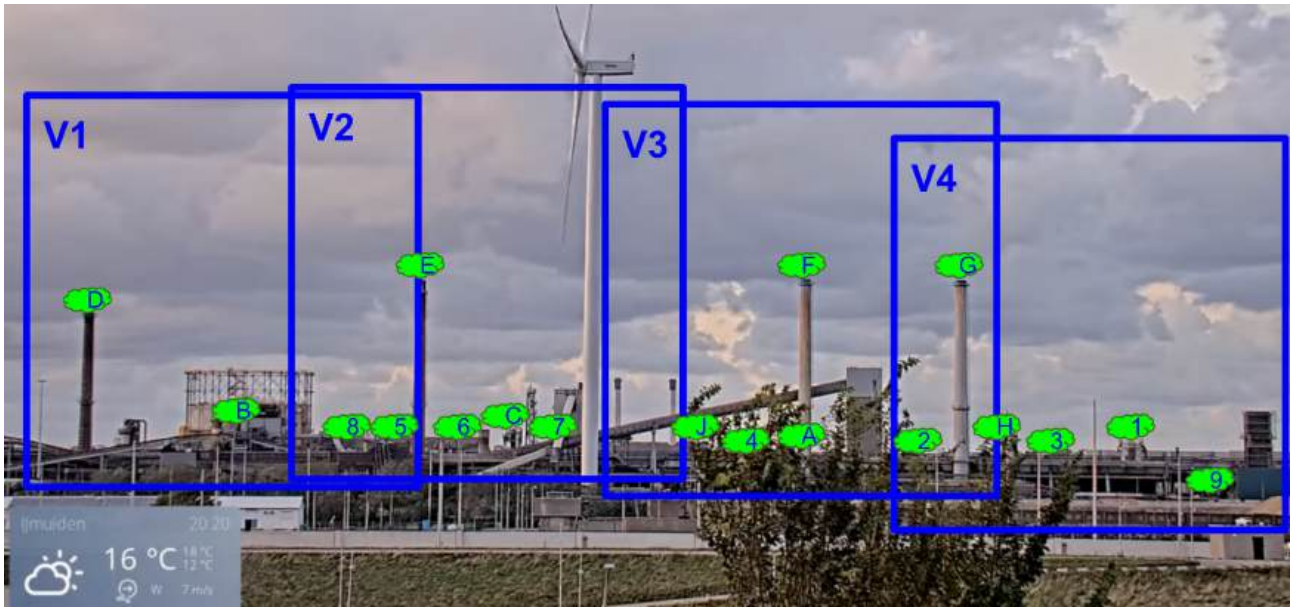
Een parlement humaine, vrij vertaald een “menselijk parlement”, wordt door organisator het Belgische Menselijk Comité voor Digitalisering omschreven als "een moment om te praten en te luisteren, onze ervaringen en ons leed te delen en elkaar te helpen met betrekking tot het digitale in ons dagelijks leven. Het is bedoeld als een plek waar we spreken vanuit onze eigen ervaring, niet vanuit theorieën of ideeën. Een ruimte waar iedereen een expert is en waar een eerlijkere wereld kan bestaan; een leiderschapsinstrument om wetten, collectieve regels of eisen op te stellen. Dit kan tijd kosten en meerdere bijeenkomsten vereisen: efficiëntie is geen prioriteit. Het doel is om tijdens een bijeenkomst en door het opstellen van wetten het menselijke aspect van digitale technologie in ons leven te verdedigen, zodat digitale technologie zich aanpast aan het menselijke aspect en niet andersom." Na het 'Parlement humain' werd een Menselijke Code van de Digitalisering gepubliceerd: een zelfverklaarde burgerwet die sinds 2021 in de maak is. Dit boek bundelt meer dan drie jaar aan acties en ontmoetingen in Brussel en Wallonië: bijeenkomsten met burgers, juridische onderzoekers, politici, intieme discussies, workshops, creaties van kartonnen decors, gezamenlijke maaltijden, demonstraties enzovoort.

BURGERCONVENTIE OVER HET GEBRUIK VAN CHATGPT

Groep studenten en docenten van Sciences Po Aix die deelnamen aan de conventie

Met de toename van het gebruik van ChatGPT door studenten in het hoger onderwijs en onderzoek, wordt het noodzakelijk om de plaats van deze tool in de pedagogische context te herzien en het gebruik ervan in de opleiding van studenten te ondersteunen. Om deze uitdaging aan te gaan, heeft de school Sciences Po Aix (Frankrijk) van 6 maart tot 12 mei 2025 een Universitaire Burgersconventie over Generatieve AI gelanceerd. Dit bijzondere initiatief heeft tot doel het debat over dit onderwerp op gang te brengen door een ruimte voor uitwisseling en discussie te creëren waarbij alle leden van de onderwijsgemeenschap van de school betrokken zijn: docenten, studenten en administratief personeel. Als resultaat van hun werkzaamheden hebben de leden van de conventie een reeks voorstellen voor opleiding en onderwijs aan Sciences Po Aix in het AI-tijdperk ingediend bij het schoolbestuur.

GEGEVENS VOOR AI LABELEN MET HET IJMONDCAM-PROJECT



Afbeeldingen van giftige wolken zoals in kaart gebracht door het neurale netwerk

Waag Futurelab en het team van Yen-Chia Hsu van de Universiteit van Amsterdam organiseerden twee bijeenkomsten waar inwoners van de regio Ijmond (Nederland) mee konden praten over het labelen van de gegevens van een AI-systeem dat rook moet detecteren. Met andere woorden: nauwkeurig bepalen wat een verdachte rookwolk is. Het AI-systeem toont een foto met een vierkante markering rond de verdachte rookwolk. Bewoners kunnen deze markering aanpassen om het systeem nauwkeuriger te maken. Omdat de gemeenschap gegevens labelt om een trainingsdataset te creëren, hebben mensen invloed op de keuzes die het model maakt. Hoewel diepe neurale netwerken nuttig zijn gebleken in verschillende toepassingen (bijvoorbeeld objectherkenning), vereist het trainen van een dergelijk model een grote hoeveelheid gelabelde gegevens. Het zelf labelen van al deze gegevens kan gemakkelijk honderden uren in beslag nemen voor een enkel onderzoek. Daarom zijn vrijwilligers nodig. Het systeem bepaalt een label op basis van alle labels die door onderzoekers en burgers zijn verstrekt. Elke video of afbeelding wordt door ten minste twee vrijwilligers en één onderzoeker bekeken. Als de vrijwilligers hetzelfde label aan de video of afbeelding hebben toegekend, slaat het systeem dit op. Zo niet, dan wordt het proces herhaald. Het resultaat wordt daarom bepaald door meerderheid van stemmen. Eerdere studies over luchtkwaliteitsmonitoring tonen aan dat het zichtbaar maken van bewijzen van rookuitstoot helpt om de aandacht van regelgevende instanties te trekken. Het helpt ook om het vertrouwen van de gemeenschap in de aanpak van luchtvervuiling te vergroten.

RISICO'S VAN AI BEOORDELEN MET DE AI AUTO-TEST VAN LABOR

Labor AI

Voordat een AI-systeem wordt geïntegreerd, is het essentieel om de mogelijke voordelen en risico's ervan te beoordelen. Als eerste diagnose kunt u met de zelfbeoordelingstool die door het onderzoeksproject Labor AI (INRIA, Frankrijk) is ontwikkeld, de kritieke aspecten van de integratie van een AI-systeem in een organisatie beoordelen en aandachtspunten identificeren. Zonder afbreuk te doen aan de wettelijke verplichtingen met betrekking tot de introductie van nieuwe technologieën in bedrijven, biedt deze tool een eerste beoordeling van de risico's die gepaard gaan met de introductie van een AI-systeem.

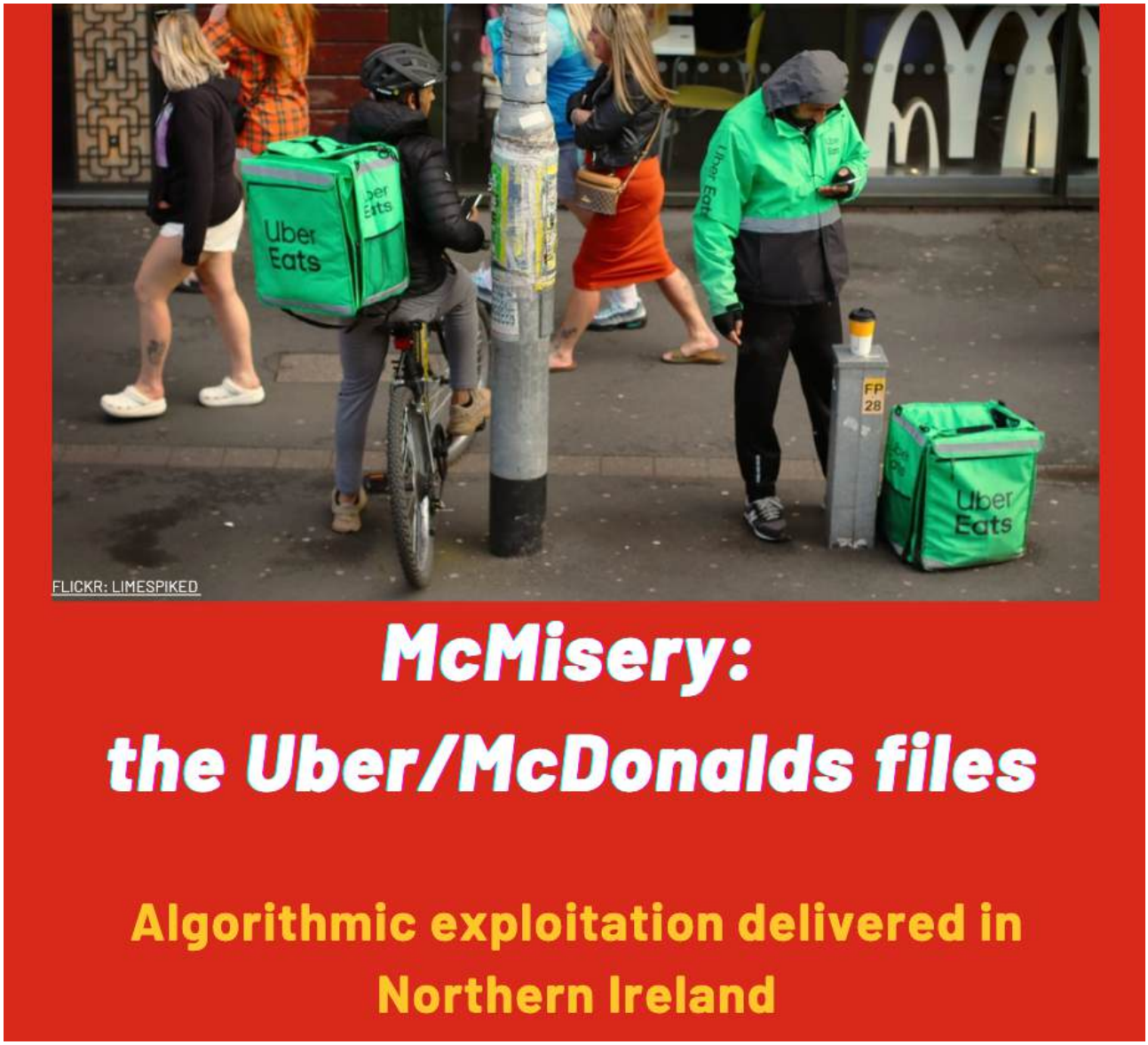
DATA CROWDSOURCING MET DE WORKERS' ALGORITHM OBSERVATORY

Workers' Algorithm Observatory

De Workers' Algorithm Observatory (WAO) is een gecrowdsourcet, onderzoeksgericht samenwerkingsverband dat is gevestigd in de VS en wordt gehost door Princeton University. WAO helpt werknemers en bondgenoten bij het onderzoeken van black-box (ondoorzichtige) algoritmische systemen. Het crowdsourcet de gegevens die nodig zijn voor zinvolle, wetenschappelijke onderzoeken en ontwikkelt tools en ondersteuning voor werknemers en hun bondgenoten. WAO werd gelanceerd in 2022. Momenteel is het een non-profitinitiatief dat wordt gefinancierd door het Mozilla Tech Fund 2023 “Auditing AI”-cohort.

WAO-projecten omvatten bijvoorbeeld FairFare, een collectief initiatief om chauffeurs, organisatoren en beleidsmakers te helpen de zogeheten ride-hailing-industrie te begrijpen door middel van het crowdsourcen van tariefgegevens van chauffeurs. Een eerder project was The Shipt Calculator, een ander door werknemers geleid onderzoek naar een black-box-algoritme dat wordt gebruikt door het bezorgbedrijf Shipt. Willy Solis, lid van het Gig Workers Collective dat het onderzoek leidde, werkte samen met Coworker.org en het team van WAO om gegevens van meer dan honderd werknemers te crowdsourcen en te analyseren met behulp van een aangepaste sms-bot. Ze ontdekten dat de lonen voor de meerderheid waren gedaald, maar voor een minderheid waren gestegen, wat suggereert dat het algoritme meer bepaalt dan alleen het loon – het beïnvloedt ook de zeggenschap en het welzijn van werknemers.

WERKNEMERS HELPEN ALGORITMEN TE BESTRIJDEN MET WORKERINFOEXCHANGE



Omslag van "McMisery", het onderzoek naar Uber-chauffeurs en McDonald's in Noord-Ierland

WorkerInfoExchange is een in het Verenigd Koninkrijk gevestigde organisatie (datacoöperatie) en werkwijze om persoonlijke gegevens en gevallen van algoritmische discriminatie te verzamelen om systemische problemen met algoritmisch beheer inzichtelijk te maken. WorkerInfoExchange heeft bijvoorbeeld aangetoond dat de invoering van dynamische beloning door Uber heeft geleid tot een lagere beloning voor chauffeurs, een hogere commissie voor Uber en een onvoorspelbaarder en ongelijker loon voor het personeel. Het zogenaamde 'dynamische' loonsysteem van Uber maakt gebruik van kunstmatige intelligentie en machine learning om in realtime het loon vast te stellen en werk toe te wijzen — zonder transparantie voor chauffeurs en zonder mogelijkheid om geautomatiseerde beslissingen over loon en werktoewijzing aan te vechten.

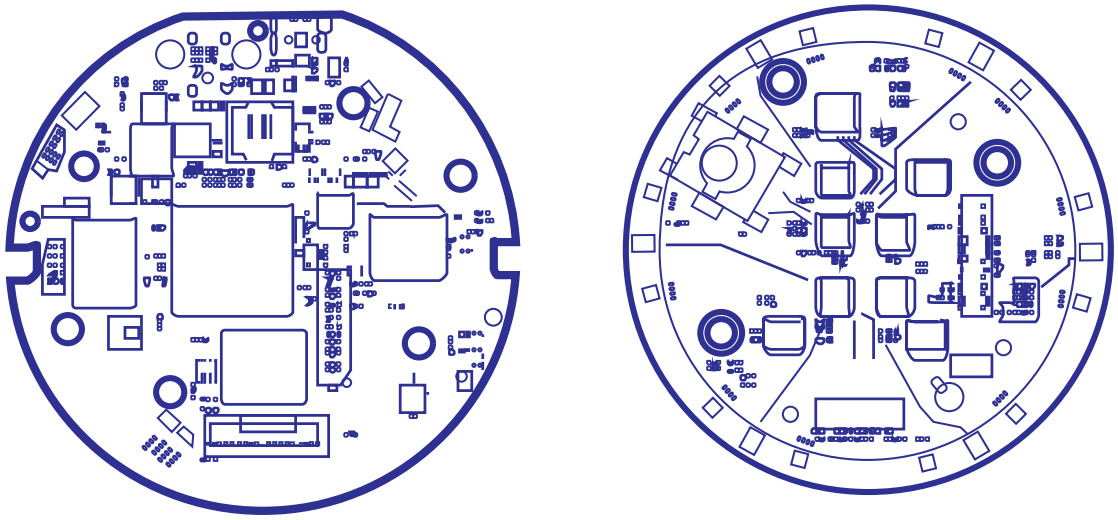
MEER DAN ROBOTS: WAAROM WE BETERE BEELDEN VAN AI NODIG HEBBEN

Pink Office van Jamillah Knowles and Digit - Betere beelden van AI

Typ 'AI-afbeeldingen' in in je zoekmachine je zult een patroon opmerken. Het resultaat is opvallend, en het is hetzelfde op fotobibliotheken en contentplatforms. Het gebrek aan variatie en de onnauwkeurigheid zijn bijna onontkoombaar. Door het overwicht van op sciencefiction geïnspireerde en geantropomorfiseerde afbeeldingen en het gebrek aan gemakkelijk toegankelijke alternatieve afbeeldingen of ideeën is het moeilijk om nauwkeurig over AI te communiceren. Deze AI-afbeeldingen dragen ook bij aan het wantrouwen van het publiek ten opzichte van AI, een groeiend probleem voor innovatie in een vakgebied dat soms als bevooroordeeld, ondoorzichtig en extractief wordt gezien. De huidige dominante afbeeldingen versterken gevaarlijke misvattingen en in het beste geval beperken zij bij het publiek het begrip van het huidige gebruik en de werking van AI-systemen, hun potentieel en implicaties.

We hebben afbeeldingen nodig die een realistischer beeld geven van de technologie en de mensen erachter en die wijzen op de sterke en zwakke punten, de context en de toepassingen ervan. Het Better Images of AI-project heeft tot doel een nieuwe verzameling betere beelden van AI te creëren die iedereen kan gebruiken, te beginnen met een verzameling inspirerende voorbeelden. De eerste fase van dit project is bedoeld om te onderzoeken hoe deze nieuwe beelden eruit zouden kunnen zien, en om mensen met verschillende creatieve, technische en andere achtergronden uit te nodigen om samen te werken aan het ontwikkelen van betere beelden. Bij het creëren van nieuwe beelden moeten we nadenken over wat een goed stockbeeld is. Waarom gebruiken mensen ze en hoe? Vertegenwoordigt het beeld een bepaald onderdeel van de technologie of probeert het een breder verhaal te vertellen? Welke emotionele reactie moet het publiek hebben wanneer het naar de afbeelding kijkt? Helpt het mensen om de technologie te begrijpen en is het een accurate weergave?

DE ANATOMIE VAN EEN AI-SYSTEEM ONTDEKKEN



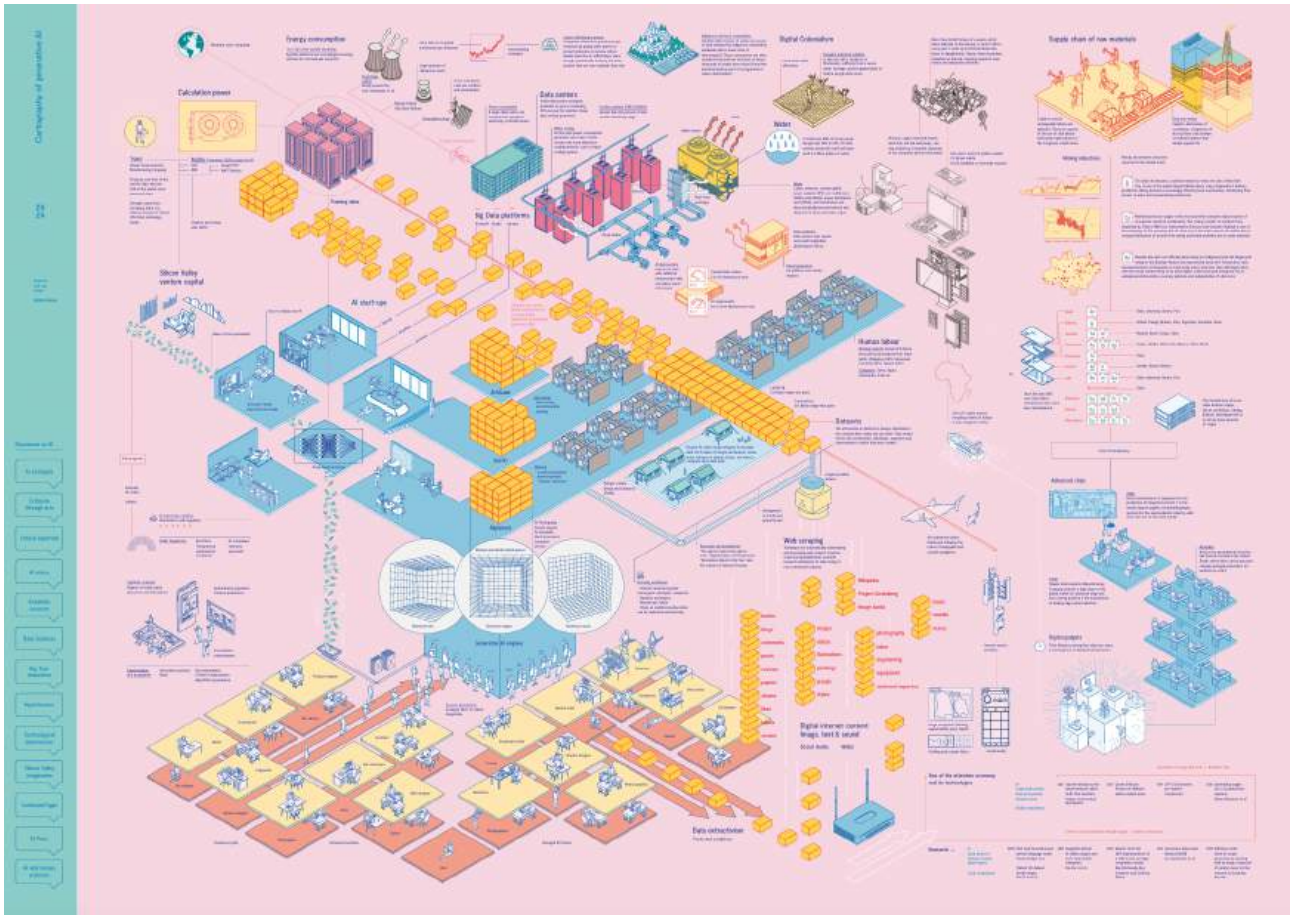
Een synthetisch overzicht van het Amazon Echo-circuit

Anatomy of an AI System is de anatomische kaart van de menselijke arbeid, gegevens en planetaire hulpbronnen die nodig zijn om het Amazon Echo-apparaat te laten werken. Het is gemaakt door kunstenaar Vladan Joler en onderzoeker Kate Crawford.

> In een kamer staat een cilinder. Hij is onbewogen, glad, eenvoudig en klein. Hij is 14,8 cm hoog en heeft een enkel blauwgroen cirkelvormig licht dat rond de bovenrand loopt. Hij staat er stil bij. Een vrouw komt de kamer binnen met een slapend kind in haar armen en richt zich tot de cilinder.
>
> 'Alexa, doe de lichten in de hal aan.'
>
> De cilinder komt tot leven. 'Oké.' De kamer wordt verlicht. De vrouw knikt lichtjes en draagt het kind naar boven.
>
> *– Een interactie met het Amazon Echo-apparaat*

In dit vluchtige moment van interactie wordt een enorme matrix van capaciteiten opgeroepen: verweven ketens van grondstofwinning, menselijke arbeid en algoritmische verwerking in netwerken van mijnbouw, logistiek, distributie, voorspelling en optimalisatie. De omvang van dit systeem gaat het menselijk voorstellingsvermogen bijna te boven. Hoe kunnen we er ons bewust van worden en de omvang en complexiteit ervan als een verbonden geheel begrijpen? Het project begint met een schets: een opengewerkte weergave van een planetair systeem in drie fasen van geboorte, leven en dood, vergezeld van een essay in 21 delen. Samen vormt dit een anatomische kaart van één enkel AI-systeem.

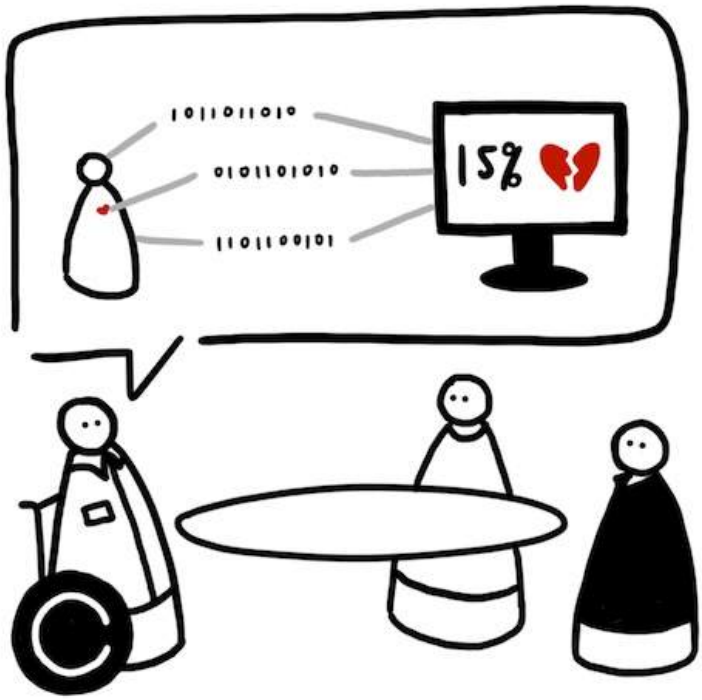
VISUALISATIE VAN DE INFRASTRUCTUUR ACHTER GENERATIEVE AI



Poster De volledige cartografie van generatieve AI

Welke reeks extracties, instanties en bronnen stellen ons in staat om online te communiceren met een tekstgenererende tool of om binnen enkele seconden afbeeldingen te verkrijgen? De cartografie van generatieve AI, gemaakt door de Spaanse ontwerpstudio Estampa, brengt de hele toeleveringsketen van een generatieve AI in kaart, van het modelleren van datasets tot de energiebehoeften, en presenteert tegelijkertijd veel van de dimensies ervan met betrekking tot rekenkracht, arbeidskrachten en economie. De verzameling relaties die hier wordt gepresenteerd, vormt een mozaïek dat moeilijk te bevatten is, omdat objecten en kennis van verschillende soorten en ordes van grootte met elkaar in verband worden gebracht. De discussies rond AI hebben vaak een sterk mythische lading en gaan gepaard met een reeks terugkerende metaforen en verbeeldingen: algoritmische instanties die losstaan van menselijk handelen of de niet-onderhandelbare technologie die ons de toekomst oplegt, de universaliteit van data. De reeks discoursen rond deze technologieën, of ze nu meer gespecialiseerd of meer algemeen toegankelijk zijn, geven deze technologieën zelf uiteindelijk op de een of andere manier mede vorm. Daarom is het project Cartography of Generative AI gebaseerd op de motivatie om een conceptuele kaart maken, die een groot deel van de actoren en middelen omvat die betrokken zijn bij het complexe en veelzijdige object dat we Generative AI noemen. Voortbouwend op een lange genealogie van kritische cartografieën, heeft deze visualisatie tot doel het fenomeen in kaart te brengen, rekening houdend met de spanningen, controverses en ecosystemen die het mogelijk maken.

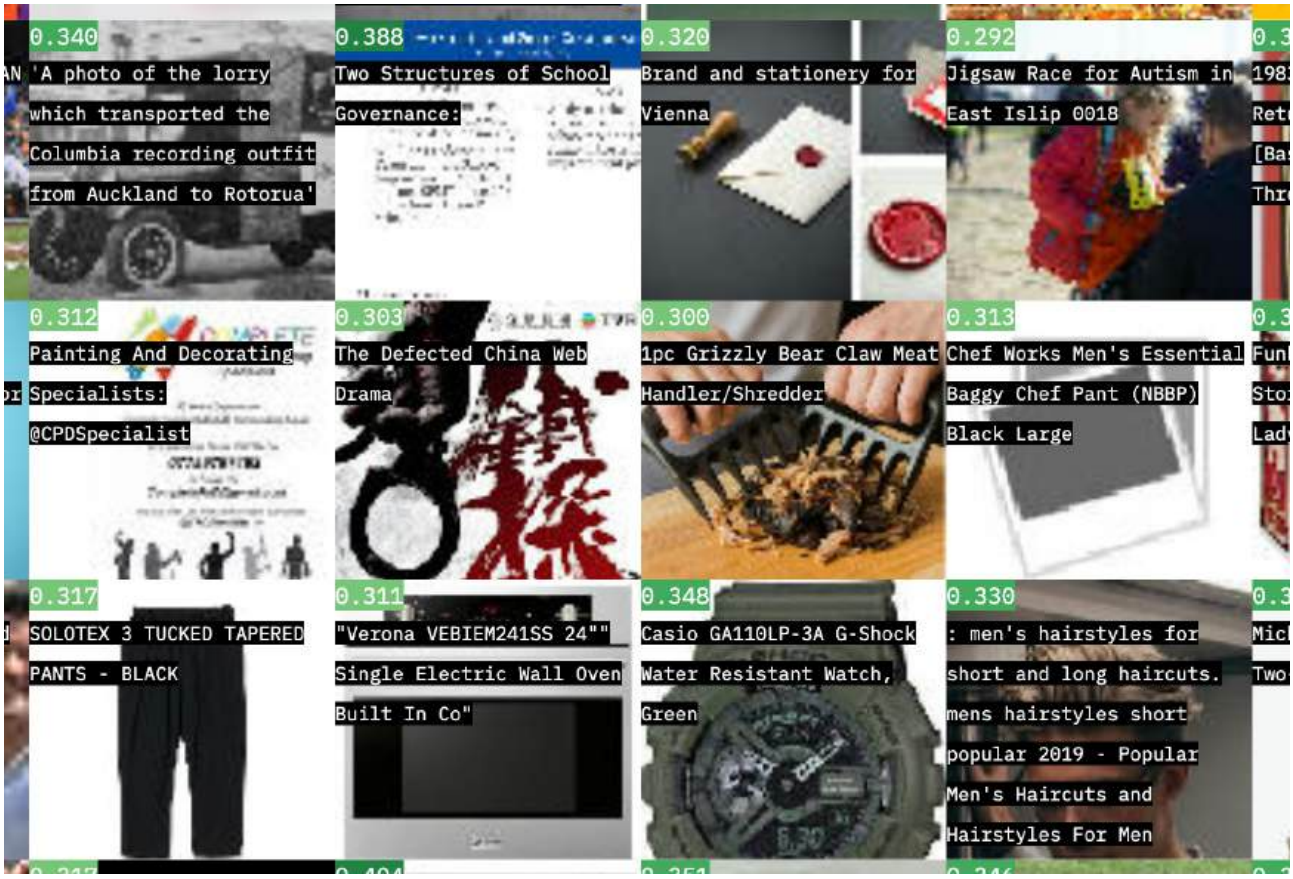
HOE WERKT MATCHING? DE NIERTRANSPLANTATIE CASUS



Voorspelling door Yaning Wu – Betere beelden van AI

Het Franse agentschap voor biogeneeskunde heeft een gids gepubliceerd om inzicht te geven in de score die bepalend is voor de toewijzing van niertransplantaten. Dit gebeurt in een context waarin het tekort aan transplantaten steeds weer moeilijke medische en ethische vragen oproept over het selectieproces voor patiënten die een transplantaat van een overleden donor toegewezen krijgen. De toewijzing van transplantaten is een complex en gevoelig proces, zowel wat het concept als de uitvoering betreft. Er is geen algemene of definitieve oplossing. Het verschilt van land tot land, afhankelijk van de toewijzingscriteria die in aanmerking worden genomen, en in de loop van de tijd, afhankelijk van de periodieke evaluatie van de resultaten. Een toewijzingssysteem moet worden aangepast, efficiënt zijn en zo rechtvaardig mogelijk zijn ten opzichte van de gezondheidsbehoeften van de betrokken patiënten. Voor sommige patiënten is een oplossing om transplantaties op basis van “prioriteit” aan te bieden. Dit is het geval voor patiënten in noodsituaties of met hyperimmuniteit, die momenteel voorrang krijgen, doorgaans op nationaal niveau. Voor de overgrote meerderheid is het onmogelijk om de ene groep voorrang te geven boven de andere. Het toewijzingssysteem moet daarom een compromis vinden tussen eerlijkheid, efficiëntie en haalbaarheid. Er moet rekening worden gehouden met verschillende toewijzingscriteria tegelijk. Een effectieve oplossing is het gebruik van een score. In deze officiële documentatie van het Franse Biomedisch Agentschap wordt uitgelegd hoe deze score wordt berekend.

WAT ZIJN DE GEGEVENS ACHTER AI?



Models all the way down met afbeeldingen van metagegevens

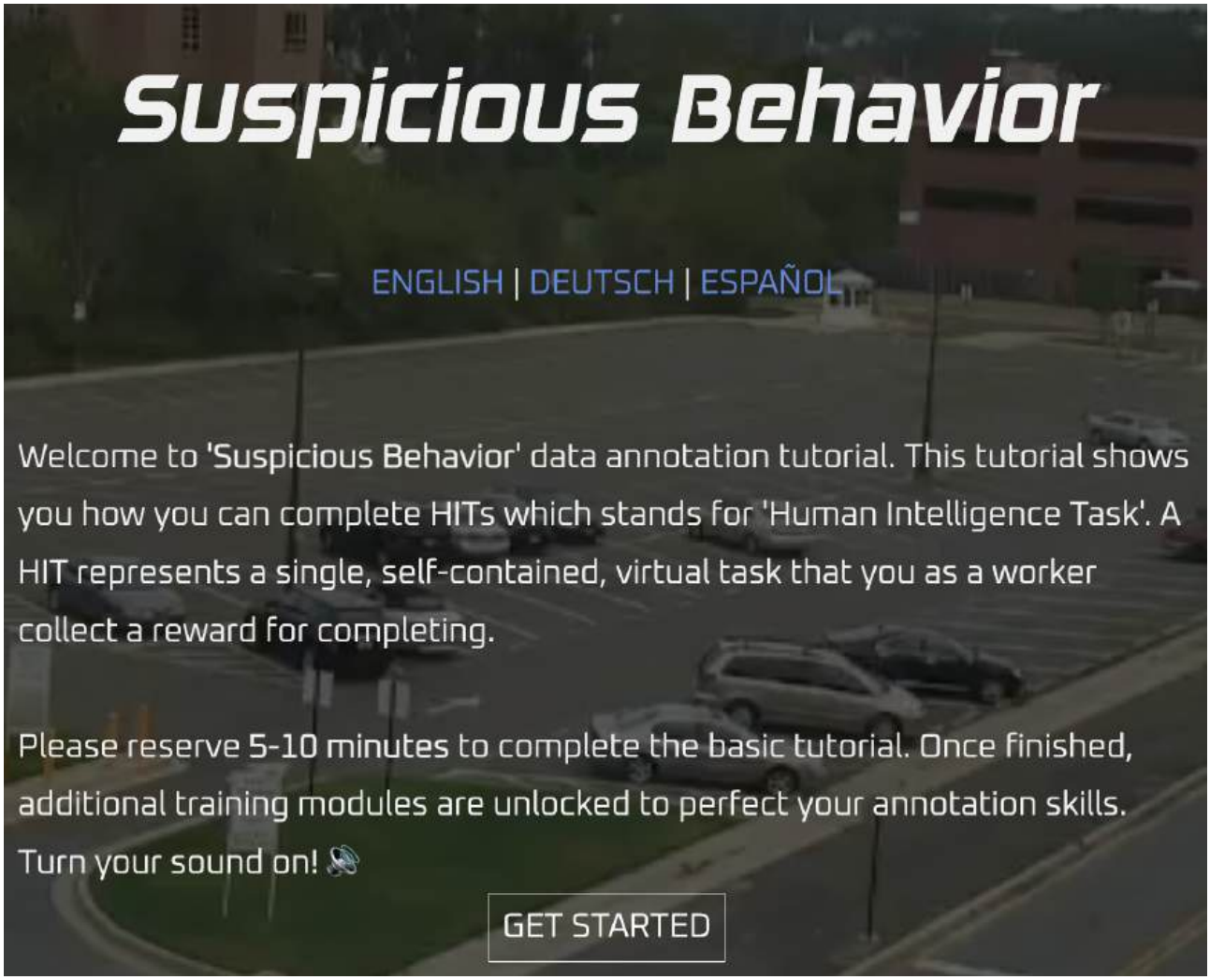
Models all the way down is een visueel verhaal dat kan worden gezien als documentatie in een verhalende vorm. Het vertelt het verhaal van LAION-5B, een open-source basisdataset die wordt gebruikt om AI-modellen zoals Stable Diffusion te trainen. Het bevat 5,8 miljard beeld- en tekstparen – een omvang die te groot is om te kunnen bevatten. In dit visuele onderzoek, gemaakt door Christo Buschek en Jer Thorpthe van het onderzoeksproject Knowing Machines, volgen we de opbouw van de dataset om de inhoud, implicaties en verwikkelingen ervan beter te begrijpen.

BETROUWBAARHEID VAN NEURALE NETWERKEN AANTONEN MET PYRAT

De PyRAT-demonstrator

Een team van onderzoekers bij CEA-List, een onderzoekslaboratorium dat deel uitmaakt van de Franse openbare organisatie voor technologie- en energieonderzoek, heeft nieuwe software ontwikkeld genaamd PyRAT. PyRAT is ontworpen om de betrouwbaarheid van de respons van neurale netwerken in bepaalde configuraties aan te tonen. Het kan helpen bij het beoordelen van de betrouwbaarheid van AI-gestuurde systemen in kritieke toepassingen, variërend van het voorkomen van botsingen tussen vliegtuigen tot medische diagnostiek. Het laboratorium wilde PyRAT graag op een zowel technische als educatieve manier aan een breed publiek demonstreren, waaronder onderzoekers, wetenschappers, industriële partners en potentiële klanten, door middel van een demonstrator die zou worden gepresenteerd in de showroom van CEA in Paris-Saclay. De Franse ontwerpstudio Units werkte samen met het laboratorium om te bepalen hoe de zeer complexe en ondoorzichtige processen in grote neurale netwerken het beste konden worden weergegeven en uitgelegd. Het resultaat is een demonstrator die de logica achter PyRAT en de verschillende wiskundige bewerkingen die daarbij een rol spelen, blootlegt, waardoor gebruikers kunnen experimenteren met verschillende analyseparameters. De interface is te zien in de showrooms van het CEA, waar hij kan worden bediend met een speciale hardwarecontroller die Units heeft ontworpen en ontwikkeld.

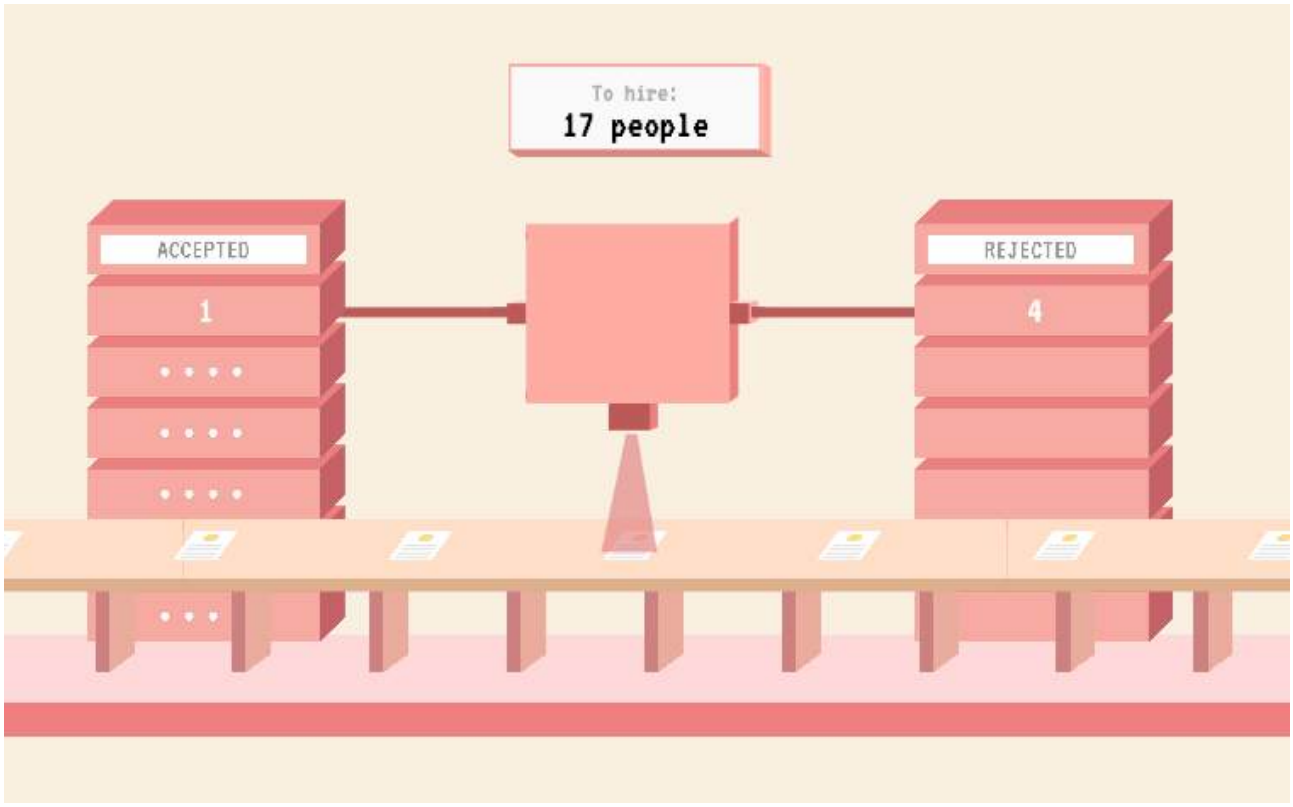
HOE INTERPRETEREN ALGORITMEN BEELDEN VAN VERDACHT GEDRAG?



Een screenshot van het spel

Het online spel Suspicious Behavior, ontwikkeld door de kunstenaars Kairus, toont een wereld van verborgen menselijke arbeid, die de basis vormt voor hoe 'intelligente' computervisiesystemen onze acties interpreteren. Via een fysieke thuiskantooropstelling en een tutorial over het labelen van beelden maakt de gebruiker kennis met het vervelende werk van geoutsourcede annotators (medewerkers die handmatig beelden van meta-data voorzien). In een interactieve tutorial voor een fictief bedrijf wordt de gebruiker gemotiveerd en geïnstrueerd om de taak van het labelen van verdacht gedrag op zich te nemen.

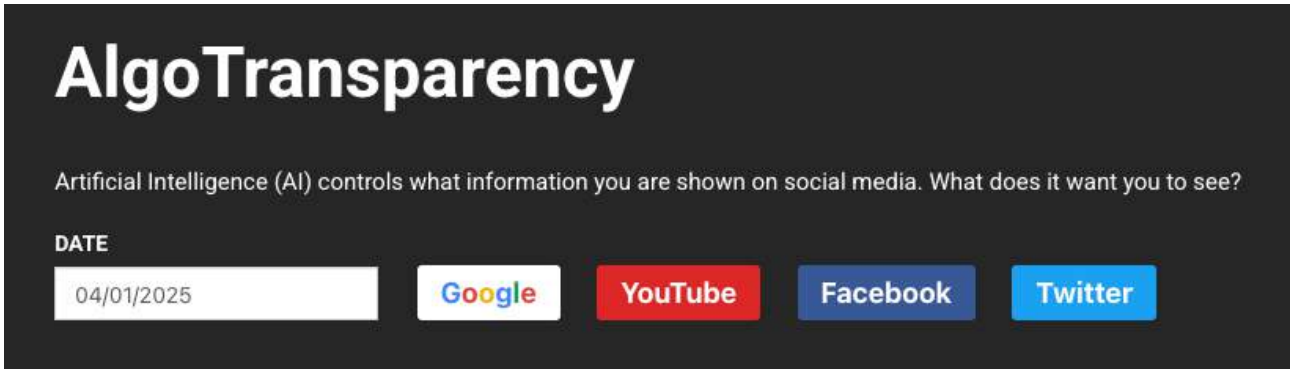
BIAS ONTDEKKEN IN WERVINGSALGORITMEN



Visual van wervingsalgoritmen

Survival of the Best Fit is een educatief spel over bias (vooringenomenheid) in AI bij het aannemen van personeel. Het legt uit hoe misbruik van AI ervoor kan zorgen dat machines menselijke vooroordelen overnemen en ongelijkheid verder vergroten. In veel publieke debatten wordt AI voorgesteld als een bedreiging die ons wordt opgelegd, in plaats van iets waar we zelf invloed op hebben. De ontwerpers van het spel wilden daar verandering in brengen door mensen te helpen de technologie te begrijpen en meer verantwoordelijkheid te eisen van degenen die steeds meer alomtegenwoordige softwaresystemen bouwen. Het project is ontwikkeld door Gabor Csapo, Jihyun Kim, Miha Klasinc en Alia ElKattan en ondersteund door de Creative Media Award van de Mozilla Foundation. Met deze game wilden de ontwikkelaars een publiek bereiken dat weliswaar niet maker van technologie is, maar dat elke dag weer de gevolgen ervan ondervindt. Het is belangrijk om mensen te laten begrijpen hoe AI werkt en hoe het hen kan beïnvloeden, zodat ze meer transparantie en verantwoordelijkheid kunnen eisen van systemen die steeds meer beslissingen voor hen nemen.

HOE WERKT HET AANBEVELINGSALGORITME VAN YOUTUBE?



Maandelijks maken 2 miljard mensen gebruik van YouTube. Het aanbevelingsalgoritme van YouTube bepaalt voor meer dan 70% van de views wat mensen bekijken. Dat is 700 miljoen uur – ofwel 1000 mensenlevens – per dag. Het project AlgoTransparency monitort meer dan 800 toonaangevende informatiekanalen, waaronder nieuwskanalen, vloggers en programma's uit het hele politieke spectrum.

Een gids gemaakt door het Algo->Lit-project, gefinancierd door het Erasmus + Agentschap, gecoördineerd door Datactivist en ontwikkeld in samenwerking met de organisaties Waag Futurelab, Fari en Mednum.

Bronvermelding van de afbeelding op de omslag: Bart Fish & Power Tools of AI / Better images of AI / Creative Commons. Voor het eerst gepubliceerd in september 2025.

Voor vragen kunt u contact met ons opnemen via: algolit@datactivist.