



Paper

Inventarisatie van beschikbare AI-indicatoren

Barteld Braaksma,
Jeffrey Hoogland,
Bart Klijs,
Fleur de Loijer,
Eric Wassink,
Sten Zijlstra

Februari 2022

Inventarisatie van beschikbare AI-indicatoren

Barteld Braaksma, Jeffrey Hoogland, Bart Klijs,
Fleur de Loijer, Eric Wassink, Sten Zijlstra

projectnummer 001038
DRD/SQS/DBD
28 februari 2022

samenvatting In deze nota presenteert het CBS de opbrengsten van de eerste inventarisatie van beschikbare AI-indicatoren voor Nederland. De inventarisatie dient als voorstudie voor een te ontwikkelen structurele AI-monitor, die een beeld moet geven van de impact van AI op de Nederlandse economie en samenleving. De uitgevoerde inventarisatie heeft een tweeledig karakter. Enerzijds is een scan uitgevoerd van mogelijk bruikbare indicatoren die beschikbaar zijn bij het CBS vanuit lopend statistisch werk. Anderzijds zijn een aantal bredere, voornamelijk grootschalige, initiatieven buiten het CBS in kaart gebracht. Voor een evenwichtig en betrouwbaar beeld van AI in Nederland moeten uiteraard nog de nodige vervolgstappen gezet worden.

trefwoorden AI-monitor, indicatoren, NLAIC

1. Inleiding

In deze nota presenteert het CBS de opbrengsten van de eerste inventarisatie van beschikbare AI-indicatoren voor Nederland. De inventarisatie dient als voorstudie voor een te ontwikkelen structurele AI-monitor, die een beeld moet geven van de impact van AI op de Nederlandse economie en samenleving. De voorstudie is uitgevoerd in opdracht van de Nederlandse AI Coalitie, in nauwe afstemming met gelijktijdig onderzoek door TNO en Elsevier/RELX op verwante terreinen.

De inventarisatie is open van aard. Dit wil zeggen dat deze is uitgevoerd zonder een vooraf bepaalde definitie van AI en zonder een vooraf bepaalde zoekstrategie. De uitgevoerde inventarisatie heeft een tweeledig karakter. Enerzijds is een scan uitgevoerd van mogelijk bruikbare indicatoren die beschikbaar zijn bij het CBS vanuit lopend statistisch werk. Anderzijds zijn een aantal bredere, voornamelijk grootschalige, initiatieven buiten het CBS in kaart gebracht waarbinnen indicatoren

worden berekend en gepubliceerd die behulpzaam kunnen zijn bij het monitoren van de ontwikkelingen op het terrein van AI in Nederland.

1.1 Bronnen en indicatoren

In de tekst die volgt beschrijven we eerst de CBS-indicatoren en vervolgens een selectie van de aangetroffen externe indicatoren op basis van een aantal belangrijke geïdentificeerde initiatieven. Per initiatief wordt ingegaan op welke indicatoren het initiatief heeft opgeleverd, welke data-bronnen hiervoor worden gebruikt, en voor welke jaren en landen de informatie beschikbaar is.

Veel van de CBS-indicatoren zijn gebaseerd op klassieke bronnen zoals overheidsregisters en steekproefsgewijze enquêtes. De gepubliceerde uitkomsten zijn gebaseerd op robuuste methodologie en voldoen aan (vaak internationale) kwaliteitsnormen. Over het algemeen wordt veel aandacht besteed aan volgtijdelijke vergelijkbaarheid en waar mogelijk aan onderlinge relateerbaarheid van cijfers. In een beperkt aantal gevallen worden (aanvullend) alternatieve nieuwe bronnen gebruikt. Hierbij kan gedacht worden aan elektronische transactiegegevens zoals scannerdata van winkelketens of gegevens verzameld via het internet, zoals kenmerken van bedrijvenwebsites.

Veel van de externe indicatoren zijn beschikbaar voor meerdere landen, wat de mogelijkheid tot benchmarking biedt. Een aantal van deze indicatoren is al beschikbaar voor meerdere jaren, waardoor ontwikkelingen over de afgelopen jaren bekeken zouden kunnen worden. Daarbij zijn de domeinen waarop de beschikbare indicatoren betrekking hebben divers; denk bijvoorbeeld aan resultaten van wetenschappelijk onderzoek naar AI, maar ook aan werkgelegenheidsaspecten.

Een belangrijke kanttekening is dat veel van de extern beschikbare indicatoren en een aantal van de CBS-indicatoren gebruik maken van databronnen waarvan het oorspronkelijke doel niet was het maken van statistieken of het in kaart brengen van ontwikkelingen over de tijd. Veel van deze bronnen zijn vrij beschikbaar, maar het is onduidelijk hoe valide en betrouwbaar de indicatoren zijn die gebaseerd zijn op dergelijke bronnen en hoe goed vergelijkingen in de tijd gemaakt kunnen worden. Een klein deel van de indicatoren is berekend op basis van enquêtedata. Voor enquêtes (van het CBS) geldt dat deze weliswaar kostbaar zijn, maar worden opgezet met het doel kwalitatief goede en representatieve informatie te leveren.

Een aantal van de geïnventariseerde indicatoren zeggen rechtstreeks iets over een aspect van AI, omdat ze afkomstig zijn uit bronnen waarin kenmerken van AI worden uitgevraagd of zijn geregistreerd. In veel gevallen zijn directe indicatoren echter niet beschikbaar. Dan kan het wel een optie zijn om op basis van geschikte aannames, met verdeelsleutels en benaderingen bruikbare indicatoren af te leiden. Dergelijke *afgeleide indicatoren* zoals we ze hieronder noemen vragen wel om een extra kritische blik met het oog op de kwaliteit.

Verder valt op dat een deel van de extern beschikbare AI-indicatoren zogenaamde *samengestelde indicatoren* zijn. Samengestelde indicatoren zijn een samenvattende maat en bestrijken meerdere AI onderwerpen of domeinen tegelijkertijd. Hoewel het aantrekkelijk is om zo met een enkel cijfer een indruk te krijgen over een hele branche, heeft het gebruik van samengestelde indicatoren ook nadelen. De precieze interpretatie van een samengestelde indicator is meestal minder eenvoudig dan die van een enkelvoudige indicator. Daarnaast is de uitkomst van een samengestelde indicator vaak afhankelijk van keuzes met betrekking tot de weging van onderliggende cijfers en heeft deze zo een zekere mate van willekeur. Het CBS is daarom vaak terughoudend met het berekenen en publiceren van samengestelde indicatoren. Enkele uitzonderingen hierop zijn specifieke en internationaal uitgebreid besproken en geharmoniseerde gevallen zoals inflatiecijfers en Bruto Nationaal Inkomen.

2. Mogelijke CBS-indicatoren

Dit deel over indicatoren binnen het CBS is iets breder ingestoken dan de delen voor de andere initiatieven. Naast de reeds beschikbare indicatoren worden enkele mogelijkheden gepresenteerd voor de ontwikkeling van nieuwe indicatoren op basis van informatie die beschikbaar is binnen het CBS. Deze ‘mogelijk te ontwikkelen indicatoren’ zijn met nadruk niet ‘ready-to-use’ en vragen een extra investering in onderzoek en ontwikkeling.

2.1 Enquête ICT-gebruik bedrijven

Binnen het CBS zijn verschillende bronnen beschikbaar die kunnen dienen als aanknopingspunt voor de monitoring van AI in Nederland. Eén van de meest direct bruikbare bronnen is waarschijnlijk de **Enquête ICT-gebruik bedrijven**. Met deze enquête verzamelt het CBS jaarlijks gegevens over het gebruik van informatie- en communicatietechnologieën (ICT) bij bedrijven op basis van een steekproef onder ongeveer 11 duizend bedrijven. Een deel van de vragen is beperkt tot bedrijven met 10 of meer werkzame personen om de rapportagelasten voor het kleinbedrijf te beperken. Uitgangspunt voor de ICT-enquête zijn Europese (minimum)afspraken over de vraagstelling, zie onder het kopje Eurostat. Landen kunnen daar extra vragen en detailleringen aan toevoegen. Een nieuw onderwerp in deze enquête (door het CBS toegevoegd vanaf verslagjaar 2019) is het gebruik van AI technologie door bedrijven. Via de ICT-enquête zijn gegevens beschikbaar over het gebruik van AI-technologieën door bedrijven van 10 of meer werkzame personen.

De zeven specifieke technologieën die in de Enquête ICT-gebruik bedrijven worden uitgevraagd zijn textmining, spraakherkenning, natural language generation, afbeeldingsherkenning en beeldverwerking, machine learning, software voor robot gestuurde procesautomatisering en autonome machines. Ook wordt gevraagd naar het doel waarvoor deze AI-technologieën worden ingezet:

marketing en verkoop, productieprocessen, administratieve processen, bestuur van bedrijven, logistiek, ICT beveiliging en human resources management. Verder worden er vragen gesteld over de manier waarop de AI-technologie is verkregen door het bedrijf: eigen ontwikkeling, aanpassing door eigen medewerkers, als gebruiksklaar systeem aangeschaft, of maatwerk door externe leverancier. En slotte, of het gebruik van AI-technologie is overwogen door het bedrijf. Via de Enquête ICT-gebruik bedrijven zijn indicatoren beschikbaar over:

- Bedrijven in Nederland die big data analyseren.
- Bedrijven in Nederland die één of meer AI-technologieën gebruikt.
- Bedrijven in Nederland die service robots gebruiken.
- Bedrijven in Nederland die een chatbot gebruiken.
- Bedrijven in Nederland die gebruik maken of overwegen gebruik te maken van AI-toepassingen.
- Gebruik van AI-toepassingen bij bedrijven.
- Toepassing van AI-technologie bij bedrijven naar activiteit.
- Intensiteit inzet van AI-technologie bij bedrijfsfunctie.
- Herkomst gebruikte AI-technologieën.
- Redenen voor bedrijven om geen gebruik te maken van AI-toepassingen.
- Intensiteit inzet bij bedrijven van AI-activiteiten.
- Toepassing van AI-technologie bij bedrijven naar bedrijfsfunctie.

Deze indicatoren worden voor Nederland gespecificeerd naar bedrijfstak of bedrijfsgrootte. Ook de komende jaren zullen er in principe gegevens over het gebruik van AI-technologieën worden opgenomen in de Enquête ICT-gebruik bedrijven. Het CBS is in ieder geval bereid om jaarlijks deze vragen op te nemen omwille van de AI-monitor. Daarmee lijkt deze bron geschikt voor monitoring (van veranderingen) op langere termijn. Doordat het hier gaat om een eigen CBS-enquête is het mogelijk om de vraagstelling aan te passen en te verfijnen, onder voorbehoud van de ermee gemoeide kosten en meetbaarheid van de uitgevraagde informatie. Hierbij is het wel van belang waar mogelijk aangesloten te blijven bij internationale ontwikkelingen rond de vraagstelling. Dat kan voor lastige keuzes zorgen als de internationale vraagstelling inhoudelijk verandert (dat is niet ongebruikelijk) terwijl er nationaal juist behoefte is om trends goed in beeld te houden.

2.2 Mogelijk te ontwikkelen indicatoren uit bestaande CBS statistieken

De enige CBS-enquête die rechtstreeks informatie over AI bevat is de hierboven beschreven Enquête ICT-gebruik Bedrijven. Naast deze klassieke aanpak op basis van een enquête werkt het CBS aan een meer innovatieve manier om AI-bedrijven te identificeren op basis van webscraping, zie hieronder. Die methode is veelbelovend maar nog volop in ontwikkeling en vormt op dit moment nog geen voldoende solide basis voor robuuste statistische informatie.

De informatie uit de ICT-enquête over welke bedrijven AI gebruiken en die uit de webscraping/modelleringsmethode over bedrijven die AI ontwikkelen vormt de ruggengraat voor het bepalen van het bedrijven-ecosysteem dat zich met AI bezighoudt. Die informatie kan op diverse manieren worden verrijkt en aangevuld met gegevens uit andere bronnen om bijvoorbeeld een beeld te schetsen van de arbeidsmarkt rond AI en de omzetten die gerealiseerd worden. In veel gevallen zal het wel nodig zijn om een beroep te doen op afgeleide indicatoren die deels op aannames en veronderstellingen gebaseerd moeten worden. Het zal extra onderzoek vragen om te kunnen bepalen wat kwalitatief en methodologisch verantwoord en voldoende betrouwbaar is. Hieronder geven we een aantal voorbeelden van mogelijkheden.

Identificatie van AI-bedrijven

Een goede identificatie van AI-bedrijven kan sterk bijdragen aan de verdere ontwikkeling van indicatoren rond AI-bedrijven (zoals de omzet, toegevoegde waarde, import export, werkgelegenheid et cetera van deze bedrijven). Het CBS heeft daarom samen met het bedrijf 'Innovatiespotter' en de 'Rijksdienst voor Ondernemend Nederland' (RVO), een onderdeel van het Ministerie van Economische Zaken en Klimaat, gewerkt aan een **methode om AI-bedrijven te identificeren** middels webscraping en modellering. Kortweg houdt deze methode in dat er op basis van een bekende set AI-bedrijven een classificatiemodel wordt ontwikkeld dat op basis van tekst op de websites van deze bedrijven de bedrijven classificeert als AI-bedrijf (ja/nee). Een verschil met de ICT-enquête is dat een AI-bedrijf hier in eerste instantie wordt gezien als een bedrijf dat AI *ontwikkelt* en niet zo zeer als een bedrijf dat AI *gebruikt*.

Uit het onderzoek bleek dat het niet eenvoudig is een goed onderscheidend model te ontwikkelen, met name doordat het aantal AI-bedrijven in Nederland klein is, de afbakening wat wel en niet mee te rekenen lastig scherp valt te krijgen en doordat het op voorhand niet eenvoudig is te bepalen wat een AI-website is (oftewel wat een website van een AI-bedrijf kenmerkt), wat van belang is voor het opstellen van een dataset waarop het model getraind kan worden. Het in het onderzoek ontwikkelde model bleek te resulteren in een groot percentage vals-negatieven, waardoor het minder bruikbaar is om AI-bedrijven te identificeren. Er zijn echter aanknopingspunten dit model te verbeteren. Zo kan er ingezet worden op verbetering van de trainingsdataset (bijvoorbeeld gebaseerd op cases geïdentificeerd in de enquête ICT-gebruik bedrijven), en kan het aantal vals positieven mogelijk teruggebracht worden door het model toe te spitsen op specifieke branches waarin een groter deel AI-bedrijven wordt verwacht (bijv. de ICT sector).

Omzet AI bedrijven

In de Enquête ICT-gebruik bedrijven wordt de bedrijfsomzet uitgevraagd. Door deze informatie in combinatie met de hierboven genoemde indicatoren op basis van de Enquête ICT-gebruik bedrijven te gebruiken kan mogelijk een indicator worden ontwikkeld die betrekking heeft op:

- Het deel van de omzet van bedrijven in Nederland die wordt gemaakt door bedrijven die gebruik maken van AI-technologieën.

AI-onderwijs

Andere indicatoren die mogelijk ontwikkeld kunnen worden hebben betrekking op AI-onderwijs. Binnen het CBS is informatie beschikbaar over alle opleidingen in Nederland. Sommige van deze opleidingen zullen vrijwel volledig gericht zijn op AI, terwijl in andere opleidingen AI één van de onderwerpen is. De eerste uitdaging bij het monitoren van AI-onderwijs is dus de afbakening van AI-opleidingen. Wanneer deze afbakening is gedaan kunnen middels een koppeling met andere binnen het CBS beschikbare gegevensbronnen (ook op persoonsniveau) indicatoren worden ontwikkeld voor bijvoorbeeld:

- Het aantal inschrijvingen bij AI-gerelateerde opleidingen in Nederland.
- Het aantal afgestudeerden bij AI-gerelateerde opleidingen in Nederland.
- Tijd tussen afstuderen bij een AI-gerelateerde opleiding en het starten van de eerste baan.
- Het aantal werkzame personen (voltijd/deeltijd) dat is afgestudeerd aan AI-gerelateerde opleidingen.

Arbeidsmarkt voor AI

Binnen het CBS zijn diverse bestanden beschikbaar die betrekking hebben op de arbeidsmarkt. De zogenaamde 'polisadministratie' afkomstig van het UWV bevat informatie over loon, uitkeringen en arbeidscontracten van alle verzekerde werknemers in Nederland. In combinatie met diverse achtergrondkenmerken en beschikbare informatie over werkgevers geeft dit rijke mogelijkheden voor arbeidsmarktanalyses. Als het mogelijk is een selectie van AI (-gerelateerde) bedrijven te maken zou dit bijvoorbeeld als aanknopingspunt kunnen dienen voor het in kaart brengen van

- Het aantal AI-gerelateerde banen in Nederland.

Hiernaast kan ook de Enquête beroepsbevolking (EBB) van het CBS aanknopingspunten bieden voor het in kaart brengen van het aantal AI-gerelateerde banen. In combinatie met bijvoorbeeld vacaturecijfers kan dit een indruk geven van de spanning op de AI-arbeidsmarkt.

Vacatures in de AI: enquête

Voor het in kaart brengen van het aantal vacatures voor AI gerelateerde banen zijn er twee aanknopingspunten. Het eerste aanknopingspunt is de vacature-enquête van het CBS. Deze enquête registreert alleen het aantal vacatures per bedrijf, niet om wat voor soort vacatures het gaat. Dat betekent dat alleen met gebruikmaking van extra aannames een afgeleide indicator valt samen te stellen rond AI-vacatures.

Vacatures in de AI: webscraping

Het tweede aanknopingspunt is een bestand dat beschikbaar is met openstaande vacatures in Nederland op online kanalen, verkregen via webscraping. Dit bestand bevat ook informatie over het type bedrijf van de vacature (SBI classificatie), het beroep (ISCO/BRC classificatie), het gevraagde opleidingsniveau (laag/middelbaar/hoog), en het contracttype (bv vast/flexibel/uitzend; voltijds/deeltijds). Ook hier zitten haken en ogen aan: een online advertentie vermeldt bijvoorbeeld

niet altijd het aantal vacatures bij een gevraagde functie, kan dubbel geteld worden omdat hij om meerdere plaatsen gepubliceerd wordt en wordt ook niet altijd direct verwijderd na vervulling.

Vaardigheden voor AI beroepen

Binnen het CBS wordt ook werk gemaakt van het in kaart brengen van de vaardigheden (ook wel 'skills' of competenties) die gevraagd worden voor de banen in de vacatures. Deze statistiek bevindt zich nog in de ontwikkelfase, maar biedt mogelijk aanknopingspunten voor het in kaart brengen van vacatures voor AI-gerelateerde beroepen (op basis van benodigde vaardigheden) of vacatures waarin gevraagd wordt om mensen met vaardigheden op het terrein van AI.

Overige onderwerpen

Hierboven zijn een aantal voorbeelden gegeven van bronnen en statistieken die het CBS beschikbaar heeft, of kan maken, die min of meer direct betrekking hebben op AI-monitoring. Daarnaast is er nog (veel) meer mogelijk. Het CBS verzamelt data en publiceert statistieken over letterlijk duizenden onderwerpen. Een paar voorbeelden van mogelijk relevante invalshoeken:

- Diverse bedrijfsenquêtes geven inzicht in bijvoorbeeld de omzetten, kostenstructuren, financiële positie, investeringen, innovatiekracht, export en groeiverwachtingen van bedrijven. De R&D enquête geeft bijvoorbeeld informatie over Research & Development activiteiten van bedrijven, instellingen en hoger onderwijs: inkomsten en uitgaven, R&D-personeel en het aantal betrokken bedrijven. Hoewel deze enquêtes niet specifiek over AI gaan geven ze mogelijk een basis voor raming van AI-gerelateerde kenmerken.
- De genoemde kenmerken kunnen mogelijk uitgesplitst worden naar vestigingsregio, grootte, leeftijd, rechtsvorm en dergelijke van bedrijven. Hierbij moet wel worden aangetekend dat de mogelijkheden voor uitsplitsing beperkt kunnen worden door omvang en aard van de steekproefwaarneming. Sommige enquêtes beperken zich bijvoorbeeld tot bedrijven met minstens tien werkzame personen.
- Naast (en deels op basis van) de genoemde statistieken die vooral een bedrijfseconomische invalshoek hebben maakt het CBS de Nationale rekeningen, waarin alle belangrijke macro-economische statistische indicatoren in een volledig geïntegreerd en internationaal geharmoniseerd stelsel samenkomen: bruto binnenlands product (bbp), economische groeicijfers, overheidsschuld, werkgelegenheid, finale consumptie, et cetera. In het kader van het begrip 'brede welvaart' worden in dit stelsel tegenwoordig ook bijvoorbeeld milieueffecten van economische activiteit meegenomen ('Carbon Footprint').
- De AI-monitor zal waarschijnlijk vooral betrekking hebben op bedrijfsgegevens vanwege de nadruk op verdienvermogen en economische groei, maar het CBS beschikt ook over veel gegevens over personen en huishoudens. Als tegenhanger van de enquête ICT-gebruik bedrijven is er de 'enquête ICT-gebruik personen en huishoudens' die een beeld geeft van het gebruik van ICT-middelen in brede zin, inclusief vaardigheden en beleving. Ook hier is het (onder voorbehoud) mogelijk om vragen of modules aan toe te voegen. Verder voert het CBS geregeld onderzoek uit over beleving, (on)veiligheidsgevoelens en maatschappelijk vertrouwen van burgers. Ook dit soort gegevens kunnen mogelijk een rol spelen om, in

combinatie met een rijke schat aan achtergrondkenmerken, een indruk te geven van bijvoorbeeld de maatschappelijke impact en mate van inclusiviteit van AI.

Nota bene

Bij bovenstaande 'Overige onderwerpen' zijn wel een aantal kanttekeningen van toepassing.

- Veel van de genoemde onderwerpen zijn bedrijfsgeoriënteerd. Om deze gegevens zinvol te kunnen gebruiken is het essentieel AI-bedrijven goed en consistent te kunnen identificeren. Aan dat identificeren zitten haken en ogen. Via de ICT-enquête kunnen AI-bedrijven geïdentificeerd worden, maar dat betreft slechts een steekproef. De experimentele methode via webscraping en textmining geeft in principe integrale resultaten maar is nog niet uitontwikkeld.
- De genoemde onderwerpen vragen in de meeste gevallen aanvullende bewerkingen en aannames om tot afgeleide indicatoren te komen. Dat vraagt om expliciete aandacht voor kwaliteitscontrole en transparantie over de gebruikte methoden.

Links

AI- en big data gebruik naar bedrijfstak (bedrijven met 10 of meer werkzame personen):

<https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84832NED/table?dl=60886>

AI- en big data gebruik naar bedrijfsgrootte (bedrijven met 10 of meer werkzame personen):

<https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84830NED/table?dl=60885>

Big data gebruik naar bedrijfsgrootte (bedrijven met 2-10 werkzame personen):

<https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84829NED/table?dl=6088F>

3. Externe indicatoren

In dit hoofdstuk worden een aantal van de belangrijkste vindplaatsen beschreven buiten het CBS die informatie geven over diverse aspecten van AI. Het hele landschap is tamelijk ingewikkeld en dynamisch, erg internationaal en lijkt in veel gevallen meer gericht op onderlinge vergelijkbaarheid van landen dan exacte cijfers voor een individueel land.

Voor de hieronder gepresenteerde resultaten is een veelheid aan databronnen, begrippen en methoden gebruikt. Een inventarisatie van de bruikbaarheid van de informatie van deze vindplaatsen voor een Nederlandse AI-monitor zou in ieder geval moeten bevatten:

1. een selectie van de relevante indicatoren;
2. een beoordeling van de gebruikte databronnen, definities/begrippen en methodologie voor deze indicatoren; met name hoe geschikt de methode is voor het weergeven van betrouwbare uitkomsten en trends voor de Nederlandse situatie;

3. een inschatting van de onderlinge samenhang en vergelijkbaarheid tussen indicatoren afkomstig van verschillende vindplaatsen;
4. een vaststelling van eventuele kosten, gebruiksvoorwaarden en waarborgen, bijvoorbeeld voor toekomstige beschikbaarheid van sleutelindicatoren.

3.1 OESO

Door de OESO is het **OECD.AI Policy Observatory** ontwikkeld. OECD.AI combineert bronnen die beschikbaar zijn binnen de OESO en partners van de OESO uit diverse groepen stakeholders. Het initiatief wil de dialoog tussen stakeholders faciliteren en voorzien in multidisciplinaire, evidence-based beleidsanalyses in de domeinen waar AI de grootste impact heeft.

Het doel van OECD.AI is (veel) breder dan monitoring alleen, maar een van de vier hoofdonderdelen op de website van OECD.AI is 'trends and data'. Dit onderdeel gebruikt veelal openbaar toegankelijke bronnen en verwerkt deze tot trends over waar, hoe en hoe snel AI wordt ontwikkeld en gebruikt, en in welke sectoren.

De OECD.AI site bevat een rijke collectie publicaties en overzichten rond AI en verwante thema's zoals digitalisering. Het is duidelijk dat de OESO aan de weg timmert op het gebied van AI-monitoring en het lijkt verstandig om goed kennis te nemen van hun inzichten.

Bij 'Trends and Data' gaat het deels om geïsoleerde indicatoren, deels om meer samenhangende collecties. De bronnen die gebruikt worden zijn nogal divers, zie hieronder. De onderwerpen waarover informatie wordt gepresenteerd zijn (tussen haakjes de bron):

- Het aantal AI beleidsinitiatieven in OESO landen (*Database of National AI policies*)
- Het aantal AI samenwerkingen tussen landen (*Microsoft Academic*)
- Onderzoek en patenten (*Microsoft Academic*)
- Wetenschappelijke publicaties (*Elsevier*)
- AI software ontwikkeling (*Github*)
- Trends in AI gerelateerde zoekopdrachten (*Google*)
- AI banen en vaardigheden (*LinkedIn Economic Graph*)
- Investerings in AI (*Preqin*)

Voor elk van deze onderwerpen is een veelheid aan specificaties en invalshoeken beschikbaar. Deze informatie is doorgaans beschikbaar voor alle OESO landen, waardoor de data zich goed leent voor benchmarking. Ook is de data veelal voor meerdere jaren beschikbaar waardoor trends in kaart gebracht kunnen worden. De jaren waarvoor data beschikbaar is variëren overigens per onderwerp. Informatie over onderzoek en patenten is bijvoorbeeld beschikbaar vanaf 1980, terwijl informatie

over AI banen beschikbaar is vanaf 2017. Nederland is opgenomen in de data maar wordt niet standaard weergegeven in de resultaten: Nederland valt veelal weg tussen de grotere landen.

Link

Landingspagina OECD.AI:

<https://oecd.ai/en/data-from-partners?selectedArea=ai-news>

3.2 Europese Commissie

AI watch

Door de Europese Commissie-onderdelen DGCNECT en JRC is de **AI watch** ontwikkeld. De AI watch monitort de ontwikkeling, het gebruik en de impact van AI in Europa. Ook hier wordt gebruik gemaakt van een veelheid van databronnen zoals 'Angellist', 'CB Insights', 'AI Job Fair', 'EPO PATSTAT', 'CORDIS', 'Crunchbase', 'Venturesource by Dow Jones' et cetera. Relevante onderdelen van het AI watch initiatief zijn het **AI landscape dashboard** en het **AI investments dashboard**.

AI landscape dashboard

Dit dashboard geeft informatie over de spelers in het AI veld, de ontplooide AI activiteiten, en informatiestromen tussen actoren in het AI veld voor de jaren 2009-2018. De informatie die gepresenteerd wordt op het AI landscape dashboard betreft:

- Aantal AI spelers (bedrijven, onderzoeksinstellingen) per land.
- Aantal AI activiteiten per land.
- Aantal AI patent aanvragen per land.
- Aantal door de EC gefinancierde projecten per land.
- Revealed comparative advantage¹ per land.

Alle informatie is ook per AI thema beschikbaar. Deze thema's zijn: AI services, Audio & Natural Language Processing, Automation, Autonomous Robotics, Computer Vision applications, Connected and Automated Vehicles, Internet of Everything, Machine Learning for Image Processing en Machine Learning Fundamentals.

¹ Het revealed comparative advantage is een maat van activiteit op een bepaald AI thema van een land, relatief ten opzichte van de gemiddelde activiteit in andere landen.

AI investments dashboard

Dit dashboard geeft informatie over investeringen in AI in Europa, vooralsnog alleen voor 2018. De informatie die gepresenteerd wordt op het AI investments dashboard betreft:

- AI investeringen per land.
- AI investeringen per land per capita.
- AI investeringen per land naar type ('talent, skills, and life-long learning', 'data, technology and infrastructure' en 'from the lab to the market').
- AI investeringen per land naar bron en doel.

Digital Economy and Society Index

Sinds 2014 volgt de Europese Commissie de digitale vooruitgang van de lidstaten en publiceert zij jaarlijks verslagen over de index van de digitale economie en maatschappij (DESI, **Digital Economy and Society Index**). Elk jaar zijn in de DESI verslagen landenprofielen opgenomen die de lidstaten kunnen gebruiken om zicht te krijgen op de gebieden waarop met voorrang maatregelen moeten worden genomen. Nederland staat in 2021 (net als in 2020) op de vierde plaats van de 27 EU-lidstaten.

De DESI-verslagen bevatten thematische hoofdstukken met een analyse op EU-niveau van de belangrijkste beleidsterreinen rond digitalisering. In de editie 2021 besteedt de Commissie in de DESI ook aandacht aan de twee belangrijke beleidsinitiatieven die in de komende jaren hun weerslag zullen hebben op de digitale transformatie in de EU: de Europese faciliteit voor herstel en veerkracht en het Digitaal kompas 2030. Hiervoor, en om rekening te houden met de nieuwste ontwikkelingen in technologie en beleid, heeft de Commissie in de editie 2021 van de DESI een aantal dingen veranderd. In de toekomst zijn nog verdere aanpassingen voorzien.

De DESI sluit aan bij de vier hoofdgebieden van het Digitaal kompas. Per hoofdgebied zijn een aantal deelindices gedefinieerd die uit diverse bronnen afkomstig zijn, zoals Eurostat en nationale statistische instituten, nationale toezichthouders op telecommunicatie en marktpartijen als Capgemini, Empirica en Ipsos (vaak na aanvullende toetsing door overheidspartijen). Hieronder staan de vier hoofdgebieden met de Nederlandse positie in de EU-27 in 2021 en 2020, en het aantal deelindices:

DESI hoofdgebied	Positie NL (2021/2020)
1. Menselijk kapitaal (7 deelindices):	3 / 4
2. Connectiviteit (10 deelindices):	2 / 6
3. Integratie van digitale technologie (11 deelindices):	5 / 4
4. Digitale overheidsdiensten (5 deelindices):	8 / 7

Een toelichting op alle indicatoren staat in de DESI methodologie-notitie.

De DESI bestrijkt het hele terrein van digitalisering en is daarom veel breder dan AI. Alleen in het derde hoofdgebied, 'Integratie van digitale technologie', is een AI deelindex opgenomen: het aantal bedrijven dat ten minste twee AI-technologieën gebruikt. De bron hiervoor is de 'European enterprise survey on the use of technologies based on artificial intelligence', zie hieronder.

Gegeven het toenemende belang van AI is de kans vrij groot dat de DESI in de toekomst meer AI-gerelateerde indicatoren gaat bevatten. Het Nationaal groeifonds wordt in het DESI-verslag over 2021 al uitgelicht als opmerkelijk en welkom initiatief.

European enterprise survey on the use of technologies based on artificial intelligence

In opdracht van DGCNECT hebben Ipsos en iCite een telefonische enquête uitgevoerd onder in totaal 9640 bedrijven, getrokken uit de ORBIS database en verdeeld over dertig landen: de EU-27, Noorwegen, IJsland en het Verenigd Koninkrijk. De bedoeling van dit onderzoek was om voor het eerst EU-breed de adoptie van AI-technologieën door bedrijven systematisch in beeld te brengen. De studie betrof de volgende vijf KPI's:

- AI bewustzijn (vier niveaus).
- AI adoptie (vier niveaus, tien AI-technologieën).
- AI herkomst (zes opties).
- Externe obstakels voor AI adoptie (acht categorieën).
- Interne obstakels voor AI adoptie (zeven categorieën).

Het veldwerk is uitgevoerd in het eerste kwartaal van 2020. In Nederland zijn 500 bedrijven bevestigd, ongeveer gelijk verdeeld over groot, midden, klein en micro. Er is geen onderverdeling vooraf naar SBI gemaakt, de kernactiviteit per bedrijf is wel uitgevraagd. Naast analyses van totaalcijfers voor de hele steekproef zijn de verschillende landen kort beschreven. Het is niet duidelijk of het de bedoeling is het onderzoek te herhalen in de toekomst.

Links

Toegang tot AI-landscape dashboard en AI-investments dashboard:

https://knowledge4policy.ec.europa.eu/ai-watch/ai-watch-dashboards_en

DESI-website:

<https://digital-strategy.ec.europa.eu/en/policies/desi>

European enterprise survey on the use of technologies based on artificial intelligence:

<https://op.europa.eu/en/publication-detail/-/publication/f089bbae-f0b0-11ea-991b-01aa75ed71a1>

3.3 Eurostat

Eurostat is het statistische bureau binnen de Europese Commissie. De Nederlandse Enquête ICT-gebruik bedrijven is onderdeel van een door Eurostat gecoördineerd programma waarbinnen informatie over het ICT-gebruik binnen alle EU lidstaten wordt verzameld. Binnen dit programma worden de vraagstellingen in de enquêtes door de lidstaten afgestemd. Het betreft een minimum-programma: landen kunnen extra vragen en detailleringen toevoegen naast wat Europees is afgesproken. Het CBS doet dat in een aantal gevallen ook om tegemoet te komen aan wensen van nationale gebruikers. Cijfers over uitsplitsingen van AI-gebruik naar bedrijfstak en bedrijfsgrootte zijn voor Nederland bijvoorbeeld beschikbaar, maar niet noodzakelijk voor andere EU-landen.

Door Eurostat wordt ook geharmoniseerde informatie verzameld en gepubliceerd over het gebruik van een aantal aan AI gerelateerde technologieën door bedrijven met 10 of meer werkzame personen in de EU-27 landen. Voorsnog is deze informatie alleen beschikbaar voor rapportagejaar 2021. In de toekomst zal het gebruik van AI waarschijnlijk² jaarlijks onderwerp blijven in de enquête die geldt voor alle EU-landen.

Vergelijkbaar met de Nederlandse enquête ICT-gebruik bedrijven gaat het om de volgende indicatoren³:

- Bedrijven in EU-27 landen die één of meer AI-technologieën gebruikt.
- Bedrijven in EU-27 landen die service robots gebruiken.
- Bedrijven in EU-27 landen die een chatbot gebruiken.
- Bedrijven in EU-27 landen die gebruik maken (of dat overwegen) van AI-toepassingen.
- Gebruik van AI-toepassingen bij bedrijven in EU-27 landen.
- Toepassing van AI-technologie bij bedrijven in EU-27 landen naar activiteit.
- Intensiteit inzet van AI-technologie bij bedrijfsfunctie in EU-27 landen.
- Herkomst gebruikte AI-technologieën in EU-27 landen.
- Redenen voor bedrijven in EU-27 landen om geen gebruik te maken van AI-toepassingen.
- Intensiteit inzet bij bedrijven in EU-27 landen van AI-activiteiten.
- Toepassing van AI-technologie bij bedrijven in EU-27 landen naar bedrijfsfunctie

Links

AI gebruik in EU landen (bedrijven met 10 of meer werkzame personen):

https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_eb_ai&lang=en

Nieuwsbericht:

<https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210413-1>

² Dit is afhankelijk van afspraken in Europees verband, waarin het CBS samenwerkt met nationale statistische bureaus van de andere EU-lidstaten en Eurostat

³ Deze lijst komt grotendeels overeen met wat is beschreven onder de ICT enquête maar is voor de volledigheid opgenomen

3.4 Universiteit van Stanford

Het **AI-index report** dat door de Universiteit van Stanford is ontwikkeld heeft als doel valide, diepgaande en uitgebreide gegevens te verstrekken aan beleidsmakers, onderzoekers, journalisten, leidinggevend en het grote publiek om een dieper begrip van het complexe veld van AI te ontwikkelen.

De resultaten worden gepresenteerd in een jaarlijks rapport, maar zijn ook interactief toegankelijk via de 'Vibrancy tool'. Via deze Vibrancy tool kunnen 26 landen (waaronder Nederland) op 22 indicatoren worden vergeleken. Ook zijn de gegevens toegankelijk via een database.

De gebruikte databronnen zijn Elsevier, Microsoft Academic Graph, Arxiv, Nesta, LinkedIn Economic Graph, Crunchbase, CapIQ en NetBase Quid. De cijfers zijn beschikbaar voor de jaren 2015-2020.

Domeinen	Indicatoren
Onderzoek	(Totaal en per capita) <ul style="list-style-type: none">• AI conferentie artikelen• AI conferentie citaties• AI Journal publicaties• AI Journal citaties• AI patenten• AI patent citaties• Deep learning publicaties
Economie	<ul style="list-style-type: none">• Relative skill penetration• AI private Investerings (totaal en per capita)• Bedrijven gefinancierd (totaal en per capita)• AI inhuur index
Inclusie	<ul style="list-style-type: none">• AI vaardigheidspenetratie (vrouwen)• Unieke AI banen door vrouwen

Links

Vibrancy tool:

<https://aiindex.stanford.edu/vibrancy/>

Data:

<https://drive.google.com/drive/folders/1CgMBVb4p6BikFBTziFqdDY4lPkcP1Cij?usp=sharing>

3.5 Universiteit van Oxford

De door de Universiteit van Oxford ontwikkelde **Government AI readiness index** is een tool om de mate van 'government AI readiness' van landen en regio's te vergelijken. De index is beschikbaar voor de jaren 2017-2021. Van jaar op jaar zijn er echter wisselingen in de methodiek, waaronder aanpassingen in de selectie van gebruikte indicatoren. De laatste index voor 2021 bevat vergelijkingen op 44 indicatoren binnen 10 dimensies die weer onderdeel zijn van één van de drie pijlers. Dit geeft het volgende beeld:

Pijlers	Dimensie
Overheid	<ul style="list-style-type: none">• Visie• Governance en ethiek• Digitale capaciteit• Aanpassingsvermogen
Technologie sector	<ul style="list-style-type: none">• Grootte• Innovatieve capaciteit• Human capital
Data en infrastructuur	<ul style="list-style-type: none">• Data representativiteit• Data beschikbaarheid• Infrastructuur

De indicatoren hebben vooral betrekking op de randvoorwaarden waaraan voldaan moet worden voor een goede implementatie van AI in de maatschappij. Deze indicatoren zijn belangrijk, maar hebben een relatief weinig specifieke link met AI. In het jaarlijks gepubliceerde rapport worden negen wereldregio's vergeleken en per wereldregio wordt telkens één 'spotlight country' uitgelicht. In het meest recente rapport (2021) zijn de scores per land beschikbaar voor de overall index en de indices per pijler.

Link

Rapport government AI readiness index 2021:

https://static1.squarespace.com/static/58b2e92c1e5b6c828058484e/t/61ead0752e7529590e98d35f/1642778757117/Government_AI_Readiness_21.pdf

3.6 Techleap

Techleap is een door het Ministerie van Economische Zaken en Klimaat gefinancierd programma dat zich richt op het bevorderen van doorgroei van scale-ups naar grote technologiebedrijven. Daarbij fungeert Techleap als expertisecentrum en onderneemt het samen met stakeholders verschillende acties om in het startup-ecosysteem toegang tot kapitaal, markten, talent en technologische of wetenschappelijke kennis te versterken.

De kennis die Techleap vergaart over wat Nederlandse scale-ups nodig hebben om door te groeien tot internationaal toonaangevende technologiebedrijven, wordt gedeeld met het hele ecosysteem in Nederland.

Onderdeel van Techleap is een **dashboard met cijfermatige informatie over het startup ecosysteem** van Nederland. Via Techleap zijn een aantal AI indicatoren beschikbaar (hieronder weergegeven). De voor deze indicatoren gebruikte databronnen zijn 'Techleap startup finder', 'dealroom.co' en 'LinkedIn'.

- Aantal AI gedreven startups en scale-ups (gebruik makend van AI) naar land en regio (in Nederland).
- Aantal werknemers van AI startups en scale-ups in Nederland.
- Aantal AI core startups en scale-ups (met ten minste één geregistreerd patent gerelateerd aan AI).
- Totale investeringen in AI startups en scale-ups naar land.
- Totale investeringen in Nederlandse AI startups en scale-ups naar land.
- Aantal investeerder in Nederlandse AI startups en scale-ups naar land.
- Aantal AI-startups met een vrouwelijke (co)founder.

Links

Landingspagina website Techleap:

<https://www.techleap.nl/>

Techleap dashboard:

<https://finder.techleap.nl/dashboard>

3.7 Tortoise

Mediaorganisatie Tortoise heeft de **Global AI Index** ontwikkeld. Deze index benchmarkt landen op het niveau van investeringen, innovatie en implementatie van AI. De index is een samengestelde maat die gebaseerd is op 143 afzonderlijke indicatoren binnen zeven domeinen. Deze domeinen zijn talent, infrastructuur, werkomgeving, ontwikkeling, overheidsstrategie en commercieel. De weging van ieder domein is vastgesteld in samenspraak met experts uit het veld.

De data die wordt gebruikt betreft veelal open access bronnen zoals LinkedIn, Github, Meetup, en stack overflow. De index is beschikbaar voor 62 landen waaronder Nederland. De totale scores en de scores per domein zijn (tegen betaling) beschikbaar voor deze landen. Het gaat om een real-time dashboard, wat betekent dat de gepresenteerde scores betrekking hebben op de huidige situatie.

Link

Landingspagina Global AI-index:

https://www.tortoisemedia.com/intelligence/global-ai/?mc_cid=20a029949f&mc_eid=9b10aeb39c

3.8 Elsevier

In het voor de Nederlandse AI Coalitie uitgevoerde onderzoek 'State of AI Research in The Netherlands' worden resultaten van onderzoek op de SCOPUS database gepresenteerd. De resultaten betreffen de jaren 2010-2020 en dekken een variëteit aan landen waaronder Nederland. In het onderzoek worden onder andere het aantal AI publicaties (en citaties hiervan), het aantal actieve patenten (en citaties hiervan), marktdekking van patenten en de mobiliteit van en samenwerkingen tussen AI onderzoekers in kaart gebracht.

- Aantal en aandeel van wetenschappelijke AI publicaties in SCOPUS database per land
- Aantal citaties van AI publicaties
- Top 1% meest geciteerde publicaties
- Aantal actieve AI patenten
- Marktdekking AI patenten
- Aantal citaties van AI patenten
- AI Publicaties door vrouwen
- Mobiliteit van AI onderzoekers
- Samenwerking tussen AI onderzoekers

Naast deze resultaten kan Elsevier nog verschillende andere indicatoren samenstellen, op basis van het bronnenmateriaal dat ze al in huis hebben.

4. Vervolgstappen

Met nadruk vermelden we dat deze inventarisatie van beschikbare bronnen niet bedoeld is als volledig overzicht. Dat is gezien de korte beschikbare tijd en de tamelijk open vraagstelling niet mogelijk. Om te komen tot een selectie van indicatoren die een evenwichtig en betrouwbaar beeld geven van de ontwikkeling van AI in Nederland, kortweg de AI-monitor, zouden de volgende stappen genomen moeten worden.

1. Het doel en de beoogde reikwijdte van de AI-monitor moet worden vastgesteld in een conceptueel kader. Deels ligt dit kader vast op basis van al gemaakte Groeifonds-gerelateerde afspraken, deels levert TNO een bijdrage in de vorm van studie parallel aan deze inventarisatie en deels zal het kader geleidelijk verder moeten groeien op basis van voortschrijdend inzicht.
2. Het in dit document gepresenteerde overzicht moet worden vergeleken met het conceptuele kader om te bepalen welke indicatoren waar passen. De domeinen waarop de indicatoren betrekking hebben zullen in kaart moeten worden gebracht. Vervolgens kan worden nagegaan in welke mate de reeds beschikbare indicatoren de domeinen in de AI monitor waarover rapportage gewenst is dekken.
3. Voor de domeinen met onvoldoende 'dekking' kan (eventueel in een later stadium) gekeken worden of er nieuwe betekenisvolle indicatoren ontwikkeld of afgeleid kunnen worden. Een deel van de huidige indicatoren kan overigens als afgeleid worden beschouwd.
4. De kwaliteit van de beschikbare indicatoren en afleidingen moet in kaart worden gebracht. Per indicator of bron zou bijvoorbeeld een indruk verkregen moeten worden van de mate waarin deze geschikt is voor het betrouwbaar in kaart brengen van ontwikkelingen over de tijd. Deze afweging zal niet zwart-wit zijn, maar wanneer er kanttekeningen bij het gebruik van bepaalde indicatoren te maken zijn, is het essentieel hierover transparant te zijn bij eventuele publicatie in de AI-monitor. Een keuze kan ook zijn om indicatoren met beperkte kwaliteit voorlopig niet te publiceren.
5. Bij de selectie van indicatoren speelt de mate waarin deze toegankelijk en te gebruiken zijn voor een doel als een AI-monitor een belangrijke rol. Voor externe indicatoren die ontwikkeld zijn door commerciële instellingen of non-profit organisaties gelden waarschijnlijk andere gebruiksvoorwaarden dan voor bestaande CBS-indicatoren.

Over het tijdsplan van deze stappen en de mate van diepgang van het proces zal moeten worden nagedacht bij het ontwikkelproces voor de AI-monitor. De voorstudiefase heeft geleerd dat het goed in beeld brengen van het hele terrein van de AI-monitor ingewikkeld en arbeidsintensief is.

Verklaring van tekens

Niets (blanco)	Een cijfer kan op logische gronden niet voorkomen
.	Het cijfer is onbekend, onvoldoende betrouwbaar of geheim
*	Voorlopige cijfers
**	Nader voorlopige cijfers
-	(indien voorkomend tussen twee getallen) tot en met
0 (0,0)	Het cijfer is kleiner dan de helft van de gekozen eenheid
2023-2024	2023 tot en met 2024
2023/2024	Het gemiddelde over de jaren 2023 tot en met 2024
2023/'24	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2023 en eindigend in 2024
2021/'22-2023/'24	Oogstjaar, enz., 2021/'22 tot en met 2023/'24

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.
Verbeterde cijfers in de staten en tabellen zijn niet als zodanig gekenmerkt.

Colofon

Uitgever

Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress

Centraal Bureau voor de Statistiek

Ontwerp

Edenspiekermann

Inlichtingen

Tel. 088 570 70 70
Via contactformulier: www.cbs.nl/infoservice

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen/Bonaire, 2024.
Vereenvoudigen is toegestaan, mits het CBS als bron wordt vermeld.