

**Goede  
Doelen  
Neder-  
land**

SAMEN VOOR EEN  
STERKE SECTOR

# Big data voor goede doelen

Verkenning van de toekomst 2021

'Big data' is in  
essentie een  
grote hoeveelheid  
georganiseerde  
informatie.

**Colofon**

Goede Doelen Nederland  
James Wattstraat 100  
1097 DM Amsterdam  
e-mail: [info@goededoelennederland.nl](mailto:info@goededoelennederland.nl)  
website: [verkenning.goededoelennederland.nl](http://verkenning.goededoelennederland.nl)  
website: [www.goededoelennederland.nl](http://www.goededoelennederland.nl)

Dit rapport is geschreven door BAKEN AI, in opdracht van  
Goede Doelen Nederland.

Vormgeving: Beeldr, Leiden  
© december 2021, Goede Doelen Nederland

Deze uitgave is met de grootst mogelijke zorgvuldigheid  
samengesteld. Goede Doelen Nederland aanvaardt geen  
aansprakelijkheid voor onjuistheden in deze uitgave.

# INHOUDSOPGAVE

## EXECUTIVE SUMMARY

## INLEIDING

### 1. WAT IS BIG DATA?

#### 1.1 Waar komt big data vandaan?

#### 1.2 Verschillende definities van big data

1.2.1 De vijf criteria van big data

1.2.2 Big data als grondstof

#### 1.3 Big data: structureren en gebruiken

#### 1.4 De laatste big data-ontwikkelingen

1.4.1 Data-ethiek

1.4.2 Regulering

1.4.3 Cloud, SaaS en open source

1.4.4 Privacy en transparantie: edge computing, federated learning en explainable AI

## 2. BIG DATA IN DE PRAKTIJK

### 2.1 Big data in de financiële sector: duurzaam en verantwoord investeren

2.1.1 De data-uitdaging van ESG

2.1.2 Big data in de ESG-praktijk

2.1.3 Impactmeting voor goede doelen

### 2.2 Big data in de zorg: toegankelijk maatwerk

2.2.1 Big data-uitdagingen in de zorg

2.2.2 Federated learning voor veilige en vertrouwde datadeling

2.2.3 Explainable AI: artsen helpen bij code zwart door COVID-19

2.2.4 Lessen uit de zorg voor goede doelen

## 3. BIG DATA VOOR GOEDE DOELEN

### 3.1 Big data voor fondsenwerving en -marketing

3.1.1 Donateur lifetime value-optimalisatie

3.1.2 Frictieloos digitaal doneren met embedded finance

3.1.3 Donateurs binden met immersive storytelling

3.1.4 Datagedreven transparantie

### 3.2 Big data voor doelbesteding

3.2.1 Big data-filantropie

## 4. AAN DE SLAG MET BIG DATA



## EXECUTIVE SUMMARY

- 'Big data' is in essentie een grote hoeveelheid georganiseerde informatie. Die kan heel nuttig zijn: door big data te structureren en vervolgens te 'gebruiken', kom je tot nieuwe inzichten.
- Het nut van big data is niet vanzelfsprekend: hoe meer data, hoe moeilijker die te structureren is. Ook het gebruik kan uitdagend zijn. Net als elke grondstof, moet gestructureerde data eerst worden verwerkt voordat je er iets aan hebt. Soms zijn daar complexe analysemethodes voor nodig – bijvoorbeeld kunstmatige intelligentie (AI).
- Je kunt big data op drie manieren gebruiken: beschrijvend (wat gebeurt er?), diagnostisch (waarom gebeurt het?) of voorspellend (wat gaat er gebeuren?).
- De mogelijkheden van big data worden enerzijds bepaald door technologie, en anderzijds door wetgeving en ethische kwesties. Niet alles wat kán, mag. **Data-ethiek** heeft te maken met zaken als privacy, transparantie, vooroordelen en governance.
- Big data-wetgeving bestaat vooral uit de Algemene verordening gegevensbescherming (AVG). De wetgeving staat niet stil, blijkt uit recente AI-regulering. Er zijn veel nieuwe richtlijnen en tools voor verantwoord data- en algoritmebeleid.
- De belangrijkste big data-technologie is **cloud computing**. Daarvoor bestaan zowel Software-as-a-Service als Open Source-aanbieders. **Edge computing, federated learning** en **explainable AI** zoeken specifieke oplossingen voor privacy- en transparantiekwesties.
- Twee cases laten het nut van big data zien. De case over **big data om de duurzaamheid van beleggingen (ESG) te meten**, toont hoe impactevaluatie en doelbesteding goede doelen kunnen helpen ontwikkelen. De case over **big data in de zorg** legt uit hoe maatschappelijke organisaties verantwoord kunnen omgaan met datadeling en transparante AI.
- In de goededoelensector wordt big data voor **fondsenwerving en marketing** gebruikt. Denk aan datagedreven werving en behoud van donateurs (donateur lifetime value-optimalisatie) en frictieloos digitaal doneren (embedded finance).
- Tools voor **donateur lifetime value**-optimalisatie vind je bij: (i) gevestigde softwareleveranciers; (ii) AI-producten voor goede doelen en (iii) start-ups en hackathons die oplossingen op maat bouwen. Onder **embedded finance** vallen bijvoorbeeld afronden-voor-een-goed-doel-apps, contactloos betalen en cross-platform campagnes op social media.
- **Data-driven storytelling** verhoogt de transparantie van goede doelen. Via **blockchain-technologie** zit die transparantie in de toekomst zelfs fundamenteel ingebouwd bij goede doelen. Met behulp van **open data** en **AI-algoritmes** zijn veel goede doelen wereldwijd in staat om hun impact groter en inzichtelijker te maken.
- **Chatbots, virtual reality** en **videogames** bieden nieuwe mogelijkheden voor goede doelen. Ze vertellen donateurs op nieuwe, krachtige en creatieve manieren over je missie. Ook data zelf kan voor maatschappelijke impact zorgen, door **big data-doelbesteding**. Data doneren (**big data-filantropie**) biedt nieuwe bestedingskansen voor goede doelen.
- De lessen van dit rapport zijn samengevat in de **Goede Doelen Big Data Roadmap**. De belangrijkste vraagstukken zijn daarin uitgewerkt, met behulp van subvragen per niveau.

# Data is niet nieuw. Wél nieuw is het volume van het datagebruik.



## INLEIDING

Tussen al het nieuws over de ongekende mogelijkheden én gevaren van datatechnologie, zou je bijna vergeten dat 'data' eigenlijk niets nieuws is. De mensheid verzamelt en analyseert al millennialang informatie: van Babylonische marktmeesters tot koloniale wereldrijken, van een volkstelling tot de telefoongids. Stuk voor stuk manieren om de wereld te doorgronden en controleren met behulp van data.

Data is dus niet nieuw. Wél nieuw is het volume van het datagebruik. Stel je het aantal Google-zoekacties wereldwijd per minuut voor (3,8 miljoen), of het aantal pintransacties per dag in Nederland (12,2 miljoen). Ons leven speelt zich steeds meer digitaal af. Allerlei sensoren nemen die groeiende digitale wereld waar en leggen een oceaan aan informatie vast. Vandaar de term: **big data**.

Om precies te zijn, draait big data om verschillende typen data beschikbaar en bruikbaar maken. En om de technologische, wettelijke en ethische kaders waarbinnen dat gebeurt.

### Big data in de goededoelensector

In dit rapport brengen we big data helder in kaart. We combineren inzicht in de laatste technieken met actuele cases van binnen en buiten de goededoelensector. Het rapport is onderdeel van het programma 'Verkenning van de toekomst' van Goede Doelen Nederland.

Medewerkers van goede doelen halen uit dit rapport:

- kennis over de vormen en toepassingen van big data in verschillende sectoren;
- ideeën voor big data-toepassingen binnen hun eigen organisatie;
- handvatten om die toepassingen zelf te realiseren.

Het eerste hoofdstuk geeft een algemeen beeld van big data. Hoe kunnen we dit grote containerbegrip het beste benaderen? En wat zijn de laatste ontwikkelingen op technisch, regulatorisch en ethisch vlak? Die vragen vormen de basis voor de rest van het rapport.

In het tweede hoofdstuk bekijken we voorbeelden uit andere sectoren. Zo meet de **financiële sector** de Environmental, Social & Governance-impact (ESG) van investeringen. Dat leidt tot concrete doelen, zoals de Sustainable Development Goals (SDG) van de VN. In de **zorg** bundelen verschillende ziekenhuizen hun data op een veilige en betrouwbare manier, zodat ze overal gepersonaliseerde zorg kunnen leveren.

In het derde hoofdstuk ligt de focus op goede doelen. Welke taken in deze sector kunnen we verbeteren met behulp van big data, en welke voorbeelden zien we hiervan wereldwijd? Die variëren van betere fondsenwerving tot sterkere donateursloyaliteit. Het rapport eindigt met een roadmap voor goede doelen van elk formaat om zelf met big data aan de slag te gaan.



## 1. WAT IS BIG DATA?

Big data wordt zó regelmatig en zó verschillend gebruikt, dat het lastig kan zijn om er een helder beeld van te vormen. Om de kansen en uitdagingen van big data voor goede doelen te begrijpen, moeten we de term eerst ontleden.

### 1.1 Waar komt big data vandaan?

Het begrip 'big data' is sinds de jaren negentig van de vorige eeuw in gebruik. Volgens de meeste bronnen heeft de Amerikaanse computerwetenschapper John Mashey de term bedacht en gepopulariseerd.

Maar de werkelijke oorsprong van het fenomeen ligt verder terug: al in de jaren zestig en zeventig werden de eerste datacenters gebouwd. Het concept van een database is toen ontstaan. Door digitale rekenkracht en gegevensopslag waren opeens veel grotere hoeveelheden informatie beschikbaar dan voorheen. Big data was geboren.

### 1.2 Verschillende definities van big data

Digitalisering en de snelle ontwikkeling van computertechnologie zijn de drijvende krachten achter de opkomst en verspreiding van big data. De term is inmiddels razend populair. Maar wat bedoelen we precies als we het hebben over big data? Laten we de term opsplitsen in 'big' en 'data', en kijken naar de veelgebruikte definities.

'Data' is niets meer dan **gestructureerde informatie**. Dat kan alles zijn: persoonsgegevens, het aantal bomen in een bos of de temperatuur van je autobanden. Om het 'gestructureerd' te noemen, moet de data wel bruikbaar zijn: je moet bij wijze van spreken aan je spreadsheet kunnen zien op welk moment je bomen moet kappen, of wanneer je banden aan vervanging toe zijn.

'Big' duidt op de **grote omvang** van de gestructureerde informatie. Vroeger telde een bosbouwer zijn bomen zelf, tegenwoordig registreren alle Nederlandse gemeenten hun bomen in databases waardoor een volledige **bomenkaart** voor Nederland mogelijk is. De hoeveelheid beschikbare data is de afgelopen decennia exponentieel gegroeid.

Hoewel data van alle tijden is, maakt digitale technologie dus een wereld van verschil. Onze leefomgeving is tegenwoordig te vertalen naar één gezamenlijk systeem van enen en nullen (binare code), waarmee we een zo goed als oneindige hoeveelheid informatie kunnen opslaan.

#### 1.2.1 De vijf criteria van big data

Big data gaat dus over omvangrijke, gestructureerde informatie. Toch kunnen we de term nog scherper formuleren.

Volgens een bekende definitie zijn er vijf criteria (of assen) die data 'big' maken:

1. **Volume:** de omvang van de data.
2. **Diversiteit:** het aantal verschillende soorten bronnen dat wordt gebruikt.
3. **Snelheid:** de snelheid waarmee data wordt gegenereerd.
4. **Waarheidsgetrouw:** de betrouwbaarheid van de data.
5. **Waarde:** in hoeverre men in staat is waarde uit de data te halen.

Hoe meer, hoe beter. Maar het aantal rijen en kolommen in een dataset is dus niet het enige wat maakt dat we over big data kunnen spreken. Het is ook belangrijk dat de informatie van verschillende soorten bronnen komt: denk aan video, audio en social media-teksten. Het ligt niet voor de hand om deze samen te brengen en te gebruiken, maar toch is het precies deze diversiteit die een dataset 'big' maakt.

Ook snelheid is een factor. Tijdens een Formule 1-race moet de bandentemperatuur bijvoorbeeld *realtime* worden gemeten en weergegeven, en niet een paar seconden later.

Data uit uiteenlopende bronnen of van verschillende snelheden, kan ook afwijken in kwaliteit. Dat kan de dataset onbetrouwbaar maken. Ook andere factoren spelen mee: meetfouten, tyfouten, spreektaal of onregelmatige afkortingen kunnen een database als bron minder betrouwbaar maken. Big data moet voldoende waarheidsgetrouw zijn.

Tot slot is big data pas écht nuttig als je er waarde uit kunt halen: je moet er iets van leren en dat gericht kunnen toepassen. Daarvoor heb je als organisatie een IT-infrastructuur en analysetechnieken als *machine learning* nodig. Neem Netflix, dat *machine learning* gebruikt om aanbevelingen te optimaliseren op basis van data over het kijkgedrag van gebruikers.

### 1.2.2 Big data als grondstof

Big data moet dus voldoen aan kenmerken als diversiteit en snelheid. En niet elke gebruiker kan ermee overweg. Maar nu big data steeds meer toepassingen heeft, wordt vaak de vergelijking met een grondstof gemaakt: big data is 'de nieuwe olie'. Een essentiële brandstof voor de wereldeconomie – of we dat nu willen of niet. Big data wordt dat ook, of is het al.

Net als bij olie is het **gebruik van big data** niet universeel. De informatie moet eerst verwerkt of *processed* worden: ruwe olie is niet meteen bruikbaar; een 'rauwe bulk' data evenmin. En zoals niet ieder voertuig dezelfde olie nodig heeft, gebruikt niet elke organisatie hetzelfde type data. De ene sector vereist meer *processing* dan de andere.

### 1.3 Big data: structureren en gebruiken

Zoals gezegd, horen er twee handelingen bij big data: grote hoeveelheden informatie structureren, en die vervolgens gebruiken. Beide zijn een potentieel struikelblok. Structureren wordt moeilijker naarmate de data 'bigger' wordt door meer volume, diversiteit of snelheid. Tegelijkertijd wordt het gebruik van big data complexer, waardoor geavanceerde analysetechnieken als AI nodig zijn.

Maar het nut van big data als informatiebron groeit ook. Organisaties baseren hun beslissingen erop. Met behulp van big data beantwoorden ze belangrijke vragen: van 'wat zijn onze operationele kosten?' tot 'welke klanten gaan ons binnenkort verlaten?'. Daarvoor zijn verschillende soorten informatie nodig: de eerste vraag is met één flinke optelsom beantwoord, maar voor de tweede is een voorspelling nodig op basis van uiteenlopende gegevens. Data kan dus op allerlei manieren worden gebruikt, met verschillende uitkomsten.

Meestal wordt big data op een van de volgende drie manieren<sup>1</sup> gebruikt:

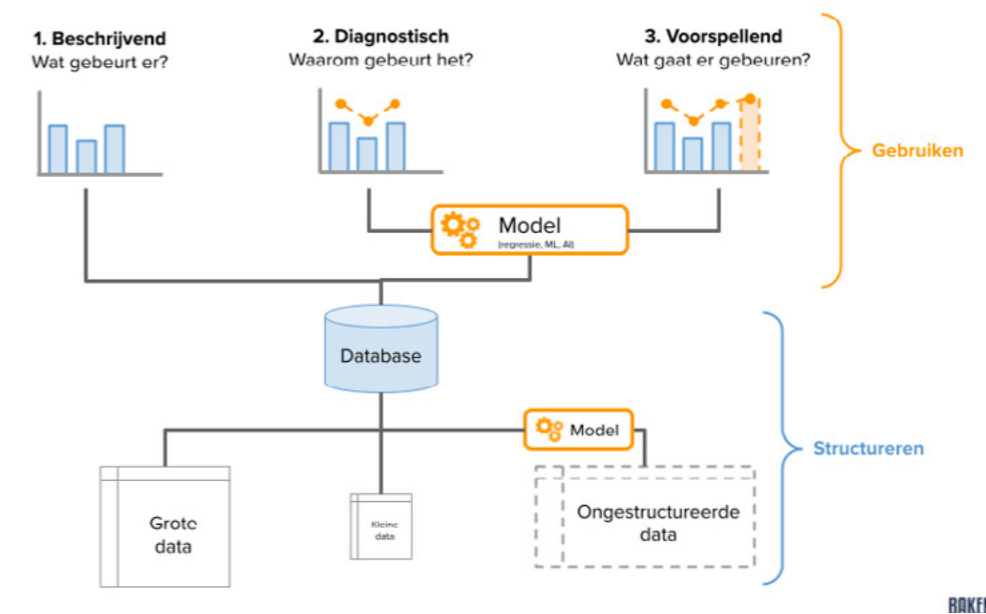
1. **Beschrijvend:** wat gebeurt er?
2. **Diagnostisch:** waarom gebeurt het?
3. **Voorspellend:** wat gaat er gebeuren?

Typisch **beschrijvend gebruik** is een dashboard dat actuele informatie samenbrengt en de huidige situatie weergeeft. Bijvoorbeeld met een grafiek die de inkomsten van je organisatie per maand laat zien, of de kosten in een bepaalde categorie.

**Diagnostisch gebruik** gaat verder en vertelt wáárom de situatie zo is. Denk aan een analyse van de factoren die je inkomsten beïnvloeden, zoals advertenties en fondsenwerving. Breng je die in kaart, dan krijg je inzicht in de oorzaken achter de cijfers. De bekendste methode voor diagnostisch gebruik is lineaire regressie (een statistische methode waarmee je een lineair verband tussen twee factoren kan aanduiden).

<sup>1</sup> Strikt genomen is er nog een vierde functie, namelijk voorschrijvend: 'Wat moet ik doen?'. Hierbij bepaalt big data ook de optimale actie van de organisatie.

**Voorspellend gebruik** gaat nóg een stap verder. Je gebruikt data niet alleen om inzicht te krijgen in de huidige situatie, maar ook in de toekomst. Zo vind je een antwoord op de vraag 'welke klanten gaan ons binnenkort verlaten?' Hiervoor gebruik je bijvoorbeeld een voorspelmodel met data van eerder vertrokken klanten, dat de vertrekans van huidige klanten inschat. Voorspellende toepassingen maken vaak gebruik van de meest complexe big data-technologie, zoals machine learning en AI.



**Figuur 1.** Zo werkt big data

**Structureren** betekent ruwe data van verschillende typen en formaten samenbrengen in één of meerdere databases of -sets. Tekstberichten zijn een typisch voorbeeld van ongestructureerde data. E-mails van klanten bevatten allerlei relevante informatie om de vraag 'welke klanten gaan ons binnenkort verlaten?' te beantwoorden. Maar die data moet er wel eerst uit worden gehaald. Vertaal je de tekst van een grote hoeveelheid e-mails bijvoorbeeld naar sentimentscores, dan wordt de data gestructureerd en bruikbaar.

Enmaal gestructureerd kun je big data op drie manieren **gebruiken**: beschrijvend, diagnostisch of voorspellend. Beschrijvend datagebruik is eigenlijk niets meer dan de informatie op een rijtje zetten; er komt weinig verwerking bij kijken. Voor diagnostische en voorspellende toepassingen moet je data *processen* met regressiemethoden of AI-modellen. Zo kun je de data gebruiken voor meer geavanceerde modellen, apps of algoritmen.

### 1.4 De laatste big data-ontwikkelingen

We weten nu wat het begrip big data inhoudt: grote hoeveelheden data structureren en gebruiken. Maar welke data kun je verzamelen, en hoe ga je er vervolgens mee aan de slag? Het domein van big data wordt gevormd door technologie, maar ook door wetgeving en ethische overwegingen. Technologie bepaalt wat *kan*, wetgeving wat *mag* en ethiek wat *verantwoord* is.

Samen sturen deze factoren de ontwikkeling van big data. In de volgende paragrafen bespreken we de actuele trends.

### 1.4.1 Data-ethiek

Ethische data-vraagstukken zijn vrijwel dagelijks in het nieuws. Denk aan datalekken, discriminerende algoritmes en gevaarlijke socialmedia-accounts die worden geblokkeerd. Stuk voor stuk hebben deze gevallen te maken met goed of verkeerd gebruik van big data. In de maatschappij zijn fundamentele zorgen en wensen over big data. Dit uit zich in zowel nieuwe wetgeving als in technologische innovatie.

Data-ethiek draait om wat *verantwoord* is op het gebied van digitaal informatiegebruik. Met andere woorden: wat binnen de grenzen blijft van wat we als samenleving aanvaardbaar vinden.

De belangrijkste kwesties op dat gebied zijn:

- privacy
- transparantie
- bias (vooroordelen)
- governance

**Dataprivacy** is de bekendste van van dit rijtje. Personen hebben rechten als het om hun persoonsgegevens gaat. Wie deze gegevens verwerkt, moet die rechten respecteren. Organisaties mogen niet zomaar beslissingen nemen op basis van informatie die ze over je hebben verzameld. Dat kan tot ongelijke behandeling leiden, wat tegen de grondwet ingaat. Daarom is het belangrijk dat je controle hebt over je persoonsgegevens. Je moet ze kunnen inzien, en beslissen wat een organisatie wel of niet van jou mag weten.

**Datatransparantie** ligt in het verlengde van dataprivacy. Naast controle over gegevens, moeten organisaties ook laten zien wat ermee gebeurt; hoe conclusies met behulp van big data tot stand komen. Ook complexe big data-modellen moeten tot op zekere hoogte uitlegbaar zijn. Stel je voor dat je geen lening kunt krijgen, maar de bank kan niet verantwoorden waarom – de beslissing is nu eenmaal met een onbegrijpelijk AI-model gemaakt. Dat zouden we niet accepteren.

Met **data bias** bedoelen we vooroordelen die er als het ware ‘insluipen’ bij databases en algoritmes. Data bias treedt op wanneer de beschikbare data niet representatief is voor wat er wordt geanalyseerd. Doordat er verklarende variabelen of waarnemingen ontbreken, of doordat de ‘bias’ al bestond tijdens het verzamelen van de data.

Big data kan vooroordelen die in de maatschappij leven weerspiegelen, en zelfs versterken. Hoe dat werkt, leggen we uit met een voorbeeld:

*Stel je een big data-analyse voor waaruit blijkt dat personen met bepaalde kenmerken (bijvoorbeeld mannen van middelbare leeftijd uit een bepaalde wijk) iets vaker dan gemiddeld een illegaal voorwerp bij zich dragen. Vervolgens krijgen politieagenten de instructie om vaker te controleren bij mensen met bovengenoemde kenmerken; de politie wil immers van die illegale voorwerpen af. Maar door te controleren en de uitkomsten te noteren, creëren de agenten nieuwe, scheef verdeelde data: mensen met de eerder genoemde kenmerken zijn vaker gecontroleerd, dus komen ze ook vaker in de resultaten voor. De volgende keer is het effect nóg sterker, enzovoorts.*

Tot slot slaat **data governance** op verantwoordelijkheid en controlemiddelen bij organisaties die data verwerken. Die moeten op een ethische manier zijn ingericht. Door middel van protocollen en een gezonde cultuur moeten organisaties dataprivacy en -transparantie garanderen en data bias tegengaan.

### 1.4.2 Regulering

Er is de afgelopen jaren veel in regelgeving vastgelegd over big data en data-ethiek.

De belangrijkste Nederlandse en Europese wetten op dit vlak zijn:

- De Algemene verordening gegevensbescherming (**AVG**) uit 2018: Europese privacywetgeving (ook bekend als GDPR).
- De Uitvoeringswet van de AVG (**UAVG**) uit 2018: Nederlandse invulling van de AVG.
- De Digital Services Act (**DSA**) uit 2020: specifiek gericht op digitale intermediairs, van internetproviders tot De Digital Markets Act (**DMA**) uit 2020: specifiek gericht op dominante online platforms, zogeheten ‘poortwachters’.
- De **AI Act** uit 2021: risicoclassificatie en bijbehorende eisen voor AI-systemen.

Uit deze lijst zijn vooral de AVG en UAVG bekend en relevant voor dit rapport. Maar de zeer recente DSA, DMA en AI Act laten zien dat Europese datawetgeving niet stilstaat.

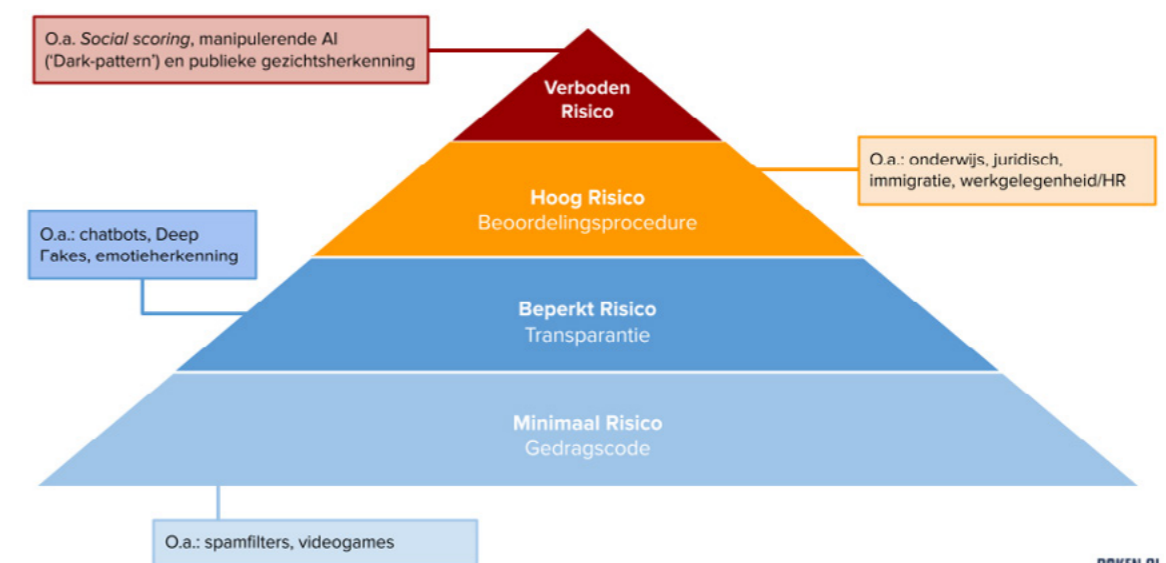
De invoering van de AVG en UAVG in 2018 ging niet zonder slag of stoot. Organisaties bleken niet voorbereid en wisten niet hoe ze de wet moesten interpreteren. Nog steeds is niet alle onduidelijkheid opgelost, het kost nu eenmaal tijd om jurisprudentie op te bouwen. Maar het besef van data-verantwoordelijkheid is wel degelijk gegroeid.

Dat heeft alles te maken met de AVG, die mensen het recht geeft op:

- inzicht, rectificatie, wijziging van toestemming en verwijdering van persoonsgegevens;
- informatie over het gebruik van persoonsgegevens.

Tegelijkertijd wordt het gebruik van persoonsgegevens door AI steeds complexer. Dat schuurt met het recht op informatie: het is niet altijd mogelijk om het precieze gebruik te achterhalen. Daarom heeft AVG-waakhond Autoriteit Persoonsgegevens aparte **adviezen** gepubliceerd over het gebruik van algoritmes. Ook de Europese AI Act is gericht op transparantie van algoritmes.

De EU-regelgeving is ook bedoeld als aanzet om zelf na te denken over verantwoord AI-gebruik. Het doel en de risico's van AI bepalen mede of het geoorloofd is. Zo is **social scoring**, waarbij je mensen beloont of afrekent op basis van hun gedrag, verboden. Maar spamfilters die van dezelfde technologie gebruikmaken, zijn wel toegestaan.



**Figuur 2.** De Europese AI-riscoschalen



De AI Act is op zijn beurt gebaseerd op de [European Ethic Guidelines for AI](#), de Europese richtlijn voor verantwoord AI-gebruik en de toetsing daarvan.

Er zijn steeds meer tools beschikbaar om organisaties hierbij te helpen. Een aantal voorbeelden:

- De Data Protection Impact Assessment ([DPIA](#)) van de Autoriteit Persoonsgegevens: verplichte toets voor projecten met een hoog risico volgens de AVG, ook bruikbaar voor andere organisaties.
- De Data Ethics Decision Aid ([DEDA](#)) van de Universiteit Utrecht: handboek (met [poster](#)) voor stakeholders binnen een dataproject.
- De [Practical Data Ethics](#) van fast.ai: gratis online cursus data-ethiek voor beginnende analisten en datawetenschappers.
- De [Ethical Toolkit for Engineering/Design Practice](#) van het Markkula Center: bekende technische toolkit binnen de datawetenschappen.
- De [Deon](#)-checklist: nuttige tool voor developers om data-ethiek continu in hun werk te betrekken.

De steeds strengere wetgeving staat dus niet op zichzelf. Er zijn genoeg tools die je kunt gebruiken om proactief en verantwoord met big data te werken.

#### 1.4.3 Cloud, SaaS en open source

Sinds het ontstaan van big data in de jaren zestig is er veel veranderd. Zaalvullende machines werden eerst personal computers (PC's) die op je bureau pasten, en vervolgens smartphones in je broekzak. Dat is mede te danken aan *cloud computing*, de laatste in een reeks innovaties die big data steeds toegankelijker maken.

**Cloud computing** is data-opslag of rekenkracht leveren via internet. De mogelijkheden zijn enorm: via iCloud heb je ontelbare terabytes aan opslagruimte, en via Google Cloud Platform voeren duizenden computers met één druk op de knop complexe berekeningen uit. Jij of je organisatie hoeft dus niet zelf meer voor opslag of rekenkracht te zorgen: die koop je gewoon in via clouddiensten als Amazon Web Services, Salesforce of Cisco.

De cloud maakt het ook mogelijk om software geheel online aan te bieden, dat noemen we **Software-as-a-Service** (SaaS). Je haalt een programma niet zelf in huis, maar koopt een abonnement op een online dienst. Voorbeelden zijn de klantbeheersystemen (CRM) van Salesforce en Microsoft, de nieuwsbrief-client van Mailchimp en [Databricks](#), een bedrijf dat data-opslag, -management en AI-analyses als clouddiensten aanbiedt.

Cloud computing maakt big data toegankelijker door opslagruimte, rekenkracht en de nodige software online aan te bieden. Je kiest makkelijk welke dienst je nodig hebt en kunt meteen aan de slag. AI hangt er natuurlijk een prijskaartje aan, want het zijn bedrijven die SaaS aanbieden. Maar er is een alternatief.

**Opensourcesoftware** is gratis en voor iedereen beschikbaar. De bedoeling is dat je de programma's gewoon kunt downloaden, gebruiken en zo nodig aanpassen. Meestal is er een online gemeenschap van developers die het programma steeds verder ontwikkelt.

Een ander belangrijk verschil tussen opensourcesoftware en SaaS, is transparantie. SaaS 'draait' op een aparte server en wordt door jouw computer van afstand 'aangeropen'. Dat maakt SaaS minder transparant – je kunt minder goed zien wat de software precies doet. Opensourcesoftware draait wél op je eigen computer. Die programma's zijn dus volledig transparant: je kunt de broncode zelf inzien.

SaaS is makkelijk te gebruiken, maar je raakt er snel van afhankelijk. Dat noemen we *vendor lock-in*: gebruikers kunnen niet makkelijk aanpassingen maken of overstappen, omdat ze de software niet zelf in huis hebben. Het nadeel van opensourcesoftware is dat het gebruik meer ervaring vraagt. Je moet de publieke broncode zelf

downloaden en implementeren. Bekende voorbeelden van opensourcesoftware zijn [Python](#) en [Wordpress](#). De komst van cloud computing heeft big data toegankelijker gemaakt. Er is een waar **ecosysteem** van SaaS- en opensource-programma's ontstaan, die leunen op big data en waar iedereen om allerlei redenen gebruik van kan maken. Sluitstuk zijn de 'open standaarden': afspraken die ervoor zorgen dat verschillende systemen goed met elkaar kunnen communiceren. Daar heb je voortdurend plezier van, bijvoorbeeld wanneer je telefoon verbinding maakt met een wifi-netwerk. Bekende big data-standaarden zijn SQL, XML, HTML en [Hadoop](#).

#### 1.4.4 Privacy en transparantie: edge computing, federated learning en explainable AI

Cloud computing heeft de ontwikkeling van big data dus versneld. Dat brengt ook nieuwe privacyrisico's met zich mee, want data gaat steeds vaker heen en weer tussen organisaties en online dienstverleners. Maar zijn je persoonsgegevens nog wel toegankelijk of veilig als ze via via bij één of meerdere SaaS-aanbieders terecht komen?

**Edge computing** is een nieuwe technologie die hierop inspeelt. Doel is om te zorgen dat er minder data wordt uitgewisseld tussen cloud en computer. Hoe dat kan? Door zo veel mogelijk handelingen en berekeningen uit te voeren 'aan de rand' van het netwerk, oftewel op je eigen apparaat. Dat betekent minder dataverkeer, en dus een lager privacyrisico. Daarbij moet gezegd dat edge computing nog voor flinke uitdagingen staat.

Een bekend voorbeeld van edge computing is de automatische piloot van een Tesla. Die moet zó snel zó veel data verwerken, dat er simpelweg geen tijd is om dat via de cloud te doen. Het moet in de auto zelf gebeuren.

**Federated learning** is niet gericht op dataverkeer, maar op datacentralisatie. Veel data op één plek verzamelen brengt namelijk ook risico's met zich mee, maar dat is wel nodig voor big data-oplossingen en -algoritmes. Door algoritmes decentraal te trainen, neem je die bezwaren weg. En dat is precies wat federated learning doet. Data hoeft niet te worden samengebracht, maar kan bij organisaties zelf blijven staan. Zo gebruiken ziekenhuizen gezamenlijke data bijvoorbeeld om algoritmes te trainen, zonder dat hun data ooit 'buiten de muren' komt (zie 2.2 voor meer informatie over federated learning).

We weten inmiddels dat transparantie een belangrijk aandachtspunt voor big data is, waar in rap tempo nieuwe wetgeving voor wordt aangenomen (zie 1.4.2). Begrijpelijk, want AI – een belangrijke technologie voor big data – staat bekend als een 'black box'. Precies navertellen hoe beslissingen tot stand komen, is vaak niet mogelijk. Maar er wordt aan oplossingen gewerkt.

**Explainable AI** is zo'n oplossing voor het black box-probleem waar datawetenschappers hard aan werken. Explainable AI doet wat de naam zegt: het legt uit welke beslissingen een AI-systeem maakt, en waarom. Het is bedoeld als manier om verantwoording af te leggen. Om de black box open te breken, komen verschillende technieken bij elkaar. Samen kunnen ze de voorspellingen van een algoritme achteraf uitleggen, zonder de kracht van het algoritme in te perken. Maar de uitkomst wordt wel transparanter. Dat maakt het algoritme juist beter bruikbaar voor organisaties (zie ook 2.2).



## 2. BIG DATA IN DE PRAKTIJK

In dit hoofdstuk bespreken we aan de hand van twee concrete voorbeelden hoe big data in de praktijk wordt gebruikt.

Allereerst kijken we naar de financiële sector, waar start-ups big data gebruiken om duurzame en verantwoorde investeringen te ondersteunen. In het tweede voorbeeld kijken we naar een semipublieke sector: de zorg. We zetten de uitdagingen op privacyvlak in deze sector uiteen en geven een inkijkje in hoe big data de zorg verbetert.

Bij beide cases leggen we een link naar de lessen die goede doelen kunnen leren van deze sectoren. In hoofdstuk 3 gaan we dieper in op de specifieke big data-toepassingen voor goede doelen.

### 2.1 Big data in de financiële sector: duurzaam en verantwoord investeren

In de financiële sector is er steeds meer aandacht voor duurzaam en verantwoord beleggen. Hierdoor groeit de vraag naar betrouwbare indicatoren die de impact van een bedrijf op milieu, maatschappij en governance meten. Deze factoren zijn lastig te kwantificeren. Daarom worden hier vaak kwalitatieve enquêtes voor ingezet, die zijn minder hard dan cijfers.

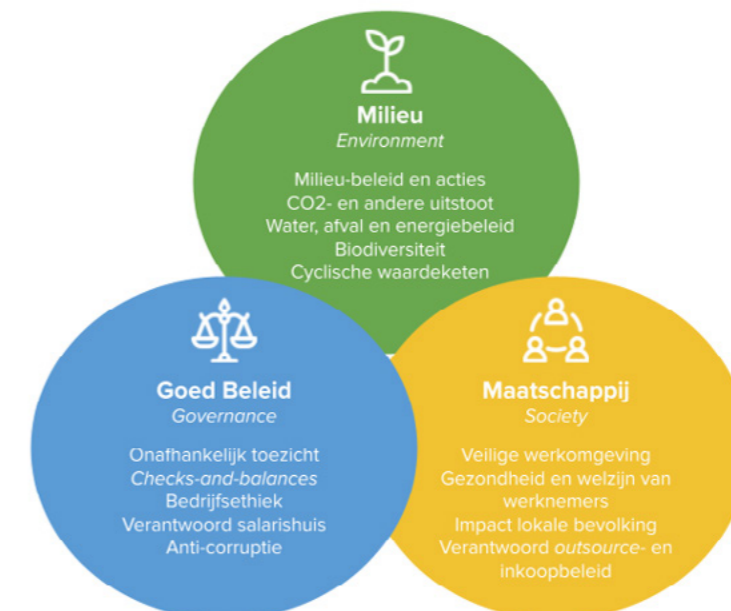
Een kwantitatieve aanpak is mogelijk, dankzij big data en de criteria voor Environment, Social en Governance (ESG). De cases in deze paragraaf laten zien wat big data kan betekenen voor impactevaluatie en doelbesteding van goede doelen.

#### 2.1.1 De data-uitdaging van ESG

Om de impact van een bedrijf op mens en milieu in kaart te brengen, wordt vaak gebruikgemaakt van de ESG-indicatoren.

Deze bestaan uit drie onderdelen:

- **Environment** (milieu) omvat indicatoren als de CO<sub>2</sub>-uitstoot, het energie- en waterverbruik en het afvalbeleid van een organisatie.
- **Social** (maatschappij) gaat over de impact van een organisatie op mensen: van eerlijke inkoop tot gender equality en mensenrechten.
- **Governance** (goed beleid) gaat over de managementstructuur van een organisatie. Is er onafhankelijk toezicht, hoe zijn de bonussen geregeld en hoe divers zijn het bestuur en personeelsbestand van een bedrijf?



Figuur 3. ESG is meer dan duurzaamheid

Het is duidelijk dat deze onderwerpen zich moeilijk laten kwantificeren. Want hoe meet je een veilige werkvloer, wat is een verantwoord salarishuis en hoe ziet een cyclische waardeketen eruit? Zelfs een afgebakende factor als CO2-uitstoot levert vragen op: meet je bijvoorbeeld alleen de uitstoot van de organisatie zelf, of ook die van leveranciers en eindgebruikers?

Wie ESG-indicatoren wil kwantificeren, loopt tegen drie problemen aan:

1. In veel gevallen ontbreekt de data. Zo geven de meeste bedrijven niet of nauwelijks informatie vrij over hun energieconsumptie of salarishuis.
2. Er is weinig standaardisatie: voor boekhouding zijn er wereldwijd twee standaarden (GAAP en IFRS), maar voor ESG zijn er wel honderden.
3. De data die wél beschikbaar is, is vaak ongestructureerd of moeilijk te analyseren. Denk aan rapporten van 300 pagina's, waar een paar alinea's over de ESG-impact in 'verstopt' zitten. Of gegevens over het energieverbruik van verschillende ondernemingen die op één hoop worden opgeslagen in grote databases.

### 2.1.2 Big data in de ESG-praktijk

Omdat ESG-indicatoren moeilijk in getallen zijn te vatten, wordt er veelal teruggevallen op kwalitatieve methoden in de zoektocht naar relevante inzichten. De Nederlandse vermogensbeheerder Robeco was een van de eerste partijen die voor deze aanpak koos door vragenlijsten over uitstoot, corporate governance en hr te ontwikkelen voor het management van grote bedrijven. Tegenwoordig zijn er zowel start-ups als gevestigde organisaties die ESG-indicatoren meetbaar maken met behulp van big data.

Zo gebruikt de start-up [Sustainalytics](#) een combinatie van satellietbeelden en temperatuurmetingen om de impact van bedrijven op de zeespiegelstijging in kaart te brengen, of de totstandkoming van droogtes en hittegolven. Satellietbeelden zijn een voorbeeld van nieuwe big data-bronnen die ESG-monitoring mogelijk maken.

Start-up [TruValue Labs](#) gebruikt big data en AI om kwalitatieve gegevens over de ESG-impact van bedrijven datagedreven/meetbaar te maken. Ze verzamelen informatie over bedrijven uit een breed scala aan bronnen in meerdere talen, waaronder nieuwsartikelen, academische papers, juridische publicaties en social media. Deze interpreteren ze met behulp van Natural Language Processing (NLP) om te bepalen of er een positief of negatief sentiment is in de verschillende bronnen over het bedrijf met betrekking tot onderwerpen als organisatie-ethiek of luchtkwaliteit.

Daarnaast zijn er onderzoeksgroepen als [GreenWatch](#), die AI gebruiken om informatie over het ESG-beleid van bedrijven te verzamelen en te analyseren. Ze gebruiken zowel tekstbronnen als de waarnemingen van sensoren in een fabriek in hun dataset. Met netwerkanalyses proberen ze diepere verbanden in de productieketen van bedrijven in kaart te brengen: hoe duurzaam zijn de leveranciers van hun leverancier eigenlijk?

Dat zo veel organisaties zich op nauwkeurige(re) ESG-data richten, is niet verwonderlijk: duurzaamheid is een belangrijk thema, en bovendien is gebleken dat ESG-beleid en kwalitatieve metingen niet altijd even betrouwbaar zijn. Organisaties doen zich nog weleens duurzamer voor dan ze zijn; 'greenwashing' noemen we dat. Big data kan helpen om de daadwerkelijke milieu-impact te onthullen.

### 2.1.3 Impactmeting voor goede doelen

De inzet van big data op het gebied van ESG helpt de vertrouwensband tussen bedrijven en investeerders onderbouwen en versterken. Maatschappelijke onderwerpen die voorheen moeilijk in cijfers uit te drukken waren, worden nu kwantificeerbaar dankzij nieuwe vormen van data: denk aan satellietbeelden om de uitstoot van specifieke fabrieken te meten.

Ook goede doelen willen weten – en dus meten – wat het maatschappelijke effect van hun inspanningen is. Ze leren hoe groot hun impact nu is, en hoe ze die kunnen vergroten. Bovendien kunnen ze 'm aan donateurs laten zien, om aan te tonen dat ze hun geld goed besteden. De technieken die nu worden gebruikt in de financiële

sector om ESG-indicatoren te meten, zijn daarmee ook relevant voor goede doelen. Voorbeelden hiervan zien we in hoofdstuk 3.2, onder meer bij het Global Emancipation Network en het Jane Goodall Institute.

## 2.2 Big data in de zorg: toegankelijk maatwerk

Net als goede doelen, dient de zorgsector een maatschappelijk belang. Dit brengt grote big data-uitdagingen met zich mee. Innovaties zijn belangrijk om de zorg toegankelijk en op niveau te houden, maar tegelijkertijd moet alle informatie verantwoord worden georganiseerd en gebruikt. Al helemaal omdat het om zeer persoonlijke gegevens gaat.

Door dit speelveld is er in de zorg specifiek aandacht voor **verantwoord data delen** en **transparante AI**. Hoe ze daar met deze uitdagingen omgaan, levert relevante inzichten op voor de goededoelensector.

### 2.2.1 Big data-uitdagingen in de zorg

De zorgsector is zeer data-intensief. Van patiëntendossiers tot resultaten van klinische onderzoeken: alles wordt geregistreerd en in databases opgeslagen. Met al deze gegevens is een groot aantal big data-toepassingen mogelijk. Van ondersteuning bij operationele administratie- en planningsprocessen tot beeldanalyses van röntgenfoto's en behandeladvies voor artsen.

Om big data-uitdagingen binnen de sector in kaart te brengen, gaan we dieper in op een specifieke toepassing: **gepersonaliseerde zorg**. Hier is sprake van wanneer big data en AI de artsen op basis van unieke patiëntkenmerken ondersteunen in hun werk.

In vaktermen noemen we dit *clinical support systems*, omdat big data wordt ingezet om artsen te helpen:

- bij diagnoses;
- het ziekteverloop te voorspellen;
- beslissingen over de behandeling te nemen.

Werden deze analyses voorheen op grote, generieke groepen toegepast, nu maken big data en AI individuele analyses mogelijk. Gepersonaliseerde zorg dus.

De potentie is groot, maar deze innovatie stelt de sector wel voor twee grote uitdagingen. Allereerst vereist gepersonaliseerde zorg **veel data**, wat betekent dat deze informatie moet worden samengebracht – en dus gedeeld. Zoals in hoofdstuk 1 beschreven, zitten er veel haken en ogen aan persoonsgegevens delen. Ook spelen zaken als privacy en toestemming van deelnemers een belangrijke rol. Bovendien zien organisaties data als een asset: waarom zou een bedrijf deze met anderen delen?

Ten tweede speelt bij gepersonaliseerde zorg ook **het zogenaamde black box-probleem**: het is niet altijd duidelijk hoe een algoritme tot een conclusie, voorspelling of berekening komt – terwijl dat juist in de zorg essentieel is. Zonder kan een arts niet op de AI varen. Naast vertrouwd en veilig data delen, is transparantie een belangrijke voorwaarde voor gepersonaliseerde zorg.

### 2.2.2 Federated learning voor veilige en vertrouwde datadeling

Voor big data-toepassingen in de zorg is dus veel data nodig, die ook nog eens erg verspreid is. Organisaties zullen dus moeten samenwerken om hier optimaal gebruik van te maken.

Een traditionele samenwerking op het gebied van big data ziet er zo uit:

- Alle stakeholders maken afspraken met elkaar: personen van wie data wordt gebruikt én organisaties die deze verwerken;
- De data komt samen op een centrale plek (database);
- Op deze database vindt de analyse plaats.

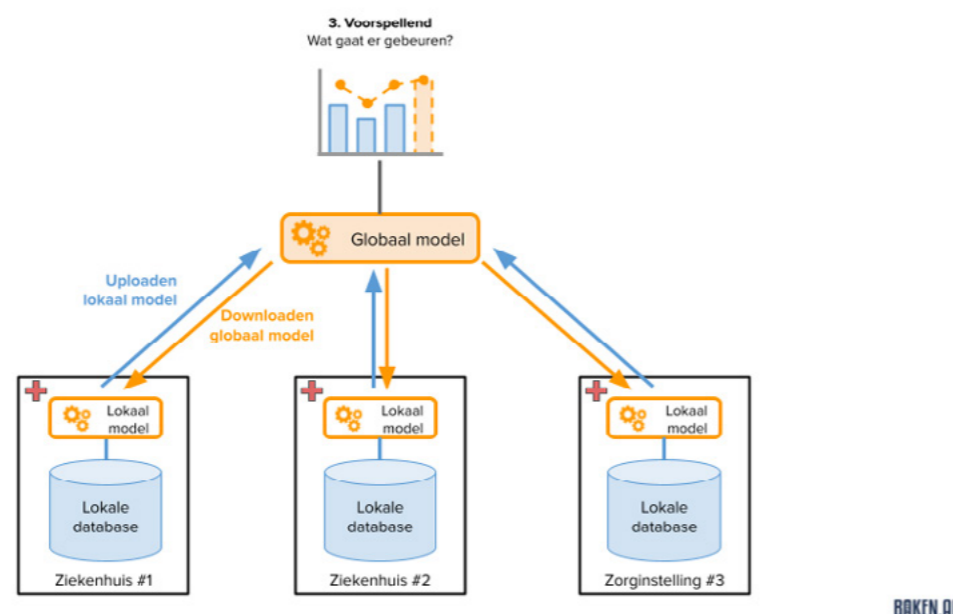
In dit proces is het belangrijk dat iedereen het met elkaar eens is én dat dataoverdracht veilig is. Voor veel sectoren een flinke uitdaging, maar zeker voor de zorg: daar is de data niet alleen erg gevoelig, maar ook verdeeld over allerlei grote en kleine zorginstellingen.

Daarom is er met name in de zorg groeiende aandacht voor **federated learning**: een privacyvriendelijke vorm van machine learning die bovengenoemde uitdagingen omzeilt door 'het model naar de data te brengen, in plaats van de data naar het model'. Deze techniek geldt als een van de relevantste voor de toekomst van gepersonaliseerde zorg.

Figuur 4 schetst hoe federated learning in de praktijk werkt.

- Data blijft altijd op locatie bij de ziekenhuizen.
- Kopieën van het globale big data-model worden met de ziekenhuizen gedeeld.
- Deze worden aangescherpt met behulp van lokale data en daarna teruggestuurd.
- De lokale modellen worden tot één nieuw globaal model samengevoegd; dit noemen we aggregeren. Bij peer-to-peer federated learning vindt deze aggregatie lokaal plaats.
- Elk ziekenhuis kan het nieuwe globale model downloaden en zo gebruikmaken van de gezamenlijke kennis over gepersonaliseerde zorg, zonder dat de data zelf het pand verlaat.

Zo stelt federated learning ook organisaties zonder grote database in staat om met big data aan de slag te gaan.



**Figuur 4.** Federated Learning: privacyproof samenwerken met big data

### 2.2.3 Explainable AI: artsen helpen bij code zwart door COVID-19

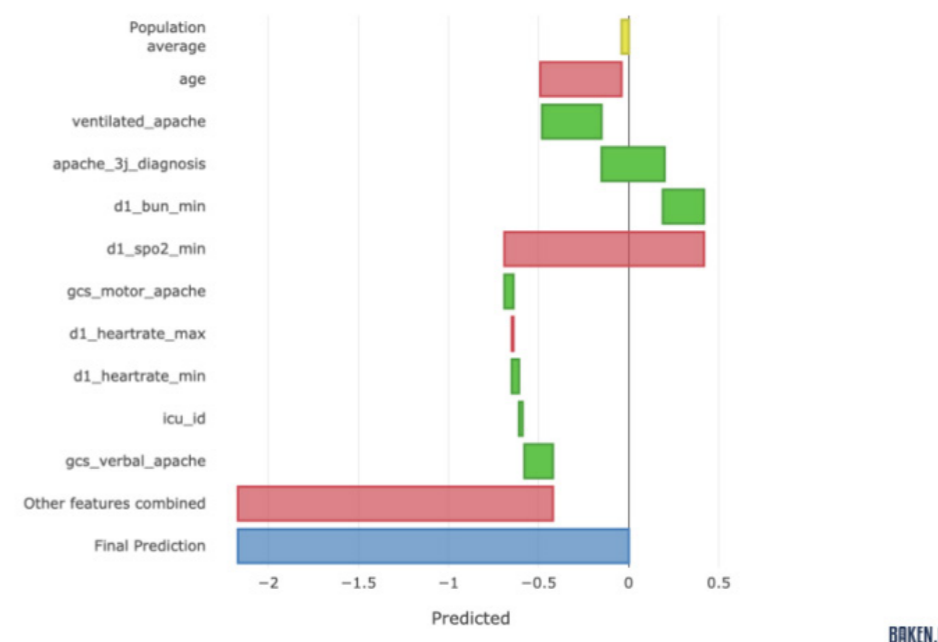
De meeste AI die we dagelijks tegenkomen, komt uit een black box gerold: denk aan de voorgestelde video's van YouTube of de advertenties op Instagram. Deze AI geeft over het algemeen goede voorspellingen, maar weinig uitleg. Hierdoor is hij maar beperkt bruikbaar voor bijvoorbeeld de zorgsector. Want als het om mensenlevens gaat, willen we weten hoe een voorspelling tot stand komt. Deze uitleg mag niet verhuld zijn in een complex of ontoegankelijk model.

Daarom is **explainable AI** in opkomst. Deze technieken openen de black box, zodat AI transparant en begrijpelijk wordt. Explainable AI legt verantwoording af aan de mens: welke beslissing maakt het systeem, en waarom? Zo zou

YouTube laten zien welke combinatie van factoren tot een bepaalde video-aanbeveling leidt. Voor gepersonaliseerde zorg houdt dit in dat je complexe voorspellingen voor individuele patiënten kunt onderbouwen. Om de potentie hiervan te illustreren, gaan we terug naar de eerste maanden van de coronapandemie.

Plotseling werd 'code zwart' een realistisch scenario voor Nederlandse ziekenhuizen: als we straks een tekort aan intensive care-bedden hebben, welke patiënten gaan we dan nog wel opnemen en welke niet? En wie beoordeelt dit? Een moeilijke beslissing, die verschillende ziekenhuizen – waaronder het Amsterdam UMC en Maastricht UMC – wilden baseren op meer kenmerken dan alleen de leeftijd van patiënten. Zo ontstond de vraag naar **een model** om de overlevingskans van COVID-19-patiënten te voorspellen.

In figuur 5 is een voorbeeld te zien van zo'n voorspelling, gemaakt met behulp van explainable AI: positieve (groen) en negatieve (rood) variabelen dragen bij aan een uiteindelijke prognose (blauw). Zo beïnvloedt de leeftijd de kansen negatief, maar is het weer voordelig dat deze patiënt aan de beademing ligt (ventilated\_apache). Via explainable AI wordt kunstmatige intelligentie transparant en bruikbaar voor de medische praktijk – al is deze techniek gelukkig nog niet nodig geweest voor code zwart.



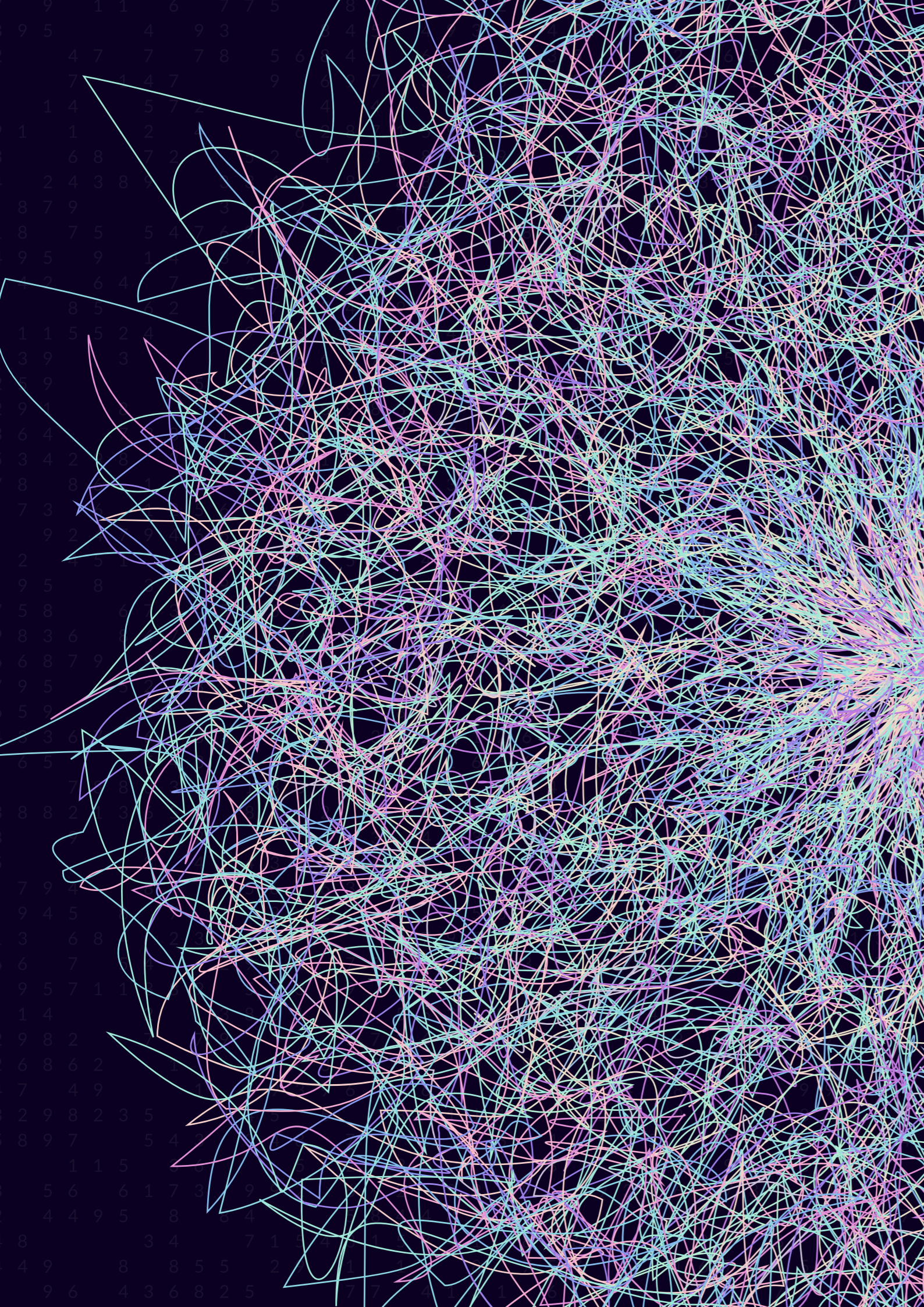
**Figuur 5.** Een voorbeeld van een explainable AI-voorspelling over de overlevingskans van een patiënt op de intensive care

### 2.2.4 Lessen uit de zorg voor goede doelen

De oplossingen die in de zorg voor big data-uitdagingen zijn gevonden, zijn waardevol voor de goedbedoelensector. Vanwege de maatschappelijke doelstelling en verantwoording aan donateurs zijn **verantwoorde datadeling** en **transparante AI** ook voor goede doelen relevant.

Federated Learning biedt goede doelen van elk formaat de mogelijkheid om data te bundelen en AI-modellen te gebruiken. Hiermee kan de techniek een interessante *enabler* zijn, bijvoorbeeld voor het vraagstuk van donateursuitstroom: factoren die een rol spelen bij uitschrijven zijn beter te achterhalen als meer data wordt samengebracht. In hoofdstuk 3 zetten we alle big data-kansen voor goede doelen op een rij.

Tegelijkertijd is het noodzakelijk dat goede doelen hun AI op een transparante en verantwoorde manier ontwikkelen. Door de black box te vermijden, kunnen ze transparante en uitlegbare AI ontwikkelen, voorsorteren op mogelijke AI-wetgeving (zie 1.4.2) en een proactief signaal afgeven aan de steeds AI-bewustere consument.



### 3. BIG DATA VOOR GOEDE DOELEN

In dit hoofdstuk kijken we naar innovaties van goede doelen die big data gebruiken voor doelbesteding en de werving van steun en donateurs.

Zo zien we hoe donateur lifetime value-optimalisatie betere analyses van donateurwerving en -behoud oplevert. Ook staan we stil bij de groeiende toegankelijkheid van digitaal doneren en nieuwe manieren om de boodschap van een fonds *immersive* te maken, zodat je donateurs als het ware 'onderdompelt' in een meeslepende ervaring, bijvoorbeeld met chatbots en virtual reality. Bovendien kijken we hoe je donateurbinding kunt verbeteren met transparantie via datagedreven storytelling en blockchain. Tot slot zien we hoe goede doelen hun doelbesteding verbeteren via satellietdata, publieke data en nieuwe AI-modellen.

#### 3.1 Big data voor fondsenwerving en marketing

We kunnen big data-toepassingen voor fondsenwerving en marketing onderverdelen in vier domeinen:

1. Verbeterde analyses om donateurs te werven en behouden
2. Datagedreven oplossingen om digitaal doneren makkelijker te maken
3. Meeslepende storytelling om de donateurbinding te vergroten
4. Transparante verantwoording aan donateurs en maatschappij

##### 3.1.1 Donateur lifetime value-optimalisatie

Datagedreven werving en behoud van donateurs noemen we **donateur lifetime value-optimalisatie (DLV-optimalisatie)**. Dit kan bijvoorbeeld door big data te gebruiken om potentiële donateurs beter te targeten en uitstroom te verminderen (zie ook de Roadmap in hoofdstuk 4).

Zo'n datagedreven aanpak berust op een combinatie van (veel) data en nieuwe analysetechnieken, zoals machine learning en AI. Door bijvoorbeeld kenmerken van potentiële en bestaande donateurs in een database samen te brengen, kan een AI-toepassing heel nauwkeurig bepalen welke personen waarschijnlijk willen doneren. Dit stelt goede doelen in staat om gericht te werven en (op termijn) de kosten van werving te verlagen. Een ander voorbeeld: dankzij marketingresponsdata (o.a. e-mailgegevens en website clicks) is het mogelijk om campagnes per doelgroep te specificeren, en zo de resultaten te verbeteren.

Tools voor DLV-optimalisatie worden als Software-as-a-Service (SaaS) aangeboden. Dit zijn gestandaardiseerde AI-diensten, geschikt voor organisaties van elk formaat. De grootste aanbieders van SaaS zijn gevestigde partijen als [Microsoft](#) en [Salesforce](#), die deze dienst koppelen aan bestaande software, zoals hun CRM-pakketten.

Naast de grote SaaS-aanbieders, zijn er ook partijen die specifieke AI-producten en -diensten voor goede doelen maken. Deze bieden een volledig alternatief op de softwarepakketten van grote aanbieders ([Gravyty](#), [Neon One](#), [GiveLively](#)). Of ze koppelen hun eigen AI-services aan software die al door de organisatie wordt gebruikt ([Boodle](#), [Quilt](#), [CivisAnalytics](#)).

Tot slot zijn er ook partijen en initiatieven die AI-oplossingen op maat maken voor goede doelen. Deze lopen uiteen van AI-consultancystart-ups tot hackathons en universiteitsprojecten: denk aan de hackathons van [DataKind](#) en – dichterbij huis – de [Analytics Academy](#) van VSAE, de studievereniging voor actuaariaat- en econometriestudenten van de Universiteit van Amsterdam.

##### 3.1.2 Frictieloos digitaal doneren met embedded finance

Big data wordt ook gebruikt voor fondsenwerving door digitaal doneren via datagedreven oplossingen makkelijker te maken. Deze focus op frictieloos digi-doneren plaatsen we in het licht van de **embedded finance**-trend.

Financiële diensten als bankieren, verzekeren of lenen zijn steeds vaker geïntegreerd in niet-financiële ecosystemen, platforms en producten. Deze ontwikkeling wordt 'embedded finance' genoemd: denk aan je laptop verzekeren via Coolblue, een bankrekening openen als je een webshop opricht via Shopify en krediet kopen via de achterafbetalen-functie van Klarna als je een bank bij Wehkamp bestelt.

Embedded finance-aanbieders maken hun financiële diensten toegankelijk en makkelijk te integreren in andere platforms. Ruim 15.000 webshops in Nederland werken al met het Zweedse Klarna. Voor goede doelen houdt embedded finance in dat ze op veel digitale plaatsen aanwezig zijn en voor frictieloze integratie zorgen. Ook in onze sector is de techniek in opkomst, blijkt uit een aantal innovaties.

De socialmediaplatforms YouTube, Facebook en Instagram zijn bekende, gevestigde plaatsen voor donatiecampagnes. Met het groeiende belang van deze platforms, zijn ook de tools die goede doelen hier kunnen toepassen in opkomst. Voorbeelden zijn [Givepanel](#), [NationBuilder](#) en [Classy](#). De software van deze bedrijven helpt goede doelen hun campagnes op één of meer social media te managen.

In Nederland is inmiddels ongeveer 70% van alle betalingen contactloos. Start-ups als [GoodBox](#) bieden betaalapparatuur voor goede doelen in de cashloze wereld.

Andere initiatieven zijn erop gericht donaties onderdeel van dagelijkse betalingen te maken. In Nederland kennen we [Peaks](#): deze beleggingsapp rondt aankoopbedragen naar boven af, het verschil wordt aan de beleggingsrekening toegevoegd. Zulke constructies bieden Britse bedrijven als [Pledjar](#) en [Roundups](#) ook aan voor goede doelen.

### 3.1.3 Donateurs binden met immersive storytelling

Om donateurs van dichtbij te laten zien waar hun geld naartoe gaat, worden steeds meer immersive – in het Nederlands: meeslepende – methoden ontwikkeld met gebruik van big data- en AI-technieken. Deze stellen goede doelen in staat om donateurs meer bij hun werk te betrekken en zo steun te werven.

Mencap laat zien hoe je mensen bijvoorbeeld kunt betrekken. Dit goede doel voor mensen met een leerstoornis gebruikt een chatbot: [Understand Me](#). Hier kunnen donateurs ‘in gesprek’ gaan met een meisje dat een leerstoornis heeft. Hierdoor komen ze op een interactieve manier meer over de aandoening te weten.

Een interactieve dialoog kan zeker helpen, maar zoals het gezegde luidt: beelden zeggen meer dan woorden – helemaal als ze bewegen. Daarom zetten goede doelen steeds vaker video’s in om donateurs dicht bij hun werk te brengen. Zo ook [Smile Train](#), een non-profitorganisatie die zich inzet voor gratis operaties voor kinderen met een hazenlip. Zij gebruiken video’s als [The Story of Nisha](#) om door de ogen van een kind te laten zien wat de impact van zo’n operatie is.

Zoals [Games for Good](#) laat zien, zijn videogames ook een uitstekend medium om het verhaal van een non-profit te vertellen. Games for Good is een platform die in samenwerking met bijvoorbeeld NGO’s games maakt voor ‘the Good’. Bekende voorbeelden van zulke games zijn [This War of Mine](#) en [Bury me, My Love](#). De eerste duikt in het leven van oorlogsslachtoffers, de tweede gaat over een Syrische vluchteling die naar Europa emigreert.

Vanaf video’s en games is het voor goede doelen een kleine stap naar augmented reality (AR) en virtual reality (VR). Bij AR worden visuele digitale elementen over de echte wereld heen gelegd: denk aan Snapchat-filters en Pokémon GO. VR is een volledig digitale wereld, waarin gebruikers met een speciale bril kunnen rondlopen en interacteren. Deze technologie wordt voornamelijk vooral voor videogames gebruikt, maar vindt steeds meer z’n weg naar andere toepassingen, zoals [behandeltherapieën](#) en [militaire trainingen](#). Voor goede doelen is VR een kans om hun verhaal op een innovatieve manier te vertellen. Ofwel: show, don’t tell 2.0.

De National Autistic Society in het Verenigd Koninkrijk gebruikt VR bijvoorbeeld om mensen te laten ervaren hoe het is om een [autistisch kind](#) te zijn. Anderhalve minuut lang word je overladen met prikkels, om vervolgens weer langzaam tot rust te komen. Een andere innovatieve VR-toepassing is die van [The Life Garden](#) van Cancer Research UK. Dit is een virtuele tuin, waarin elke donateur die het fonds iets heeft nagelaten een eigen bloem heeft. De organisatie eert zo ruim 100.000 mensen. The Life Garden blijft oneindig groeien en nabestaanden kunnen er altijd naartoe om de digitale bloem van hun geliefde in alle

rust op te zoeken. Ze hoeven er alleen een VR-bril voor op te zetten.

Brengen we VR, videogames en de kansen van embedded finance (zie 3.1.2) samen, dan zien we een toekomst voor nieuwe, immersive donatie-ervaringen. Donateurs belevend zich in een interactieve VR-omgeving met digitale ervaringen of spelelementen, gebaseerd op de missie van het goede doel. Hier kunnen ze direct doneren dankzij de integratie van een betalingssysteem. Een eerste voorbeeld hiervan is [Krikey](#): een AR-app waarin spelers een babygorilla moeten beschermen en verzorgen, à la [Tamagotchi](#).



Figuur 6. Een screenshot uit AR-donatiegame Krikey

Vanuit de videogame kunnen ze doneren aan The Ellen DeGeneres Wildlife Fund.

### 3.1.4 Datagedreven transparantie

Goede doelen doen een beroep op steun van het publiek. Ze moeten verantwoording afleggen aan hun donateurs en de maatschappij over hun werk en resultaten. Big data stelt goede doelen in staat om hier transparant en ‘in cijfers’ over te communiceren: met datagedreven storytelling.

Door data in hun communicatie te gebruiken, maken goede doelen hun impact en missie transparant met cijfers. Zo maakt [charity: water](#) het mogelijk om projecten te volgen via een Google Maps-achtige kaart op de website. De [Bill & Melinda Gates Foundation](#) gebruikt interactieve grafieken om de gendergap te illustreren die door COVID-19 op de arbeidsmarkt is ontstaan. Zo geven ze inzicht in loon- en werkgelegenheidsverschillen tussen mannen en vrouwen, die niet te verklaren zijn door factoren als leeftijd en opleidingsniveau.

Een koploper op het gebied van datagedreven transparantie is [Our World in Data](#) van het Global Change Data Lab. Deze non-profit maakt data en onderzoek toegankelijk en kaart zo de grootste wereldproblemen aan: van inkomensongelijkheid tot luchtvervuiling. Dankzij hun heldere uitleg en inzichtelijke grafieken zijn ze een leidende bron van data. Zo hebben ze een uitgebreide tracker voor de [Sustainable Development Goals](#) van de Verenigde Naties.

Data is dus een tool die transparantie kan bevorderen. Maar zoals we in hoofdstuk 1 beschreven, levert datagebruik ook steeds meer ethische en juridische uitdagingen op. Als goede doelen de ethische standaarden uit 1.4.1 implementeren en bijvoorbeeld bij gebruik van explainable AI inzetten, kunnen ze een signaal afgeven richting donateurs over hun verantwoorde databeleid. Dit wordt ook wel **privacy by design**

genoemd: technisch en organisatorisch databeleid dat verantwoord gebruik van data en algoritmes afdwingt.

De toekomst van transparante communicatie en privacy by design voor non-profits ligt in de blockchain-hoek. Blockchain is de technologie achter Bitcoin. Korte samenvatting: het is een systeem dat data opslaat in een keten van blokken, zonder dat iemand zo'n blok beheert en controleert. Zo kun je onderling en zonder tussenpersoon transacties uitvoeren en bijhouden.

Vertalen we dit principe naar de goedbedoelensector, dan krijg je bijvoorbeeld dat de uitkering van een donatie is gekoppeld aan een bepaald resultaat. De betaling, uitkering en afspraken over het gewenste resultaat zijn vastgelegd in de blockchain. Het systeem kan niet anders dan de afspraken opvolgen, waardoor er geen besliser of organisatie meer nodig is – de ultieme vorm van transparantie. **Alice** biedt zulke blockchain-applicaties aan goede doelen aan, zoals een geldinzameling voor dakloze jongeren. **Endaoment** gaat een stap verder: deze organisatie stelt mensen in staat hun eigen fonds op te zetten in de blockchain, waarbij deelnemers zelf bepalen hoe het geld wordt uitgekeerd.

De blockchain is mogelijk relevanter voor goede doelen dan voor de gemiddelde bank. Vanwege ontwikkelingen van de technologie, maar ook door het belang bij efficiëntie en transparantie van donateurs. De blockchain stelt goede doelen in staat hun te laten zien hoe hun donatie wordt besteed, zonder veel operationele kosten.

### 3.2 Big data voor doelbesteding

Ten slotte gaan we dieper in op de rol van big data en AI bij de doelbesteding van non-profitorganisaties. Kunnen ze bijvoorbeeld bijdragen aan fraudebestrijding? En hoe kan big data chimpansees helpen? Aan de hand van voorbeelden schetsen we hoe goede doelen big data en AI gebruiken.

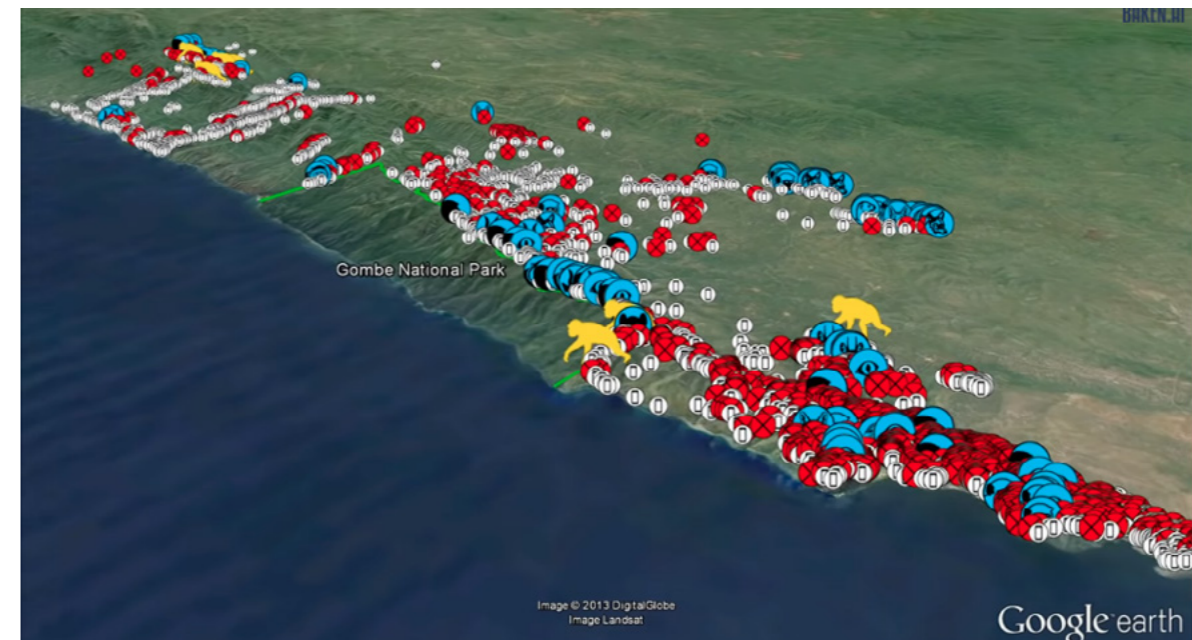
Het **Jane Goodall Institute** richt zich op de conservatie van chimpansees. Het instituut had veel informatie over de dieren, maar minder over hun leefomgeving en hoe deze verandert – bijvoorbeeld door ontbossing. Door hun eigen data aan die van de publieke **Google Earth Engine** te koppelen, kunnen ze de chimpansees beter helpen en is behoud van hun natuurlijke habitat een nieuw doel dat de organisatie nastreeft én monitort.

Een ander mooi voorbeeld van een non-profit die data bij doelbesteding inzet, is het Britse **Crisis Text Line**. Zij bieden 24/7 gratis mentale zorg en crisishulp via een sms-lijn. Deze organisatie werkt met getrainde vrijwilligers die AI gebruiken om de binnenkomende sms'jes te prioriteren. Hun model is getraind om, op basis van het eerste bericht, direct mogelijke zelfmoordsituaties te herkennen. Hierdoor zijn ze nu in staat om 94% van de sms'jes met een hoog risico binnen 5 minuten te beantwoorden.

In 2.1 hadden we het al kort over het gebruik van satellietdata in de financiële sector voor ESG-doeleinden. Ook in de goedbedoelensector zien we dit terug, onder andere bij **Médecins Sans Frontières** (MSF, de Franse tak van Artsen zonder Grenzen). Met behulp van het geo-analyselab van de Universiteit van Salzburg vertalen ze deze data in kaarten op maat, populatieschattingen of analyses bij overstromingen. Handige hulpmiddelen voor de uitdagende gebieden waar MSF werkt, helemaal omdat nauwkeurige informatie en omgevingskaarten daar vaak niet voorhanden zijn.

Ook het **Global Emancipation Network** gebruikt big data, in hun strijd tegen mensensmokkel. In samenwerking met Microsoft ontwikkelde het netwerk **Minerva**: een platform dat data verzamelt door onder meer socialmediaposts, deepwebadvertenties, Google Maps-gegevens en afbeeldingen van slachtoffers te bundelen.

De non-profit **Conservation International** ontwikkelde AI-tool **Colandr** om wetenschappelijk onderzoek naar conversatie te analyseren. Deze neemt constant nieuwe, relevante wetenschappelijke papers door en vat ze samen. Vanuit de opensource-gedachte is deze tool ook beschikbaar voor andere organisaties. De non-profit **Zidisha** – gericht op microfinanciering – maakt gebruik van big data om fraude bij aanvragen van microkrediet tegen te gaan. Hierdoor nam het misbruik met een derde af en groeide het totaal aan



**Figuur 7.** De data van het Jane Goodall Institute gecombineerd met Google Earth

uitgeleend geld met 10%.

#### 3.2.1 Big data-filantropie

In de toekomst zullen particulieren of bedrijven vaker niet alleen geld, maar ook data doneren voor de doelbesteding. Hiermee kunnen goede doelen analyses uitvoeren die ze anders niet hadden kunnen doen, bijvoorbeeld vanwege kosten of privacy.

Er zijn een paar voorbeelden van datadonaties die particulieren kunnen doen voor doelbesteding. Zo maakte **Fitbit** het mogelijk om data te doneren voor onderzoeken naar COVID-19 en kon je met **Donata** zelf bepalen welke data je aan welke organisatie wilde geven.

Daarnaast doneren grote bedrijven hun big data steeds vaker aan goede doelen om hun doelbesteding te ondersteunen. Dit noemen we big data-filantropie. Een goed voorbeeld hiervan is het **Harvard-onderzoek** naar de verspreiding van malaria. Safaricom (Vodafone Kenya) doneerde hiervoor de geanonimiseerde telefoongegevens van miljoenen mensen in Kenia. Deze data werd zo voor het eerst gekoppeld aan lokale infectiegegevens, om inzicht te krijgen in de mobiliteit van mensen en de bijdrage hiervan aan de verspreiding van malaria.

Een ander voorbeeld van big data-filantropie is de donatie van satellietbedrijf **DMCii**, dat data doneerde aan de **National Institute for Plant Protection** van Algerije. Met deze gegevens kan het instituut effectief optreden tegen sprinkhanenplagen, waar het land vaak last van heeft. Door satellietbeelden van de vegetatie met weersvoorspellingen te combineren, werd het mogelijk om sprinkhanenzwermen te voorspellen. Dankzij deze prognoses kunnen gericht pesticiden worden ingezet.

# Big data is pas écht nuttig als je er waarde uit kunt halen.



## 4. AAN DE SLAG MET BIG DATA

De Goede Doelen Big Data Roadmap beschrijft de onderzoeksvragen die goede doelen tegenkomen bij de belangrijkste big data-analyses. Om het voor fondsen van elke omvang toegankelijk te maken, is de Roadmap

Big data-niveaus → en -analyses ↓	Beschrijvend Data verzamelen en uitzetten	Diagnostisch Doelvariabele verklaren op basis van kenmerken	Voorspellend Doelvariabele voorspellen op basis van kenmerken	
<b>Donateur lifetime value</b>	Optimalisatie donatiebedrag	Wat zijn de kenmerken van donateurs per donatiebedrag?	Welke kenmerken verklaren het gemiddelde donatiebedrag? Wat zijn relevante doelgroepen?	Wat is het optimale donatiebedrag per doelgroep? Hoeveel gaat een donateur op basis van kenmerken doneren?
	Optimalisatie donatiefrequentie	Wat zijn de kenmerken van donateurs per donatiefrequentie?	Welke donateur- of campagnekenmerken verklaren de donatiefrequentie?	Wat is de optimale donatiefrequentie per doelgroep?
	Optimalisatie instroom	Hoeveel instroom is er per doelgroep?	Wat zijn de 'unieke' kenmerken van ingestroomde donateurs: specifieke campagne, doelgroep?	Wat is de kans dat een potentiële donateur daadwerkelijk gaat doneren?
	Optimalisatie uitstroom	Hoeveel uitstroom is er per doelgroep?	Wat zijn de 'unieke' kenmerken van uitgestroomde donateurs: specifieke campagne, doelgroep, regio, enz.?	Wat is de kans dat een huidige donateur stopt met doneren?
<b>Campagne-analyse</b>	Impact campagne	Wat zijn relevante datapunten van een campagne? Hoe meet je het resultaat?	Welke verklaren het resultaat? Hoe verhouden deze zich tot andere campagnes? En is er verschil in effectiviteit tussen doelgroepen?	Wat is het verwachte resultaat? Wat is de beste campagne per doelgroep?
<b>BAKEN.AI</b>				

Figuur 8. Goede Doelen Big Data Roadmap



© 2021

**Goede Doelen Nederland**

Websites

[www.goededoelennederland.nl](http://www.goededoelennederland.nl)

[www.goededoelen.nl](http://www.goededoelen.nl)

Postadres

Postbus 94488

1090 GL Amsterdam