

Data zijn geen productiefactor, maar wel productiviteitsverhogend

Sinds de jaren tachtig groeit de mondiale hoeveelheid data sterk. Ook data-intensieve bedrijven als Google en Amazon zijn steeds belangrijker geworden in het economisch verkeer. Hoe moeten we opkomst van data kwalificeren?

JOOST WITTEMAN

Onderzoeker bij
SEO Economisch
Onderzoek

ERIK BROUWER

Clusterhoofd bij
SEO Economisch
Onderzoek en hoog-
leraar aan Tilburg
University

TOM SMITS

Onderzoeker bij
SEO Economisch
Onderzoek

De opkomst van data en aanpalende onderwerpen als datadeling staan maatschappelijk volop in de belangstelling. Maar hoewel veel opinieartikelen data als de ‘olie van de 21e eeuw’ (FD, 2016) of de ‘vierde productiefactor’ bestempen (Financial Times, 2012), is er maar weinig écht bekend over de rol die data spelen in de economie.

Om te beoordelen of een zaak als verplichte datadeling de economie daadwerkelijk zal veranderen, is het daarom belangrijk om eerst op een rijtje te zetten wat we op dit moment weten over het belang van data. Zijn data bijvoorbeeld een grondstof, of zijn ze misschien zelfs een productiefactor?

Op basis van de wetenschappelijke literatuur geven we in dit artikel een antwoord op deze vraag. Eerst laten we zien hoeveel data er eigenlijk zijn. Daarna stellen we een meer conceptuele vraag, namelijk of data eerder een productiefactor (‘een argument in de productiefunctie’) zijn of factorproductiviteitsverhogend (‘een parameter van de productiefunctie’).

HOEEVEELHEID DATA EN GROEI

Zicht krijgen op de totale hoeveelheid beschikbare data is niet eenvoudig. Anders dan de productiefactoren arbeid en kapitaal zijn data bijvoorbeeld geen onderdeel van de nationale rekeningen. Daarnaast zijn er maar weinig goede schattingen en is ook de geijkte meetmethode nog onderwerp van debat (Hilbert, 2012).

Betere schattingen komen van Hilbert en López (2011) en Hilbert (2015). Zij baseren zich op tellingen van de hoeveelheid data die men kan opslaan op de verschil-

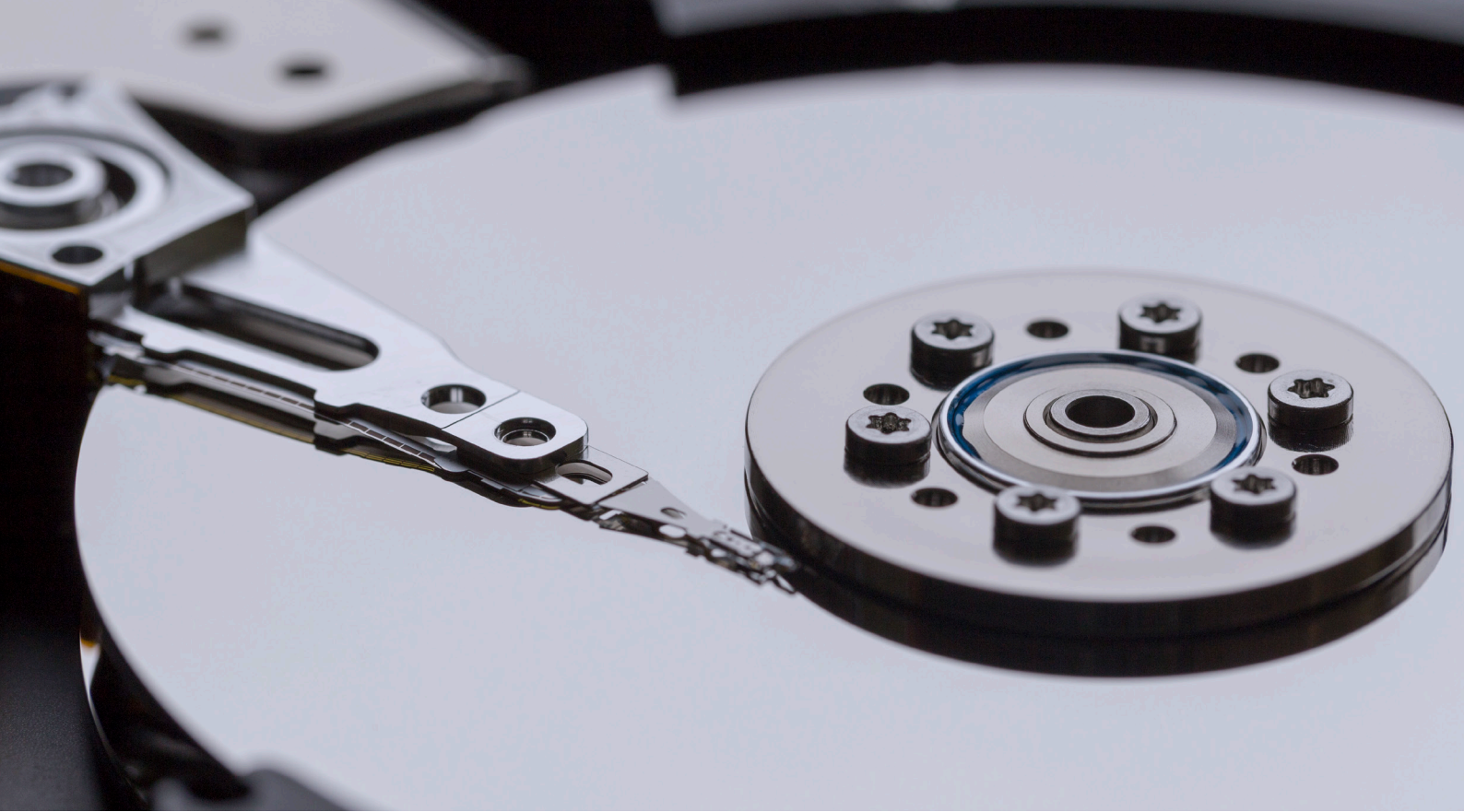
lende opslagmedia. Als er tien pagina's tekst zijn en elke pagina bevat 1 kilobyte aan data, dan heb je dus 10 kilobyte aan data. Door dergelijke tellingen te doen voor 25 verschillende opslagmedia wereldwijd, ontstaat er een tijdsreeks van de mondiale opslag van data tussen 1986 en 2014.

Deze tijdsreeks wordt zichtbaar in figuur 1, waarbij de schaal van de y-as logaritmisch is. Het eerste gelabelde veelvoud betreft een exabyte, ofwel een miljard gigabyte. Het tweede gelabelde veelvoud – de zettabyte (ZB) – is nog eens 1.000 keer groter. Bij de figuur vallen drie zaken op. Ten eerste dat er bijzonder veel data zijn: zo zijn er 62,5 miljoen iPhones met 16 gigabyte geheugen nodig om 1 exabyte aan data op te slaan. Ten tweede valt het op dat de hoeveelheid opslagcapaciteit enorm gegroeid is: van minder dan 3 exabyte in 1986 naar meer dan 4 zettabytes in 2014. Dit is een toename van bijna een factor 1.500 over minder dan dertig jaar, ofwel een jaarlijkse groei van ongeveer dertig procent. Ten derde komt de groei van de wereldwijde data-opslag vanaf ongeveer de millenniumwisseling vrijwel volledig voor rekening van ‘digitale’ opslagmedia zoals harde schijven, in plaats van voor ‘analoge’ opslagmedia zoals boeken. In opslagtermen is de ‘digitale eeuw’ hiermee vrij letterlijk rond de eeuwwisseling van start gegaan, wat samenhangt met de opkomst van het internet en de toenomen adoptie van mobiele telefoons en laptops.

Ook lijkt deze groei nog altijd niet te stikken. Zo schat IDC (2017) dat er in 2016 16 zettabytes aan data waren en dat de groei zal doorzetten tot 163 ZB in 2025. Beide schattingen zijn in lijn met de door Hilbert en López (2011) en Hilbert (2015) geobserveerde groei. Het constante karakter van de groei lijkt een wetmatigheid in de groei te suggereren, vergelijkbaar met die van de Wet van Moore wat betreft de toename van de computerrekenkracht.

VERGELIJKING MET ARBEID EN KAPITAAL

De groei van de wereldeconomie, ofwel de productiefactoren arbeid en kapitaal, steekt wat bleek af tegen de stormachtige



groei van de opslagcapaciteit. Tussen 1986 en 2014 nam het wereldwijde reële bruto binnenlands product (bbp) toe van ongeveer 37 biljoen naar ruim 100 biljoen dollar. De mondiale kapitaalvoorraad groeide in reële termen van 95 biljoen dollar naar meer dan 220 biljoen dollar over dezelfde periode. De groei van beide reeksen is hiermee, met drie tot vier procent per jaar, weliswaar gezond maar niet explosief (IMF, 2017). Hetzelfde geldt voor de mondiale potentiële beroepsbevolking die over deze periode met 1,7 procent per jaar toenam (VN, 2017).

Interessant is verder dat de groei van data niet conjunctuurafhankelijk lijkt. De groei in het wereldwijde reële bbp stakte tijdens de Grote Financiële Crisis (IMF, 2017), maar dit is niet zichtbaar bij de groei van de hoeveelheid data in figuur 1. Opvallend is daarbij de parallel met de reële wereldwijde kapitaalvoorraad, die ook weinig krimp vertoonde

tijdens de crisis (IMF, 2017). Blijkbaar is het, ondanks macro-economische tegenwind, niet 'ongewoon' voor productiefactoren om mondiaal in niveau toe te nemen.

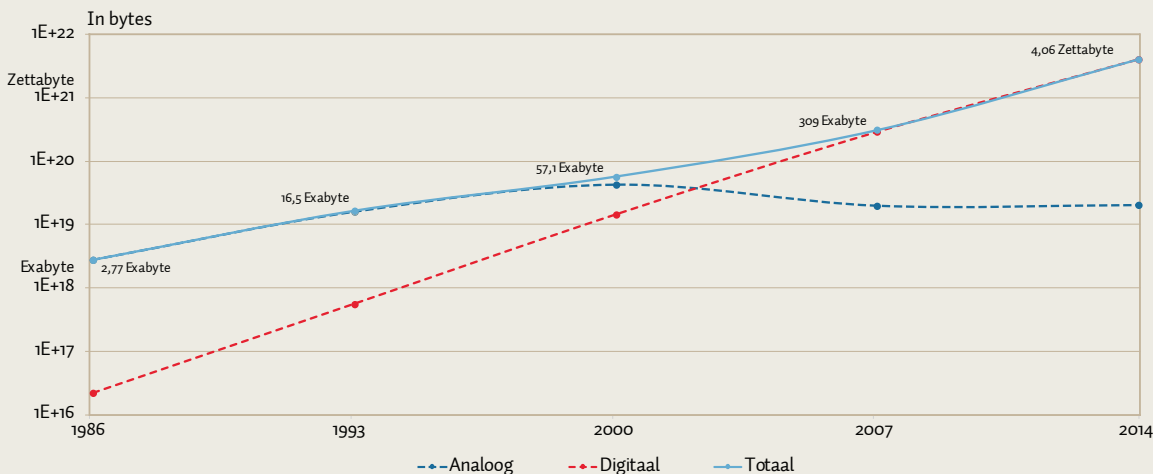
DATA IN PRODUCTIE

Het gebruik van data in de productie vereist dat deze geanalyseerd worden. Het Centraal Bureau voor de Statistiek laat zien hoeveel werknemers in bepaalde bedrijfstakken werken met computers en hoeveel bedrijven bigdata-analyses uitvoeren. We gebruiken dit als proxy voor het datagebruik van sectoren.

De ICT- en data-intensiteit is hoog in bepaalde hoogwaardige dienstensectoren. In de verzekeringsbranche werkt bijvoorbeeld 97 procent van de werknemers met computers en heeft 52 procent van de bedrijven bigdata-analyses uitgevoerd. Andere sectoren aan de bovenkant van

Schatting van de totale wereldwijde dataopslag

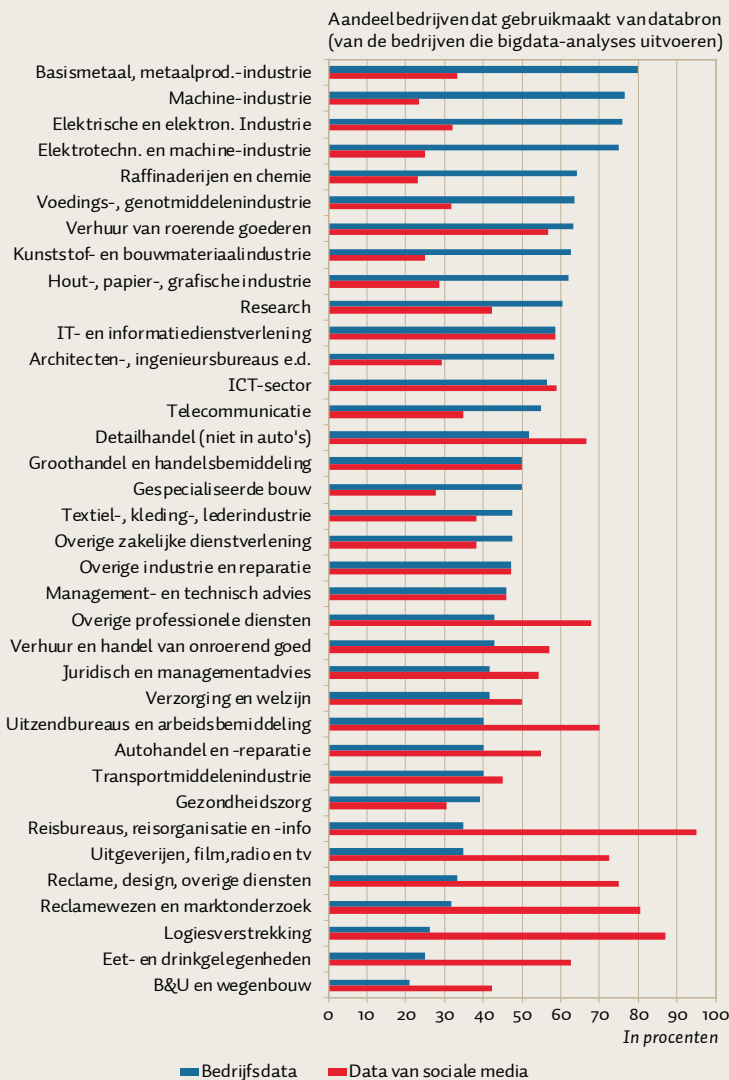
FIGUUR 1



Bron: Hilbert en López (2011); Hilbert (2015)

Verschillen in gebruik databronnen per sector

FIGUUR 2



Noot: Figuur op basis van tweecijferige SBI2008, bedrijfstakken C, N-Q.

Bron: CBS

dit spectrum zijn bijvoorbeeld het reclamewezen, de reisbranche of de telecommunicatiesector.

Beduidend minder data-intensief zijn de (maak)industrie, handel, detailhandel en de simpelere dienstverleningen. Minder dan de helft van de werknemers in eet- en drinkgelegenheden werkt met een computer, en maar acht procent van de bedrijven in deze bedrijfstak doet aan bigdata-analyse. Andere sectoren die minder data-intensief zijn, zijn bijvoorbeeld de verzorgings- en welzijnssector, de kunststof- en bouwmaterialindustrie, of de transportmiddelenindustrie.

Verdere sectorale verschillen zijn er met betrekking tot wat voor soort data gebruikt worden door de bedrijven die bigdata-analyses uitvoeren. Met name bedrijven in de (maak)industrie die aan bigdata-analyse doen, zijn geneigd gebruik te maken van data die voortkomen uit het eigen bedrijf. Sectoren op het gebied van (professionele) dienstverlening zoals uitzendbureaus, het reclamewezen en uitgeverijen maken daarentegen veel vaker gebruik van data van sociale media om analyses uit te voeren (figuur 2).

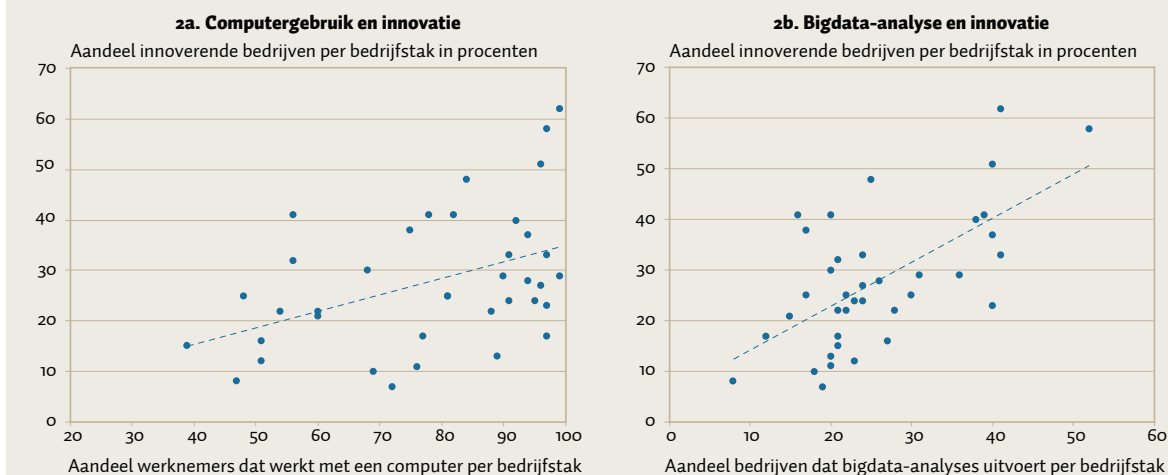
Wat deze bedrijven precies doen met hun bedrijfsdata en/of de data die verkregen zijn via sociale media, is niet (systematisch) bekend. Gelet op de beperkte ICT-intensiteit van bepaalde sectoren en het feit dat er in veel sectoren maar weinig bedrijven bigdata-analyses uitvoeren, is het echter niet direct evident dat data een productiefactor als arbeid of kapitaal zijn, zonder welke productie niet mogelijk is. In elk geval blijken sommige bedrijfstakken, gezien het aandeel werknemers dat hier gebruikmaakt van computers voor hun werk of dat bigdata-analyses uitvoert, helemaal niet zo ICT- en dus data-intensief.

DATA WERKEN PRODUCTIVITEITSVERHOGENDE

In plaats van op arbeid of kapitaal lijken data mogelijk meer op informatietechnologie (IT), waarvan bekend is dat het de productiviteit verhoogt (Syverson, 2011). De waarde van data schuilt dan meer in het vergroten van het gebruik van andere productiefactoren. Het kan in dat geval bijvoorbeeld helpen bij het oplossen van *search-and-match*-problemen op markten, of bij het nemen van

Relatie tussen ICT-gebruik en innovatie in Nederlandse sectoren

FIGUUR 3



Noot: Figuur op basis van viercijferige SBI2008, productinnovatie, bedrijfstakken C, N-Q

Bron: CBS

betere bedrijfsbeslissingen over productie.

In Nederland lijken bijvoorbeeld sectoren waarin meer medewerkers van computers gebruikmaken of meer big-data-analyses uitvoeren vaker innovatief te zijn – zowel in termen van producten als processen (figuur 3). Dit strookt met het idee dat innovatie op de lange termijn een belangrijke motor van de productiviteitsgroei is (Syverson, 2011), en tevens met de claim dat de ‘digitale economie’ momenteel mogelijk de belangrijkste bron van innovatie is, met een zeer hoge surpluswaarde (Arrieta-Ibarra et al., 2018; Brynjolfsson et al., 2018). Ook sluit dit aan bij de opkomst van *superstar firms* als Google en Amazon, die zowel hoog-productief als zeer data-intensief zijn (Autor et al., 2017). Empirisch wijzen Brynjolfsson et al. (2011) erop dat bedrijven met een ‘datagestuurde besluitvorming’ vijf tot zes procent productiever zijn dan bedrijven zonder.

CONCLUSIE

Hoewel data soms gezien worden als de ‘nieuwe olie’ of de ‘vierde productiefactor’ lijken ze op dit moment veel meer op een nieuwe ‘technologie’, die in de eerste plaats vooral productiviteitsverhogend werkt. Meer onderzoek hiernaar is zeer gewenst, ook al omdat het voorgaande duidelijk heeft gemaakt hoeveel we in feite nog niet weten over data in de economie. Het opvullen van deze leemte is belangrijk, niet alleen academisch maar vooral ook maatschappelijk. Zo ziet RAND Europe (2015) in het productiviteitseffect van data een reden tot meer datadeling. En ook mededingingsoverwegingen kunnen hiertoe aanleiding geven (Ecorys, 2017). Tegelijkertijd wijzen Rosenboom en Smits (2018) erop dat het niet (hoeven) delen van data een prikkel tot innovatie kan bevatten. Daarbij zijn er voor een zorgvuldige afweging van deze effecten feitelijk meer data over data vereist.

LITERATUUR

- Arrieta-Ibarra, I., L. Goff, D. Jiménez-Hernández et al. (2018) Should we treat data as labor? Moving beyond ‘free’. *American Economic Association Papers & Proceedings*, 108, 38–42.
- Autor, D., D. Dorn, L. Katz et al. (2017) *The fall of the labor share and the rise of superstar firms*. NBER Working Paper, 23396.
- Brynjolfsson, E., F. Eggers en A. Gannamaneni (2018) *Using massive online choice experiments to measure changes in well-being*. NBER Working Paper, 24514.
- Brynjolfsson, E., L. Hitt en H. Kim (2011) *Strength in numbers: how does data-driven decision-making affect firm performance?* SSRN Working Paper, 24 april.
- Ecorys (2017) *Big data and competition*. Rotterdam: Ecorys. Rapport te vinden op www.rijksoverheid.nl.
- FD (2016) Big data is geen goud, maar olie. *Het Financieel Dagblad*, 28 mei.
- Financial Times (2012) Why ‘big data’ is the fourth factor of production. *Financial Times*, 27 december.
- Hilbert, M. (2012) How to measure ‘how much information’? Theoretical, methodological, and statistical challenges for the social sciences. *International Journal of Communication*, 6, 1042–1055.
- Hilbert, M. (2015) *Quantifying the data deluge and the data drought*. SSRN Working Paper, 1 april.
- Hilbert, M. en P. López (2011) The world’s technological capacity to store, communicate, and compute information. *Science*, 332(6025), 60–65.
- IDC (2017) *Data age 2025. The evolution of data to life-critical: don’t focus on big data; focus on data that’s big*. Framingham, MA: IDC. Te vinden op www.seagate.com.
- IMF (2017) *Investment and capital stock dataset, 1960–2015*. Te vinden op [data.world: https://data.world/imf/investment-and-capital-stock-i](https://data.world/imf/investment-and-capital-stock-i).
- RAND Europe (2015) *The digital catapult and productivity: a framework for productivity growth from sharing closed data*. Cambridge, VK: RAND Europe.
- Rosenboom, N. en T. Smits (2018) Vrij toegankelijke data van Goliath helpt David niet. *Het Financieel Dagblad*, 2 februari.
- Syverson, C. (2011) What determines productivity? *Journal of Economic Literature*, 49(2), 326–365.
- VN (2017) *World population prospects: the 2017 revision*. Working Paper, ESA/P/WP/248.

In het kort

- ▶ De mondiale hoeveelheid data groeit sterk sinds de jaren tachtig.
- ▶ Niet alle bedrijven zijn data-intensief, maar data-intensieve bedrijven zijn wel innovatiever.
- ▶ Het lijkt er daarmee op dat data vooral de productiviteit verhogen.

VAN ZANTEN

Roger Klaassen

