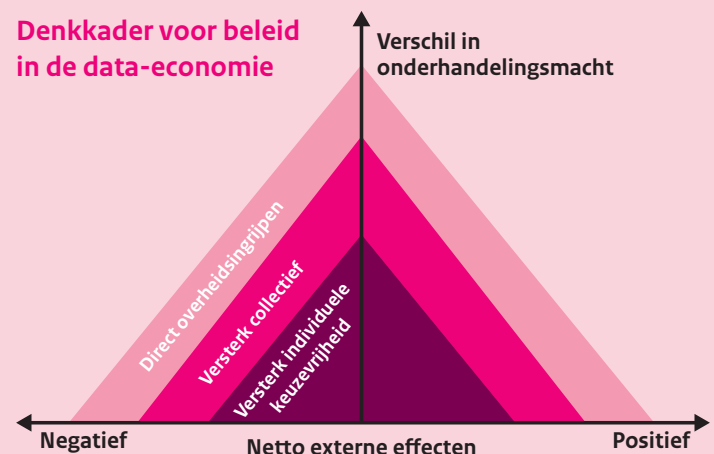




Brave new data Databeleid in een imperfecte wereld

Data spelen een steeds belangrijkere rol in onze economie. Door twee knelpunten werkt de data-economie niet optimaal: externe effecten en ongelijke machtsverhoudingen tussen dataleveranciers en dataverwerkers. Externe effecten kunnen positief en negatief zijn en leiden ertoe dat data te veel of te weinig gedeeld worden. Het gevolg van ongelijke machtsverhoudingen is dat de dataverwerker de voorwaarden bepaalt van de datatransactie.



We introduceren een denkkader dat verschillende oplossingsrichtingen om deze knelpunten aan te pakken ordent.

CPB Policy Brief

Ramy El-Dardiry, Bastiaan Overvest, Milena Dinkova,
Rob Aalbers

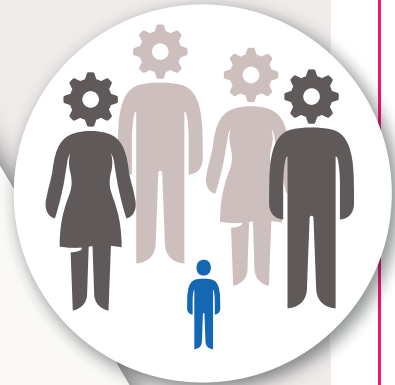
mei 2021

Denkkader voor databeleid

De rol van data in de economie is groot en groeit. Door externe effecten en ongelijke machtsverhoudingen werkt de data-economie niet optimaal. Ons denkkader – de piramide rechts – geeft inzicht in verschillende oplossingsrichtingen.

Knelpunt: ongelijke machtsverhoudingen

Vaak hebben dataverwerkers een machtspositie. Consumenten en kleine bedrijven hebben daardoor weinig keuzemogelijkheden en een slechte onderhandelingspositie.



Groot verschil in onderhandelingspositie

Probleem: dataverwerkers zijn te machtig

Voorbeeld
Verplichte data-ontbundeling

Voorbeeld
Collectieve data-overeenkomst

Voorbeeld
Prikkel voor datadelen

Voorbeeld
Gestandaardiseerde privacy-overeenkomsten

Voorbeeld
Publieke databanken

Drie oplossingsrichtingen voor een sterkere data-economie. Welke wordt gekozen, verschilt per casus en is afhankelijk van de grootte van beide knelpunten.

3. Direct overheidsingrijpen

2. Versterk collectief

1. Versterk individuele keuzevrijheid

Voorbeeld
Beperkingen op verzameling sommige data

Klein verschil

Negatief netto extern effect

Probleem: data wordt teveel gedeeld

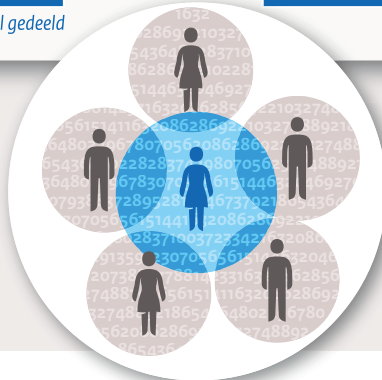
Geen extern effect

Positief netto extern effect

Probleem: data wordt te weinig gedeeld

Knelpunt: externe effecten

Doordat data ook invloed hebben op de privacy of gebruikerservaring van anderen ontstaan externe effecten.



Samenvatting

Data spelen een steeds belangrijkere rol in onze economie. Zo kunnen bedrijven met behulp van data hun producten en dienstverlening verbeteren en innoveren. Denk bijvoorbeeld aan medische datasets die worden gebruikt om algoritmes te trainen met als doel tumoren op te sporen. Of aan bedrijven die door analyse van gebruikersdata gerichtere advertenties kunnen tonen. Ook consumenten gebruiken data om beter geïnformeerde beslissingen te nemen over aankopen, bijvoorbeeld via recensies of ratings.

Hoe data worden gedeeld bepaalt vaak wie data kunnen gebruiken. In tegenstelling tot meer traditionele productiefactoren of goederen, zoals arbeid en consumptieartikelen, zijn data niet-rivaal: het gebruik van data door de één gaat niet ten koste van gelijktijdig gebruik door een ander. Toch kunnen er redenen zijn om anderen uit te sluiten van datatoegang. Bijvoorbeeld omdat data gevoelige informatie bevatten of omdat exclusieve toegang tot data lucratief is. Of data door meerdere organisaties of personen kunnen worden gebruikt, is dus vaak afhankelijk van de bereidwilligheid om data te delen.

De keuze om data te delen, ligt voornamelijk bij *individuele personen of organisaties*. Zowel de Algemene Verordening Gegevensbescherming (AVG) als de Nederlandse visie op datadelen tussen bedrijven heeft individuele zeggenschap als uitgangspunt. De AVG geeft individuele rechten aan personen zoals het recht op dataportabiliteit en verplicht bedrijven om per individu toestemming te vragen voor gebruik van persoonsgegevens. Volgens de Nederlandse datadeelvisie komt datadeling tussen bedrijven bij voorkeur vrijwillig tot stand – al worden verplichtingen niet uitgesloten.

Anderen kunnen profiteren of juist schade ondervinden van de keuze om data te delen zonder hiervoor te betalen of gecompenseerd te worden. Er treden dan externe effecten op. Zo kan de beslissing van individuen om data te delen onbedoelde gevolgen hebben voor de privacy van anderen. Bijvoorbeeld wanneer een foto met meerdere personen op sociale media wordt geplaatst of wanneer gedeelde DNA-data informatie prijsgeven over familieleden. Startups die sneller en goedkoper innoveren door de beschikbaarheid van open data zijn een voorbeeld van profijt dat door datadelen ontstaat.

Een tweede knelpunt bij datadelen is dat de onderhandelingsmacht vaak ongelijk is verdeeld. Datalleveranciers (doorgaans consumenten) hebben minder invloed op de uiteindelijke datatransactie dan dataverwerkers, omdat de laatsten vaak marktmacht hebben en de belangen *per individuele consument* klein zijn.

Zo ontstaat het probleem dat data inefficiënt gedeeld worden: soms te veel, soms te weinig. Om te begrijpen of meer datadelen per saldo positief of negatief uitpakt, is een analyse nodig van zowel de directe baten van gebruikers als de positieve en negatieve impact op anderen. Hoe deze balans uitpakt, wordt onder andere bepaald door het type data. Zo spelen privacyoverwegingen lang niet altijd een rol bij industriële data. Meer datadelen is dan vanuit maatschappelijk oogpunt sneller wenselijk. Breed toegankelijke socialemediaprofielen die gebruikt worden om verkiezingen te beïnvloeden zijn een voorbeeld van data die mogelijk te veel worden gedeeld.

Beleid dat de keuze om data te delen bij individuele partijen neerlegt, is niet geschikt om dit efficiëntieprobleem op te lossen. Individuele consumenten of bedrijven nemen de impact van datadelen op anderen onvoldoende mee en beschikken niet altijd over voldoende onderhandelingsmacht om echt te kunnen kiezen. Het huidige databeleid, dat nu juist gericht is op individuele verantwoordelijkheid, is in die zin ontoereikend. Standaardbeleidsopties om externe effecten te internaliseren zoals beprijzen of toekennen van eigendomsrechten werken niet goed bij data. Zo is het vaststellen van accijnzen voor iedere datatransactie complex. Daarnaast is het toekennen van eigendomsrechten niet eenvoudig wanneer data betrekking hebben op meerdere partijen en is het onwaarschijnlijk dat consumenten hun data exclusief verhandelen met een enkele partij.

Als de externe effecten of de verschillen in onderhandelingsmacht substantieel zijn, zien wij twee oplossingsrichtingen voor een efficiëntere inrichting van de data-economie: verschuiving van keuzebevoegdheid naar collectieven of direct overheidsingrijpen. Deze oplossingsrichtingen zijn geen panacee. Zo is het maar de vraag of de overheid altijd beschikt over de benodigde informatie en is de effectiviteit van bepaalde oplossingsrichtingen nog onzeker. Niettemin kan het in bepaalde gevallen zinvol zijn om hier toch beleid op te voeren, al was het maar omdat in sommige gevallen de negatieve gevolgen van niets doen groot zijn.

1 Inleiding

Of course there will be challenges. But these digital assistants will be so useful that everyone will want one, and the scare stories you read today about privacy concerns will just seem quaint and old-fashioned. – Hal Varian, 2013

Under the regime of surveillance capitalism [...] scientists are not recruited to solve world hunger or eliminate carbon-based fuels. Instead, their genius is meant to storm the gates of human experience, transforming it into data and translating it into a new market colossus that creates wealth by predicting, influencing, and controlling human behavior. – Shoshana Zuboff, 2019

Digitale producten en diensten veranderen ons leven. Data staan aan de basis van deze verandering. Bovenstaande citaten illustreren het levendige debat dat over datagebruik is ontstaan¹. Enerzijds bieden datagedreven toepassingen economische en maatschappelijke kansen. De Europese datastrategie² is erop gericht deze kansen te benutten. Digitalisering maakt het bijvoorbeeld mogelijk om onderwijs en zorg te personaliseren of op afstand aan te bieden (zie bijv. Miller e.a. (2019)). De mogelijkheden van automatisering worden groter doordat algoritmes worden getraind op grote datasets. De zelfrijdende auto is hier een voorbeeld van.

Maar ook de zorgen over de verzameling en het gebruik van data nemen toe. De autonomie van het individu staat onder druk doordat algoritmes steeds beter in staat zijn ons gedrag te voorspellen en te beïnvloeden. Bovendien hebben een handvol techbedrijven de meeste data in handen. Het is onzeker in welke mate onze economie en democratie kunnen profiteren van digitalisering.

In Nederland bracht de coronacrisis verschillende opvattingen over datagebruik aan het licht. De suggestie om een app te gebruiken om mogelijk besmette personen te traceren, leidde tot onrust. Wetenschappers stuurden een brandbrief³ naar het kabinet waarin werd opgeroepen zorgvuldig met fundamentele vrijheden om te gaan. Als samenleving zoeken we kennelijk nog naar manieren waarop het gebruik van data in balans gebracht kan worden met grondrechten en gemeenschappelijke belangen.

In deze *Policy Brief* verkennen we beleidsopties voor dataregulering waarin een afweging tussen gemeenschappelijke belangen en individuele grondrechten kan plaatsvinden. Door naast beleid gericht op individuele consumenten en bedrijven ook naar collectieve instrumenten voor datadelen te kijken, ontstaat een breder palet aan beleidsopties. Centraal staat telkens de vraag of het vanuit maatschappelijk perspectief wenselijk is om meer of minder data te delen ten opzichte van de huidige situatie. De analyses en aanbevelingen in dit rapport komen voort uit een literatuurstudie en een reeks interviews met wetenschappers uit binnen- en buitenland. Een uitgebreidere literatuurbespreking is te vinden in het achtergronddocument.

Deze beleidsnotitie geeft richtingen, geen antwoorden. In hoofdstuk 4 bespreken we welke oplossingsrichtingen onder welke omstandigheden het meest kansrijk lijken. Hierbij tekenen we aan dat over de uitvoerbaarheid nog veel onduidelijk is. Uiteindelijk zal alleen met behoedzaam

¹ Zie Zuboff (2015) voor een nadere analyse van de lezing van Hal Varian ([link](#) naar lezing)

² Europese Commissie, *A European strategy for data*, 19 februari 2020 ([link](#))

³ [Link](#)

experimenteren en onderzoek hierover meer duidelijk worden. Sommige beleidsopties vragen daarbij ook om nieuwe wettelijke en institutionele kaders. Regulering waarbij de zeggenschap over data verschuift van individuen naar collectieven vereist bijvoorbeeld verruiming van de Algemene Verordening Gegevensbescherming (AVG).

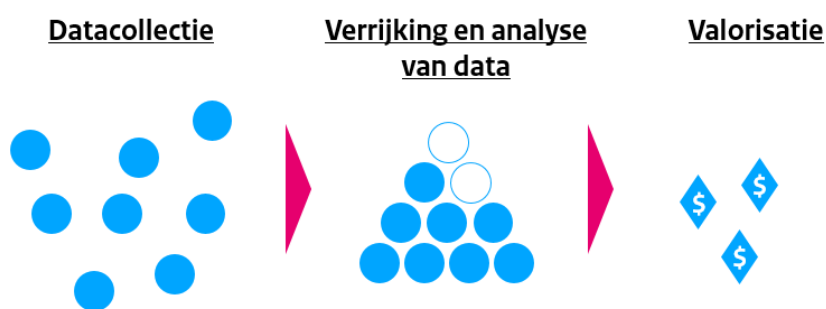
2 Kenmerken van de data-economie

De hoeveelheid data neemt snel toe. Per dag worden ongeveer 300 miljard e-mails verstuurd, 400 miljoen foto's op sociale media geplaatst en 5 miljard zoekopdrachten gedaan.⁴ Miljarden apparaten verzamelen data en wisselen data uit.

Verder wordt het vermogen om deze groeiende hoeveelheid data te analyseren steeds groter. Door toenemende reken capaciteit en ontwikkelingen op het gebied van data science kunnen bijvoorbeeld teksten, beelden en geluid steeds beter digitaal worden ontleed en tot nut worden gemaakt.

Data worden gebruikt om producten en diensten te maken en zijn daarmee een productiefactor geworden. Voorbeelden zijn verkeersdata die een algoritme voeden voor zelfrijdende auto's, internetdata die gepersonaliseerde advertenties mogelijk maken of data voor optimalisatie van bedrijfsprocessen. In de 'datawaardeketen' (figuur 1) worden data eerst verzameld en opgeslagen. Vervolgens veranderen data in informatie, via aanvullende databronnen en analyses. De informatie wordt uiteindelijk omgezet in waarde via producten en diensten. Met de term data-economie verwijzen we naar dat deel van de economie dat geraakt wordt door activiteiten binnen de datawaardeketen⁵.

Figuur 1 De datawaardeketen



In de data-economie ontstaan verschillende typen data. Data komen vaak voort uit administratieve handelingen of transacties. Denk bijvoorbeeld aan inschrijvingen bij de gemeente of foto's plaatsen op sociale media. Data kunnen echter ook ontstaan door te observeren of door

⁴ Bron: Raconteur ([link](#)).

⁵ Hiermee sluiten we aan bij de definitie zoals gehanteerd door de Europese Commissie ([link](#)): "The data economy measures the overall impacts of the data market – i.e. the marketplace where digital data is exchanged as products or services derived from raw data – on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis, elaboration, delivery, and exploitation of data enabled by digital technologies."

patronen af te leiden uit grote hoeveelheden data. De data die vrijkomen, lopen uiteen in de mate van geordendheid en persoonsgevoeligheid.

Data zijn ‘niet-rivaal’. Dat betekent dat het gebruik van data door de één niet ten koste gaat van het (gelijktijdig) gebruik door een ander (Carrière-Swallow en Haksar, 2019). Data zijn dus geen klassieke private goederen zoals voedsel, meubilair en kleding.

Vaak moeten data eerst gedeeld worden, voordat data kunnen worden gebruikt. Data verschillen in de mate waarin ze voor derden toegankelijk zijn. Door deze uiteenlopende toegankelijkheid zijn data ook niet direct vergelijkbaar met klassieke publieke goederen zoals schone lucht of dijken. Soms zijn er wettelijke verplichtingen, zoals bijvoorbeeld bij een resultatenrekening, die ervoor zorgen dat data openbaar toegankelijk zijn. In andere gevallen zijn data zo eenvoudig te bemachtigen, dat uitsluiten ondoenlijk is. Denk bijvoorbeeld aan data over het weer. Meestal moet er echter een keuze worden gemaakt om data voor anderen beschikbaar te maken. Deze keuzes vinden onder andere plaats wanneer consumenten transacties aangaan met bedrijven of wanneer bedrijven besluiten data met elkaar uit te wisselen.

Het huidige beleid legt de keuze om data wel of niet te delen in eerste instantie bij individuele consumenten of bedrijven. De eerste grondslag van de AVG stelt dat persoonsgegevens verwerkt mogen worden wanneer consumenten hiervoor toestemming verlenen⁶. De dataverwerker dient hierbij de consument vooraf voldoende te informeren over onder andere de noodzaak van de toegang tot data. In de Nederlandse datadeelvisie⁷ bespreekt het kabinet de economische mogelijkheden van datadelen tussen bedrijven en benadrukt daarbij dat datadelen bij voorkeur vrijwillig tot stand komt.

De waarde van data voor de economie hangt ook af van de verdeling van gebruiksrechten. Data over een machine van een bedrijf zijn waardevol voor dat bedrijf zelf, maar kunnen ook waardevol zijn voor de machineproducent, die daarmee aanvullende diensten kan leveren of de machine kan verbeteren. Data van miljoenen socialemediagebruikers of streamingdiensten helpen om gerichte advertenties te tonen of een gepersonaliseerd media-aanbod te bieden. Datadelen kan ook het publieke belang dienen, zo kunnen data van weggebruikers nuttig zijn voor de beheerders van infrastructuur of openbaar vervoermaatschappijen.

Het gebruik van data laat afnemende meeropbrengsten zien. Data over een enkele gebruiker hebben voor dataverwerkers weinig waarde, tenzij die data specifieke toepassingen voor een gebruiker mogelijk maken. Pas naarmate een bedrijf meer data weet te verzamelen, worden data waardevol en kunnen data gebruikt worden in het productieproces. Net als bij andere productiefactoren kennen data voor gebruik binnen een bepaalde toepassing op den duur ook afnemende meeropbrengsten. Een extra datapunt levert dan steeds kleinere prestatieverbeteringen op.

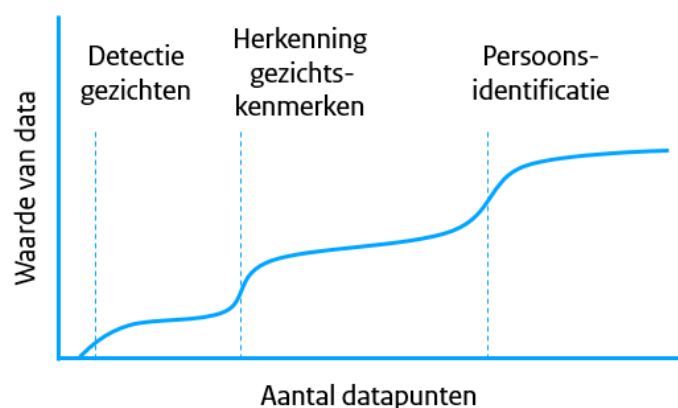
Maar soms kunnen meer data nieuwe toepassingen mogelijk maken; de meeropbrengsten nemen dan toe. Neem bijvoorbeeld een algoritme dat als doel heeft menselijke gezichten op foto's te detecteren. Vanaf een bepaalde hoeveelheid data zal dit algoritme nauwelijks nog nauwkeuriger

⁶ Dit staat in artikel 6 van de AVG.

⁷ Zie kamerbrief 20 februari 2019 ([link](#))

worden. Toch kan het interessant zijn om meer data in de vorm van foto's te verzamelen, omdat door de extra data andere, complexere toepassingen mogelijk worden. In het geval van foto's kunnen algoritmes bijvoorbeeld worden getraind om kenmerken van gezichten te onderscheiden en uiteindelijk personen te identificeren. De waarde van data verloopt dan met sprongen, zie figuur 2.

Figuur 2 Illustratie van impact extra data op waarde van data



Bron: vrij naar Posner en Weyl (2018)

Waar wringt het? Op zichzelf zijn de eigenschappen van data zoals beschreven in dit hoofdstuk niet uniek. Niet-rivaliteit en beperkte uitsluitbaarheid komen vaker voor. Lucht of dijken zijn bijvoorbeeld ook niet-rivaal. Er ontstaat een probleem, of er worden kansen gemist, wanneer er marktfalen zijn. Het volgende hoofdstuk gaat in op specifieke marktfalen in de data-economie.

3 Knelpunten in de data-economie

Door twee knelpunten werkt de data-economie niet optimaal: externe effecten en ongelijke machtsverhoudingen tussen dataleveranciers (vaak consumenten of kleine bedrijven) en dataverwerkers. Externe effecten treden op doordat individuele keuzes ook de privacy en gebruikerservaring van anderen beïnvloeden en doordat organisaties data vrijgekomen uit datatransacties van derden tot nut kunnen maken. Deze knelpunten worden in dit hoofdstuk verder toegelicht. Het huidige beleid voor data heeft nog geen goed antwoord op deze knelpunten. Dit suggereert dat via beleidsaanpassingen welvaartswinst geboekt kan worden.

3.1 Externe effecten

Externe effecten ontstaan wanneer iemands persoonlijke gegevens ook invloed hebben op de privacy of gebruikerservaring van anderen. Informatie over personen is vaak vervlochten: denk aan foto's waar meerdere personen op staan, gesprekken tussen personen of DNA dat onthullende informatie bevat over verwanten. Allemaal voorbeelden van data die het individu overstijgen. Daarnaast kunnen persoonlijke gegevens gebruikt worden om over andere personen voorspellingen te doen. Dit laatste gebeurt wanneer dataverwerkers met grote hoeveelheden data profielen maken over personen over wie ze zelf geen data hebben (MacCarthy, 2010). Uit data van een groep personen kunnen bedrijven zeer uiteenlopende informatie afleiden over andere personen, zoals politieke

voorkeuren, seksuele geaardheid⁸ of etnische achtergrond⁹. Een belangrijk gevolg van deze gecorreleerde informatie is dat de keuze van mensen om data te delen met een bedrijf (“de datatransactie”) impact heeft op de privacy of gebruikerservaring van anderen die niet bij de datatransactie betrokken zijn (Acemoglu e.a., 2019). Deze externe effecten kunnen voor individuen zowel negatief als positief zijn. Zo zal een *influencer* op sociale media het wellicht toejuichen als een foto van hem of haar op internet wordt gedeeld, terwijl een ouder het online plaatsen van foto’s met zijn of haar kinderen niet kan waarderen. Als medische scans worden gebruikt om diagnostische algoritmes te trainen waardoor de kwaliteit van de diagnose over de tijd verbetert, zal een volgende patiënt daar baat bij hebben. De grootte van het externe effect is daarnaast afhankelijk van het type data en de manier waarop de data wordt gebruikt.

Ook wanneer organisaties datatransacties aangaan, kunnen externe effecten ontstaan. Als data van een bedrijf of (semi-)publieke instelling toegankelijk worden voor anderen, kunnen deze op nieuwe manieren toegepast worden. Neem data uit de gezondheidszorg. Bedrijven gespecialiseerd in kunstmatige intelligentie kunnen op basis van scans of echobeelden algoritmes trainen en daarmee de nauwkeurigheid van diagnostische hulpmiddelen verbeteren. En medische wetenschappers kunnen met data uit meerdere zorginstellingen beter bepalen welke behandelingen werken of welke bijwerkingen optreden. Externe effecten kunnen ook bij datagebruik door organisaties negatief zijn. Wanneer een dataverwerker bijvoorbeeld investeert in cyberveiligheid weegt deze verwerker de effecten van een datalek op anderen mogelijk onvolledig mee (CPB, 2016).

Door deze externe effecten worden data inefficiënt gedeeld. Zoals beschreven in hoofdstuk 2 is het uitgangspunt van het huidige databeleid dat consumenten en bedrijven individueel bepalen of zij data delen. Een consument is vrij om zijn of haar persoonsgegevens te delen met een dataverwerker, ook wanneer deze data (indirect) informatie bevatten over andere personen. Een bedrijf zal, uit vrees voor concurrentie, veelal niet geneigd zijn om data te delen of publiek toegankelijk te maken (Jones en Tonetti, 2020). De potentiële maatschappelijke waarde van data wordt dan onvoldoende benut. Vanuit maatschappelijk oogpunt kunnen door de aanwezigheid van externe effecten afhankelijk van het type data en het gebruik ervan zowel te weinig als te veel gegevens gedeeld worden.

Er zijn geen voor de hand liggende oplossingen voor dit probleem. De gebruikelijke oplossingen om externe effecten te internaliseren, werken niet goed bij data. Een klassieke oplossing is om eigendomsrechten te definiëren. De uitstoot van schadelijke stoffen zoals CO₂ of stikstofoxiden wordt bijvoorbeeld in verschillende landen gereduceerd via het systeem van verhandelbare emissierechten. Eigendomsrechten exclusief toekennen is in het geval van data niet altijd eenvoudig. Neem data die voortkomen uit financiële transacties: beide zijden van de transactie beschikken automatisch over deze data. Daarnaast is het onwaarschijnlijk dat consumenten persoonsgegevens exclusief verhandelen met een enkele partij. Consumenten kunnen data immers meerdere malen aanbieden zonder dat de data hun functionaliteit verliezen.

De aanwezigheid en grootte van externe effecten is bovendien afhankelijk van de wijze waarop data later worden gebruikt. Neem bijvoorbeeld data die ontstaan bij gebruik van sociale

⁸ Zie Jernigan, C. en Mistree, B. F., 2009, Gaydar: Facebook friendships expose sexual orientation. *First Monday*, 14(10).

⁹ Zie [dit](#) artikel op *Ars Technica*.

mediaplatformen. Platformen kunnen deze gebruikersdata op velerlei manieren inzetten. Zo kunnen ze hun eigen gebruikersinterface veranderen of gepersonaliseerde advertenties aanbieden. In beide gevallen zijn er externe effecten – het product verandert immers voor alle gebruikers, maar consumenten zullen deze externe effecten anders beoordelen. Het platform heeft ook de keuze om de data vooralsnog niet te gebruiken waardoor externe effecten, in ieder geval tijdelijk, achterwege blijven.

Ook het beprijzen van datatransacties met externe effecten is complex. Beprijzen vormt een tweede klassieke oplossing voor externe effecten. Dit kan bijvoorbeeld in de vorm van een subsidie voor transacties met positieve externe effecten en een accijns voor transacties met een negatief extern effect. Wederom is deze oplossing in de praktijk lastig te implementeren, omdat de omvang van het externe effect zowel afhangt van het type data (gevoelig of niet), de individuele privacyvoorkeuren van consumenten, en de toepassing (wordt data bijvoorbeeld gebruikt voor wetenschappelijk onderzoek of voor het beprijzen van verzekeringen?). Bovendien is het monitoren van alle datatransacties en de wijze waarop data gebruikt wordt praktisch gezien vrijwel onuitvoerbaar.

3.2 Ongelijke machtsverhoudingen

Naast externe effecten vormen ongelijke machtsverhoudingen een tweede knelpunt in de data-economie. In deze paragraaf bespreken we hoe deze ongelijkheid tot stand komt.

Onderhandelingsmacht is in de data-economie vaak ongelijk verdeeld. Als consument, werknemer of burger verstrekken mensen voortdurend data, maar vaak hebben ze beperkte zeggenschap of inzicht over hoe hun data gebruikt worden door de dataverwerker. Deze ongelijke onderhandelingsmacht heeft verschillende oorzaken: marktmacht, informatie-asymmetrie en cognitieve beperkingen. Het gevolg is dat vooral de dataverwerker binnen de wettelijke randvoorwaarden bepaalt welke persoonsgegevens verzameld worden en hoe deze data gebruikt worden, en niet de dataleverancier.

In sommige sectoren zijn de keuzemogelijkheden voor consumenten beperkt vanwege marktmacht. Sommige platformdiensten hebben een dominante positie, waardoor een consument moeilijk “met de voeten kan stemmen” over het databeleid van het platform. Als al je vrienden op Facebook of WhatsApp zitten, zal je sneller geneigd zijn om de privacy-voorwaarden te accepteren dan wanneer je meer keuze hebt. Ook in andere sectoren kan dit probleem spelen: overheidsdiensten hebben vaak een monopolie en kunnen binnen de wettelijke grenzen eenzijdig de datavooraanwaarden aanpassen.

Bovendien kunnen informatie-asymmetrie en cognitieve beperkingen de onderhandelingspositie van consumenten verslechteren. Consumenten zien niet wat bedrijven met hun data doen. Ook kunnen ze niet alle datagebruikersvoorwaarden lezen of de gevolgen ervan overzien. De sterke informatiepositie van bedrijven kan uitgebuit worden door de voorkeuren van consumenten te manipuleren en hen te verleiden tot de aanschaf van producten of het delen van meer data (Acquisti e.a., 2015). Voorbeelden zijn te vinden bij boekingswebsites en online retailers die doelbewust schaarste suggereren om consumenten tot aanschaf te verleiden.

Ongelijke machtsverhoudingen doen zich ook voor tussen bedrijven. Bedrijven beschikken soms over data die mogelijk door andere bedrijven gebruikt kunnen worden om nieuwe diensten of producten te ontwikkelen. De data van het ene bedrijf kunnen een essentiële input zijn voor het andere bedrijf. Wanneer een bedrijf als enige beschikt over de benodigde data, dan hebben ontvangende bedrijven weinig onderhandelingsmacht. Zo heeft Facebook na de Cambridge Analytica-casus zijn gebruikersvoorwaarden voor de API (application programming interface) aangepast. Strengere voorwaarden waren deels nodig om risico op misbruik te verkleinen. Tegelijkertijd werden kleinere bedrijven en ontwikkelaars door deze strengere gebruikersvoorwaarden van essentiële data afgeknepen en kwamen daardoor vervolgens in de problemen. Soms ontzeggen platformen opkomende concurrenten toegang tot data om ze daarmee uit te sluiten van concurrentie¹⁰.

Europees beleid is erop gericht een deel van de ongelijke machtsverhoudingen weg te nemen. De AVG geeft iedere dataleverancier (consument) individuele rechten, maar omdat het aantal dataleveranciers groot is en de belangen per individuele dataleverancier klein zijn, is de onderhandelingspositie met de dataverwerker ongelijk. Het bestaande beleid (wellicht met uitzondering van het mededingingstoezicht) laat de machtsverhoudingen tussen bedrijven ongemoeid. De Europese Commissie ontwikkelt daarom, o.a. via de *Data Governance Act*, de *Digital Services Act* en de *Digital Markets Act*, nieuwe regels om machtsverhoudingen gelijkjer te trekken. Zo worden grote platformen verplicht gesteld om onder voorwaarden derde partijen toe te staan services te ontwikkelen die gebruikmaken van hun platformen (interoperabiliteit). Ook voorziet deze voorgestelde wetgeving in de mogelijkheid om organisaties op te zetten die de belangen van kleine bedrijven en consumenten bundelen.

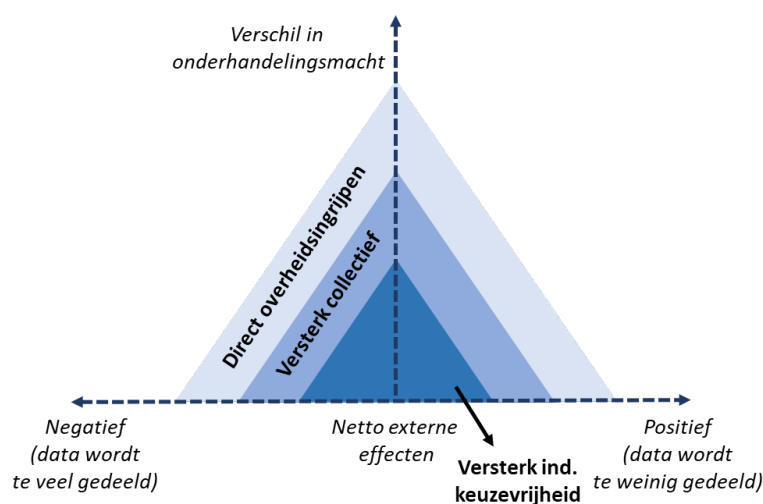
¹⁰ Twitter heeft bijvoorbeeld het data-analyse bedrijf DataSift toegang ontzegd nadat het zelf een vergelijkbaar bedrijf had overgenomen. Zie [dit](#) bericht voor een bespreking van meerdere relevante casussen.

4 Richtingen voor beleid

4.1 Denkkader voor beleid in de data-economie

We introduceren een denkkader dat de verschillende oplossingsrichtingen ordent. Zie figuur 3. Deze figuur is opgebouwd uit drie lagen. De vorm waarin de overheid ingrijpt, is daarbij afhankelijk van zowel de omvang als het teken van de (netto) externe effecten. In de binnenste laag van de piramide zijn de knelpunten klein: per saldo zijn de externe effecten en de verschillen in onderhandelingsmacht gering. In een dergelijke situatie kunnen mensen en bedrijven het beste zelf de afweging maken tussen de voor- en nadelen van datadelen.¹¹

Figuur 3: Denkkader voor beleid in de data-economie



Collectieve instrumenten komen in beeld als de externe effecten of het verschil in onderhandelingsmacht groter worden. De centrale gedachte hierbij is dat groepen consumenten gezamenlijk met een andere partij in onderhandeling treden. Het klassieke instrumentarium van economen, waarbij consumenten en bedrijven, bijvoorbeeld via belastingen, geprikkeld worden om hun gedrag te wijzigen werkt hier immers niet. Dit wordt gerepresenteerd door de middelste laag van het beleidskader. We bespreken de beleidsopties voor deze middelste laag in paragraaf 4.3.

Direct ingrijpen kan in beeld komen als de externe effecten verbonden aan data of het verschil in onderhandelingsmacht (zeer) groot zijn. In dit geval kan de overheid direct ingrijpen door datagebruik te beperken, verplichtingen op te leggen of zelf data te verzamelen en toegankelijk te maken, zie paragraaf 4.4. Voordat we individuele beleidsopties bespreken, leggen we in de volgende paragraaf uit hoe het kader in de praktijk gebruikt kan worden en welke issues daarbij spelen.

¹¹ Hoewel er wel een aantal opties zijn om dit beleid verder te versterken, besteden we daar in deze Policy Brief geen aandacht aan. Zie het bijbehorende AD voor beleidsrichtingen die passen bij deze situatie.

De analyse in dit hoofdstuk beperkt zich tot het economische perspectief. We merken op dat het niet louter een economische keuze is *waar* in de datawaardeketen wordt ingegrepen. Er kunnen bijvoorbeeld ethische redenen zijn om bepaalde typen data überhaupt niet te verzamelen, zoals data over etniciteit.

4.2 Toepassing van het denkkader in de praktijk

Het denkkader helpt bij het vinden van oplossingen, maar voor een specifieke casus is het vaak onduidelijk waar het probleem zich in het denkkader bevindt. Neem bijvoorbeeld het externe effect van privacyschending dat ontstaat door datadeling. De overheid heeft in veel gevallen onvoldoende informatie om de omvang van de externe effecten te bepalen. Het belang dat individuen aan privacy hechten verschilt namelijk per persoon en is bovendien afhankelijk van het type data, wie de data bezit en wat er met de data gebeurt. Bovendien zal de rationale of noodzaak tot overheidsingrijpen deels afhangen van de context.

Niettemin kan inzicht in de aard en omvang van het probleem behulpzaam zijn. Startpunt voor het nadenken over oplossingsrichtingen is een goed begrip van het type data en het gebruik ervan. Dit maakt het vervolgens mogelijk om te kijken welke externe effecten zich voordoen en wat de gevolgen zijn van een eventueel verschil in onderhandelingsmacht. Zodra de problemen voor een bepaald type data en toepassing helder in beeld zijn gebracht, kan maatschappelijk worden afgewogen hoe *groot* deze problemen zijn en welk soort overheidsingrijpen gerechtvaardigd is.

Bij externe effecten is de netto-impact bepalend voor de vraag of de overheid datagebruik voor een toepassing moet stimuleren of juist moet inperken. Externe effecten verbonden aan het gebruik van een bepaald type data kunnen zowel positief als negatief zijn. Neem het voorbeeld van DNA-data die door ziekenhuizen gebruikt worden voor wetenschappelijk onderzoek. Voor een deel van de bevolking kan het delen van hun DNA-data gepaard gaan met een verlies van privacy; voor een ander deel van de bevolking kan het delen van deze data juist positief zijn. Bovendien hebben ziekenhuizen niet altijd de juiste prikkels om de data die in hun bezit zijn met elkaar te delen. De overheid dient deze verschillende externe effecten te wegen om tot een besluit te komen of DNA-data voor deze toepassing juist meer of minder gedeeld moet worden en onder welke randvoorwaarden.

Dataregulering is complex. Omdat de externe effecten van data(gebruik) afhankelijk (kunnen) zijn van het type data, de oorspronkelijke eigenaar, het doel van de toepassing en de (rechts)personen waar de data betrekking op hebben, is het niet eenvoudig om de netto-omvang van de externe effecten vast te stellen. Het aantal casussen dat moet worden onderzocht, is schier oneindig. Het is noch realistisch noch doelmatig om voor al deze casussen separaat regulering vast te stellen. Om met deze complexiteit om te gaan is prioritering nodig – welke toepassingen lijken *a priori* het meest problematisch en voor welke toepassingen volstaat bestaande regulering? Bovendien ontstijgen knelpunten in de data-economie vaak nationale grenzen. Grote dataverwerkers bevinden zich dikwijls buiten Nederland of de Europese Unie. Effectieve regulering en toezicht vereisen daarom internationale coördinatie of afspraken.

Gegeven deze complexiteit is opbouw van kennis bij de overheid onontbeerlijk om maatschappelijke discussies over dataregulering te kunnen voeren. De Tweede Kamer heeft in 2020 een commissie ingesteld om meer grip te krijgen op digitalisering. Hoewel meerdere ministeries, planbureaus, adviesraden, toezichthouders, universiteiten en hogescholen informatie aanreiken over data en digitalisering, blijft het voor volksvertegenwoordigers lastig om (de impact van) digitalisering voldoende te doorgronden¹². De versnippering van kennis en onderzoek bij meerdere organisaties brengt het risico met zich mee dat kennisopbouw op het gebied van datagebruik onvoldoende is geborgd en dat de maatschappelijke discussies over externe effecten en onderhandelingsmacht te weinig vorm krijgen¹³. Om kennis gericht te ontwikkelen is het raadzaam om onderzoeksagenda's beter te coördineren of de oprichting van specifieke kennisinstellingen te overwegen.

4.3 Versterking van het collectief

Collectieve actie is een kans op een beter werkende data-economie. Externe effecten kunnen beter worden meegewogen wanneer onderhandelingen over datagebruik (deels) verschuiven van individuen naar gemeenschappen. Door de samenwerking tussen dataleveranciers te versterken, kunnen externe effecten zoals op het gebied van privacy zichtbaar worden. Voor de overheid is het voordeel verder dat zij geen inzicht hoeft te hebben in de precieze omvang van externe effecten.

Daarnaast kunnen door krachten te bundelen dataleveranciers hun onderhandelingsmacht ten opzichte van dataverwerkers vergroten. Bij sommige vormen van dataverzameling en gebruik is individuele toestemming bovendien onrealistisch. Denk bijvoorbeeld aan 360-gradenopnames van straten.

Een route voor meer collectieve actie is de vorming van collectieve data-overeenkomsten (cdo's). Bij een cdo sluit een groep dataleveranciers (zoals consumenten) een overeenkomst met een dataverwerker (zoals een bedrijf) over datatransacties¹⁴. Het belangrijkste idee hierbij is dat gezamenlijk optrekken van een groep dataleveranciers het waarschijnlijker maakt dat een economisch optimale uitkomst wordt bereikt. Zo'n collectieve overeenkomst kan op meerdere manieren vorm krijgen. Het kan bijvoorbeeld gestalte krijgen via consumentvriendelijkere algemene voorwaarden of – in een meer extremere vorm – via een collectief contract. Bij een dergelijk contract wordt, net als bij cao's in arbeidsrelaties, individuele contractvrijheid opgegeven. Omgevingen waar geëxperimenteerd kan worden met cdo's zijn plekken waar al natuurlijke groepen van "dataleveranciers" bestaan, zoals scholen (ouderraden), werk (ondernemingsraden) of sportverenigingen (besturen of algemene ledenvergaderingen).

Cdo's zijn nieuw en de praktische uitvoering is onzeker. Aangezien cdo's nog niet bestaan, is onbekend wat praktische belemmeringen zijn voor succesvolle onderhandelingen en aan welke voorwaarden een cdo moet voldoen. We zien twee grote uitdagingen. Ten eerste hebben individuen waarschijnlijk een lage prikkel om zich te organiseren in een collectief. Een data-transactie is voor

¹² Tweede Kamer, Rapport tijdelijke commissie Digitale toekomst - Naar meer parlementaire grip op digitalisering, 2020 ([link](#))

¹³ Zie ook recent AWTI-rapport, Rijk aan kennis - Naar een herwaardering van kennis en expertise in beleid en politiek, 2021 ([link](#))

¹⁴ Dit idee heeft parallellen met *community consent* in geneeskundig onderzoek. Jeni Tennison beschrijft in een blogpost *community consent* voor de data-economie ([link](#)). Deze blogpost ([link](#)) van Anouk Ruhaak bespreekt *collective consent*.

individuele consumenten vaak een verwaarloosbare gebeurtenis. Meeliftgedrag kan ertoe leiden dat niemand tijd steekt in een cdo-onderhandeling. De belangen van gebruikers kunnen daarnaast uiteenlopen. Deze problemen doen zich bijvoorbeeld voor bij gebruikers van sociale media of zoekmachines. Ten tweede kunnen voordelen van een cdo voor dataverwerkers te laag zijn. Deze voordelen liggen grotendeels in het verlagen van transactiekosten die het huidige informed consent model met zich meebrengt.

Naast deze organisatorische uitdagingen vragen cdo's ook om nieuwe wettelijke en institutionele kaders. De voorgestelde Europese *Data Governance Act* biedt ruimte voor de oprichting van datacoöperaties. Een mogelijkheid is dat de cdo een uitbreiding wordt van de AVG. Dit kan een vrijstelling onder voorwaarden van de AVG zijn of een wettelijke verplichting om een cdo af te sluiten in specifieke gevallen. Daarnaast kan de prikkel om een cdo af te sluiten worden vergroot door de *outside option* – de situatie waarbij geen cdo wordt afgesloten – te veranderen; denk bijvoorbeeld aan beperkingen voor gepersonaliseerde advertenties. De nieuwe kaders zullen ook moeten beschrijven hoe een cdo gemonitord kan worden en wat er gebeurt bij conflicten.

Ook bij datadeling tussen organisaties kan collectieve actie knelpunten verkleinen.

Samenwerking tussen bedrijven biedt een kans om externe effecten in de data-economie te internaliseren. Zo verbetert het delen van kennis en data over cyberaanvallen- en bedreigingen de beveiliging van data voor het collectief. Ongelijke onderhandelingsposities tussen organisaties doen zich onder andere voor wanneer data van één partij essentieel zijn voor een groot aantal andere bedrijven of wanneer kleinere organisaties data afstaan aan een grote dataverwerker, vaak een techbedrijf. Ook hier kan het verschil in onderhandelingsmacht tussen organisaties mogelijk worden verminderd door collectieve onderhandelingen. Een alternatief is om discriminatoire toegangsvoorwaarden te verbieden.

Wanneer de netto externe effecten van datadelen tussen bedrijven groot (en positief) zijn, liggen extra prikkels voor datadelen voor de hand. Dit kan bijvoorbeeld door een bedrijf alleen toegang tot een publieke databank te verlenen als het bedrijf ook eigen data inbrengt of om datadelen als voorwaarde te stellen bij ontvangst van een subsidie of bij markttoetreding. Een mogelijk nadeel van dit type maatregelen is dat organisaties weinig prikkels ondervinden om data met organisaties te delen die zelf niet over data beschikken. Hierdoor blijven potentiële welvaartswinsten van datadelen onbenut.

4.4 Direct overheidsingrijpen

Wanneer netto externe effecten of verschillen in onderhandelingsmacht verder toenemen, is direct overheidsingrijpen wenselijk. Dit is niet uniek voor de data-economie. Fundamenteel onderzoek kent bijvoorbeeld positieve externe effecten en wordt daarom gesubsidieerd door de overheid. Wetgeving op het terrein van consumentenbescherming verkleint verschillen in onderhandelingsmacht. Daarnaast borgen mededingingsautoriteiten concurrentie in een economie. Recent zijn er verschillende voorstellen gedaan voor de versterking van dit, deels extraterritoriale, toezicht in de data-economie (Crémer et al., 2019 en EZK, 2019-2).

In de data-economie kan de overheid – naast verschillende vormen van toezicht – op drie manieren directer ingrijpen: via wettelijke verplichtingen, beperkingen of door zelf data te

verzamen en toegankelijk te maken. Het doel van deze paragraaf is om een aantal mogelijke beleidsrichtingen en de voor- en nadelen daarvan verder toe te lichten. In het achtergronddocument komen enkele van deze opties nader aan bod.

Binnen de datawaardeketen heeft de overheid verschillende opties om via verplichtingen onderhandelingsmacht beter in balans te brengen. De overheid, nationaal of Europees, kan bijvoorbeeld (platform)bedrijven verplichten om alleen persoonsgegevens te verwerken die decentraal opgeslagen zijn in een persoonlijke dataruimte van consumenten¹⁵. Een persoonlijke dataruimte geeft consumenten meer zeggenschap over data en verlaagt overstapkosten tussen diensten. De onderhandelingsmacht van consumenten wordt daardoor vergroot. In de tweede stap van de datawaardeketen – waar data wordt geanalyseerd en verrijkt – kan de overheid ingrijpen door bedrijven te verplichten scheidingen aan te brengen tussen verschillende datagedreven diensten ('mandatory data unbundling', zie Condorelli en Padilla (2020)). Het is voor consumenten makkelijker te begrijpen wat er met data gebeurt wanneer de data slechts voor één bepaald product of dienst worden gebruikt. De informatie-asymmetrie wordt dan verkleind. Verschillen in onderhandelingsmacht tussen bedrijven kunnen worden aangepakt door bijvoorbeeld het al bestaande recht op dataportabiliteit – het recht om gegevens over te dragen – ook te verlenen aan zakelijke gebruikers van platformen of om te eisen dat diensten interoperabel zijn.

Bij netto negatieve externe effecten liggen maatregelen om minder data te verzamelen voor de hand. Een verplichting tot het aanbieden van een bepaald type businessmodel, zoals abonnementen, vermindert de prikkels voor bedrijven om data te verzamelen. Een nadeel van deze optie is dat bepaalde producten en diensten mogelijk alleen voor een kleinere groep consumenten beschikbaar komen of dat privacy enkel voor welgestelden bereikbaar wordt. Het al eerder genoemde verplicht ontbundelen van data is een andere manier om de prikkel zoveel mogelijk data te verzamelen te verkleinen. Een voordeel van maatregelen zoals ontbundelen en het verplicht stellen van een bepaald businessmodel is dat ze generiek zijn: de overheid heeft geen specifieke informatie over de externe effecten van data nodig om ze toe te passen¹⁶. Wel is onzeker in welke mate deze maatregelen externe effecten daadwerkelijk verminderen en hoe groot nadelige effecten uitvallen.

De overheid kan voor specifieke data verboden opleggen. Zo kan de overheid bijvoorbeeld eisen dat organisaties voor de opslag en analyse van DNA-data een vergunning nodig hebben. Verder kan de overheid bepaalde datatoepassingen verbieden, terwijl ze het gebruik van diezelfde data voor wetenschappelijk onderzoek toestaat.¹⁷ Deze vorm van regulering wordt echter al snel complex. Externe effecten worden namelijk niet alleen bepaald door het type data maar ook door de toepassing en door de vraag wie de data gebruikt en beheert. Hierdoor ontstaan veel combinaties die in potentie aparte regelgeving vereisen. Daarnaast beschikt de overheid niet altijd over de benodigde informatie om de omvang van de externe effecten goed in te kunnen schatten.

¹⁵Dit voorstel komt o.a. van Tim Berners-Lee ([link](#)).

¹⁶Ook op het gebied van cybersecurity kunnen verplichtingen externe effecten verkleinen. Denk bijvoorbeeld aan eisen voor unieke wachtwoorden bij Internet of Things-apparaten en verplichte versleuteling voor netwerkverkeer.

¹⁷Een voorbeeld van het verbieden van een toepassing is de Genetic Information Nondiscrimination Act in de V.S. die in 2008 is aangenomen. Deze beschermt burgers tegen het gebruik van genetische informatie bij verzekeringen of in een arbeidsrelatie.

Tot slot is, in die gevallen waarbij de netto externe effecten positief zijn, het opzetten van publieke databanken een beleidsoptie. Een databank maakt data toegankelijk, zorgt ervoor dat inzichten zich breder kunnen verspreiden en maakt nieuwe toepassingen mogelijk. Op die manier kan de samenleving profiteren van de non-rivaliteit van data. Hét voorbeeld van bestaande publieke databanken is het CBS, dat data uit verschillende bronnen op geaggregeerd niveau toegankelijk maakt voor een breed publiek en, onder strikte voorwaarden, op microniveau voor wetenschappelijk onderzoek. In verschillende sectoren zijn nieuwe maatschappelijk waardevolle databanken voorstelbaar. Denk bijvoorbeeld aan data die vrijkomen bij gebruik van deelplatformen. Met gedeelde mobiliteitsdata krijgen overheden bijvoorbeeld inzicht in infrastructuurknelpunten en kunnen bedrijven innovatieve toepassingen ontwikkelen¹⁸.

Een risico bij publieke databanken is de mogelijke inbreuk op grondrechten zoals de rechten op privacy of een gelijke behandeling. Een succesvolle databank kan alleen bestaan door zorgvuldig datamanagement dat de cyberveiligheid, de kwaliteit en de privacy borgt. Het kan van strategisch belang zijn om als overheid deze investeringen te doen, bijvoorbeeld om doelstellingen op het gebied van maatschappelijke uitdagingen te realiseren.

5 Ten slotte

Digitalisering is een belangrijke bron van economische groei en biedt kansen om maatschappelijke uitdagingen op te lossen. Data zijn daarbij essentieel. De markt voor data werkt echter niet goed door externe effecten en ongelijke machtsverhoudingen. Het gevolg hiervan is dat soms te veel en soms te weinig data gedeeld worden. Het beleidskader en de beschreven oplossingsrichtingen in deze *Policy Brief* zijn niet het definitieve antwoord op hoe de data-economie te ordenen. Er is nog veel onzeker, bijvoorbeeld over onder welke omstandigheden collectieve data-overeenkomsten een oplossing bieden of over welke type data meer gedeeld zou moeten worden. Nieuwe inzichten ontwikkelen zich gaandeweg. De snelheid en impact van de digitale transformatie nopen daarbij tot een solide kennisbasis en wendbaarheid.

¹⁸ oneTransport is hier een voorbeeld van in het VK ([link](#)).

Referenties

- Acemoglu, D., A. Makhdoumi, A. Malekian, en A. Ozdaglar, 2019. Too much data: Prices and inefficiencies in data markets, NBER working paper 26296.
- Acquisti, A., L. Brandimarte, en G. Loewenstein, G., 2015, Privacy and human behavior in the age of information, *Science*, vol. 347, nr. 6221, pag. 509-514.
- Carrière-Swallow, M.Y. en M.V. Haksar, 2019, The economics and implications of data: an integrated perspective, Internationaal Monetair Fonds (IMF).
- Condorelli, D. en J. Padilla, 2020, Harnessing platform envelopment in the digital world, *Journal of Competition Law & Economics*, vol 16, nr. 2, pag. 143-187.
- Crémer, J., Y.A. de Montjoye en H. Schweitzer, 2019, Competition policy for the digital era, Rapport voor de Europese Commissie.
- CPB, Risicorapportage Cyberveiligheid Economie, 2016.
- CPB, Risicorapportage Cyberveiligheid Economie, 2018.
- EZK, 2019, De Nederlandse visie op datadeling tussen bedrijven.
- EZK, 2019-2, Discussienotitie: toekomstbestendigheid mededingingsbeleid in relatie tot online platforms.
- Jones, C. I. en C. Tonetti, 2020, Nonrivalry and the economics of data, *American Economic Review*, vol. 110, nr. 9, pag. 2819-58.
- MacCarthy, M, 2010, New Directions in Privacy: Disclosure, Unfairness and Externalities, *I/S: A Journal of Law and Policy for the Information Society*, vol. 6, pag. 425-512.
- Miller, A.C., Obermeyer, Z. en Mullainathan, S., 2019. A comparison of patient history-and ekg-based cardiac risk scores, *AMIA Summits on Translational Science Proceedings*, 2019, p.82.
- Posner E.A. en E. Glen Weyl, 2018, *Radical Markets*, Princeton University Press.
- Zuboff, S., 2015, Big other: surveillance capitalism and the prospects of an information civilization, *Journal of Information Technology*, vol. 30, nr. 1, pag. 75-89.
- Zuboff, S., 2019, *The age of surveillance capitalism*, Profile Books.