

Opzet van een datavoorzieningsfunctie ter ondersteuning van datagedreven toezicht

Wouter van Aerle, Ronald Damhof en Frank Ouddeken*

Toezichthouders hebben te maken met een steeds grotere stroom aan data. Dit vraagt om toenemende discipline in de verwerking ervan. Uit de principes van goed toezicht volgen niet-functionele eisen – als traceerbaarheid, uitwisselbaarheid en betrouwbaarheid – waaraan de gegevensverwerking moet voldoen. Voor het realiseren hiervan is een datavoorzieningsfunctie nodig, vergelijkbaar met andere bedrijfsfuncties als Finance en HR. Binnen DNB is het datakwadrantenmodel gebruikt om deze realisatie vorm te geven. Het model is een begripsvormend kader, dat de gegevensverwerking vanuit twee perspectieven beschouwt: de aanbod- en vraagzijde van gegevens en een opportunistische versus een systematische manier van werken. Op basis van het model zijn keuzes gemaakt ten aanzien van de technische voortbrenging van data, inrichting van de organisatie, besturing en benodigde kennis en competenties.

Inleiding

De opkomst en het belang van data zijn overduidelijk zichtbaar in de samenleving. Naast nieuwe mogelijkheden worden ook risico's en kwetsbaarheden openlijk zichtbaar, zoals recent bleek uit het datalek bij Face-

book.¹ Toezichthouders staan voor soortgelijke uitdagingen. Uit de praktijk wordt duidelijk dat het gebruik van data structurele ondersteuning vergt. Gebaseerd op ervaringen bij De Nederlandsche Bank (DNB) en gespiegeld met ervaringen bij andere toezichthouders zetten we in dit artikel uiteen waarom deze ondersteuning vereist is en wat de overwegingen bij de inrichting ervan zijn. Dat begint bij een nadere duiding van wat data en de groei hiervan betekenen. Vanuit dit begrip beschrijven we de kaders waarbinnen de toezichthouder de gegevensverwerking dient in te richten. Met het kwadrantenmodel introduceren we vervolgens een begripsvormend raamwerk waarmee de inrichting van een datavoorzieningsfunctie gestuurd kan worden. We sluiten af met een beschouwing hoe de inrichting bij DNB concreet is vormgegeven.

Dataficatie: exponentiële groei van gegevens

Het verzamelen van informatie is voor toezicht een kernactiviteit.² Verzamelde informatie vormt de basis voor oordeelsvorming door de toezichthouder en eventueel daaropvolgend ingrijpen. Met de toenemende groei van gegevens³ – vaak aangeduid met de term *dataficatie* – ontstaan nieuwe informatiebronnen die voor toezicht van belang kunnen zijn. Dataficatie wordt door

* Ir. W.A. van Aerle is managing partner bij Deltiq, adviesbureau voor informatiemanagement. Drs. R.D. Damhof was tot 1 juli 2018 werkzaam als enterprise data architect bij De Nederlandsche Bank. Drs. F.E.M. Ouddeken is werkzaam voor De Nederlandsche Bank als manager Data Management Office, Divisie Statistiek. De auteurs schrijven op persoonlijke titel.

1. 'Data van Facebook misbruikt door firma die voor Trump werkte', www.nrc.nl/nieuws/2018/03/19/data-van-facebook-misbruikt-door-firma-die-voor-trump-werkte-a1596172.
2. R. Velders, *Introductie in toezicht en handhaving*, Den Haag: Boom Lemma uitgevers 2011.
3. De termen *gegevens* en *data* worden beschouwd als synoniem en worden in dit artikel door elkaar heen gebruikt.

Mayer-Schonberger en Cukier⁴ gedefinieerd als het verschijnsel ‘to put a phenomenon in a quantified format so it can be tabulated and analyzed’. In ruimere zin wordt de term dataficatie vaak gebruikt om de toename van data in de maatschappij te duiden. Feitelijk is dit het gevolg van dataficatie: ‘through datafication (...) we can now capture and calculate at a much more comprehensive scale the physical and intangible aspects of existence (...)’.⁵

Deze groei van gegevens kent verschillende verschijningsvormen. Voortdurende digitalisering zorgt voor de overgang van analoge naar digitale vastlegging van informatie over bedrijfsactiviteiten. De Boordcomputer Taxi (BCT)⁶ bijvoorbeeld vervangt de papieren rittenadministratie van taxi’s, wat de toezichtlast vermindert en de toezichthouder (ILT) in staat stelt gericht toezicht te houden. Daarnaast ontstaan nieuwe aanbieders van producten en diensten die vrijwel geheel digitaal opereren zoals Bunq (bank), Adyen (betaaldienstverlener) of Brand New Day (pensioenbeleggen). Een andere relevante trend is het internet der dingen (Internet of Things of IoT): de ontwikkeling dat steeds meer alledaagse voorwerpen verbonden worden met het internet en gegevens ontvangen en versturen over hun functioneren en gedrag. Voorbeelden zijn slimme energiemeters, vaatwassers en lampen, maar ook auto’s, medische apparatuur, straatverlichting, vuilnisbakken of agribots. IoT is enerzijds een verdergaande vorm van digitalisering en de data die op deze manier worden gegenereerd zijn een potentiële informatiebron voor toezicht. Anderzijds brengt het IoT ook nieuwe toezichtvraagstukken met zich mee, zoals het rechtmatig gebruik van persoonsgegevens of de beveiliging van apparaten. Sociale media ten slotte worden hier beschouwd als het verschijnsel dat burgers en bedrijven vrijwillig data delen en beschikbaar stellen. Hiermee ontstaat voor toezichthouders een nieuwe bron om signalen uit de praktijk te krijgen. Berichten op Twitter bijvoorbeeld of discussies op consumentenfora als Radar vormen zo een additionele informatiebron om indicaties te krijgen van mogelijke gedragingen die strijdig zijn met wet- en regelgeving.

Gevolgen voor toezicht

De voortdurende groei in de beschikbaarheid en gebruiksmogelijkheden van data heeft indringende gevolgen voor toezicht. Niet alleen verandert de externe omgeving waarbinnen de toezichthouder zelf opereert en de wijze waarop de onder toezicht staande instellingen daarbinnen invulling geven aan hun activiteiten, ook de verwachtingen van het publiek over wat het als goed

toezicht beschouwt, wijzigen hierdoor. Voor de toezichthouder zelf vertaalt zich dit in nieuwe kansen en mogelijkheden, een hoger verwachtingsniveau bij de buitenwacht ten aanzien van toezicht en daarmee in grotere risico’s en verantwoordelijkheden voor de toezichthouder.

Kansen en mogelijkheden

Traditioneel beschikken toezichthouders over wettelijke bevoegdheden om bij de onder hun toezicht staande instellingen noodzakelijke informatie op te vragen. Bij prudentiële toezichthouders is daarbij veelal sprake van reguliere, gestandaardiseerde rapportages. Hoewel de traditionele informatievoorsprong van toezichthouders door de toenemende dataficatie in het gedrang leek te komen, zijn verplichte rapportages de laatste jaren als gevolg van – veelal Europese regelgeving⁷ – sterk in omvang gegroeid. Steeds vaker is hierbij sprake van *granulaire datasets*⁸ die op periodieke basis beschikbaar moeten worden gesteld door financiële instellingen.⁹ De reikwijdte van deze verplichte rapportages¹⁰ neemt hierbij steeds vaker toe, vooral bij het markt- c.q. gedrags-toezicht. Zo is na het begin van de financiële crisis de verplichting tot afwikkeling van transacties via centrale tegenpartijen substantieel vergroot en hebben de markt-toezichthouders de beschikking gekregen over data met betrekking tot individuele beleggings- en derivaten-transacties, inclusief het betrokken onderpand.¹¹ Ook is hierdoor meer en gedetailleerdere informatie beschikbaar gekomen die kansen biedt voor versterking van het toezicht.¹²

Zo is in diverse rapportages het gebruik van de internationaal overeengekomen Legal Entity Identifier (LEI)¹³ verplicht gesteld. De LEI-code komt terug als identificerend sleutel in toezichtrapportages voor banken (CRD IV) maar is bijvoorbeeld ook verplicht onder EMIR (derivaten-transacties) en AnaCredit (granulaire leningsposities en onderpand, zie verderop). Het gebruik van de LEI-code biedt zo de mogelijkheid om deze datasets eenvoudig(er) te combineren.¹⁴ Op deze combinatie van gegevens kan met behulp van algoritmes een risico-

4. V. Mayer-Schonberger en K. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, New York: Houghton Mifflin Harcourt 2015.
5. Mayer-Schonberger en Cukier 2015, p. 96.
6. De BCT is sinds 2012 verplicht en daarmee een voorbeeld van dataficatie die al enkele jaren geleden is ingezet, www.ilent.nl/onderwerpen/boordcomputer-taxi.

7. Bijvoorbeeld EMIR, MifID II, SFTR en AnaCredit.
8. Met granulariteit wordt het detailniveau van data de bedoeld. Granulaire (d.w.z. gedetailleerde) data zijn het tegenovergestelde van geaggregeerde (d.w.z. samengevatte) data.
9. D.M. Bholat, ‘The future of central bank data’, *Journal of Banking Regulation* 2013, nr. 3-4, p. 185-194.
10. Bijvoorbeeld de prudentiële rapportages van banken en verzekeraars (CRDIV resp. Solvency II).
11. Voorbeelden hiervan zijn de MifID/MifIR II en EMIR.
12. Zo wordt uit studies van de European Systemic Risk Board (ESRB) duidelijk dat granulaire data betere mogelijkheden bieden om bijvoorbeeld schaduwbankieren in kaart te brengen.
13. De LEI-code is een internationaal toegepast uniek nummer bedoeld om juridische entiteiten die participeren in financiële transacties, eenduidig te identificeren, www.gleif.org/en/about-lei/introducing-the-legal-entity-identifier-lei.
14. Europese ontwikkelingen als het European Reporting Framework (ERF) en Banks Integrated Reporting Dictionary (BIRD) ten aanzien van het harmoniseren van data over verordeningen heen passen in dit kader, www.ecb.europa.eu/stats/ecb_statistics/co-operation_and_standards/reporting/html/index.en.html.

Tabel 1 Principes van goed toezicht

OECD Best Practice Principles for Regulatory Policy	Zes principes van goed toezicht
1. Evidence-based enforcement	a. selectief toezicht
2. Selectivity	b. slagvaardig toezicht
3. Risk focus and proportionality	c. samenwerkend toezicht
4. Responsive regulation	d. onafhankelijk toezicht
5. Long-term vision	e. transparant toezicht
6. Co-ordination and consolidation	f. professioneel toezicht
7. Transparent governance	
8. Information integration	
9. Clear and fair process	
10. Compliance promotion	
11. Professionalism	

profiel opgebouwd worden om vroegtijdig signaleringen te doen.

Bij het bancaire toezicht in Europa geldt dat door de centralisatie in SSM¹⁵-verband en de onderbrenging ervan bij de ECB ook monetaire statistische datasets beschikbaar zijn gekomen voor het toezicht. Ook deze statistische rapportages worden steeds meer granular, zoals de effectenstatistiek van het ESCB¹⁶ die het houderschap van individuele effecten meet en de Ana-Credit¹⁷-verordening. Deze verordening leidt binnen afzienbare tijd tot een register van individuele bancaire bedrijfskredieten op Europees niveau. Al met al beschikken toezichthouders daarmee in toenemende mate over een unieke en steeds rijkere informatiepositie. De mogelijkheden om van de informatie die in deze databronnen zit gebruik te maken voor het toezicht nemen snel toe. Analysetechnieken nemen een steeds hogere vlucht. In eerste instantie werden data daarbij op gestructureerde wijze toegankelijk gemaakt om bestaande toezichtprocessen beter te ondersteunen met behulp van ‘dashboards’. Tegenwoordig wordt steeds meer gebruikgemaakt van algoritmes die voorspellende eigenschappen hebben.¹⁸ Een verdere doorontwikkeling daarvan zijn analytische technieken zoals vormen van kunstmatige intelligentie. Hierdoor voltrekt zich een transformatie in de toezichtprocessen die een transitie naar steeds meer *datagedreven toezicht* bewerkstelligt. Datagedreven toezicht wordt gekenmerkt door het systematisch inzetten van databronnen en evaluatie van de effectiviteit van daarop gebaseerde acties van de toezichthouder. Op basis van ervaring in diverse organisaties zijn de volgende effecten zichtbaar:

- efficiëntere inzet van schaarse toezichtcapaciteit door betere risicoanalyses;

- monitoring van hele toezichtpopulaties en netwerken in plaats van steekproeven;
- verbeterde onderbouwing van standpunten en besluiten op basis van feiten, bijvoorbeeld bij thematisch onderzoek;
- verschuiving van reactief naar meer preventief toezicht door meer gebruik te maken van voorspellende en zelflerende analysetechnieken zoals vormen van kunstmatige intelligentie.

Hoger verwachtingsniveau

De omvang, snelheid en gedetailleerdheid waarmee informatie tegenwoordig publiek beschikbaar komt en bijvoorbeeld via sociale media wordt verspreid, leidt ertoe dat het publiek steeds sneller is geïnformeerd of meent dat te zijn. Hierdoor ontstaat gemakkelijk de verwachting in de maatschappij dat de toezichthouder ook altijd van alles op de hoogte is en alles tijdig voorziet. Dit kan in de praktijk gemakkelijk leiden tot de opvatting dat de toezichthouder ‘het had kunnen weten’. Hoe minder data onbekend of niet toegankelijk zijn, des te hoger zullen de publieke verwachtingen zijn ten aanzien van wat de toezichthouder vermag te bereiken. Kortom, dataficatie schept ook hoge verwachtingen ten aanzien van de toezichthouder.

Impact voor toezicht

De toenemende beschikbaarheid en daarmee gepaard gaande gebruiksmogelijkheden van data brengen ook grotere verantwoordelijkheden én risico’s met zich mee. Deze verantwoordelijkheden en risico’s manifesteren zich primair in de wijze waarop gegevens door de toezichthouder worden verwerkt¹⁹ ter ondersteuning van zijn toezichtstaken.

15. Single Supervisory Mechanism: de wijze waarop het banktoezicht in Europa is vormgegeven. Zie ook: www.bankingsupervision.europa.eu/about/thessm/html/index.en.html.
 16. Securities Holdings Statistics (SHS).
 17. www.ecb.europa.eu/stats/money_credit_banking/anacredit/html/index.en.html.
 18. Zie bijvoorbeeld ‘Tekst van een jaarverslag zegt meer over fraude dan gedacht’, *Het Financieele Dagblad* 29 december 2017.

19. Onder het *verwerken* van gegevens vallen handelingen als het verzamelen, vastleggen, opslaan, wijzigen, opvragen, raadplegen, gebruiken, verstrekken, wissen en vernietigen van data.

De principes van goed toezicht²⁰ kunnen worden beschouwd als een algemene leidraad voor toezichthouders voor de vormgeving van hun toezicht (zie Tabel 1). Het ligt voor de hand om te stellen dat deze principes niet alleen van toepassing zijn op het toezichtproces zelf maar ook op hiervoor benodigde (geautomatiseerde) gegevensverwerking.

De implicaties van deze principes worden uitgedrukt in *niet-functionele eisen*. Dit soort eisen specificereert criteria om het functioneren van de gegevensverwerking te beoordelen maar beschrijft geen specifieke gegevensverwerking zelf, zoals de selectie van specifieke databronnen ten behoeve van analyse in het kader van thematisch onderzoek. De principes van de OECD en die van het ministerie liggen in lijn met elkaar en worden hieronder gecombineerd beschouwd in de context van datagedreven toezicht.

Bewijs en feiten

Het eerste principe van de OECD gaat over *evidence-based enforcement*: ‘deciding what to inspect and how it should be grounded on data and evidence’. Het tweede principe van het ministerie is slagvaardig toezicht: ‘zakelijke houding gebaseerd op feiten’.

Het belang van data komt in deze principes expliciet naar voren en onderstreept de noodzaak voor toezichthouders om een informatiepositie op te bouwen. De niet-functionele eisen die hiermee samenhangen zijn *relevantie* van de data voor een specifiek toezichtvraagstuk. Omgekeerd dient *misbruik of onjuist gebruik* van data, zoals een verkeerde duiding van de data, te worden voorkomen. Een tweede belangrijke eis is de *betrouwbaarheid* van data. Het waarborgen van de *kwaliteit* van data en de analyses waarin deze data worden gebruikt is van groot belang, zodra data ten grondslag gaan liggen aan toezichtmaatregelen. Dit vereist dat de toezichthouder ‘in control’ is over de kwaliteit van de gegevens. Dit impliceert niet direct dat (alle) data altijd van de hoogste kwaliteit moeten zijn alswel dat er inzicht is in wát de kwaliteit van de data is, en de toezichthouder op grond van deze kennis een weloverwogen beslissing neemt een dataset al dan niet te gebruiken.

Daarnaast dienen de gegevens ook daadwerkelijk voor de toezichthouder *toegankelijk* te zijn. Dit lijkt triviaal maar in de praktijk blijkt vaak dat data opgesloten zitten in systemen en slecht benaderbaar zijn voor gebruikers.

Selectief en risicogebaseerd

Selectiviteit is nodig omdat ‘inspections and enforcement cannot be everywhere and address everything’.²¹ Het is dus nodig om keuzes te maken. Voor het maken van deze keuzes staat een risico-oriëntatie voorop: ‘enforcement needs to be risk-based and proportionate’.²²

20. OECD, *Regulatory Enforcement and Inspections*, Parijs: OECD Publishing 2014; ministerie van Binnenlandse Zaken en Koninkrijksrelaties, *Minder last, meer effect. Zes principes van goed toezicht*, 2005.

21. OECD 2014, p. 24.

22. OECD 2014, p. 27.

Het maken van keuzes is bij uitstek een beslissing die in hoge mate met data en analyse ondersteund kan worden. Door algoritmes en analytische modellen te gebruiken kunnen risico’s in kaart worden gebracht en beoordeeld. Deze toepassing stelt eisen aan de *bruikbaarheid* en *betrouwbaarheid* van (beschikbare) data, gewenste *wendbaarheid* om gemakkelijk nieuwe databronnen te ontsluiten en *traceerbaarheid* van de gegevensverwerking om de uitkomsten van risico-analyses en daarop gebaseerde keuzes te kunnen verantwoorden.

Samenwerking

Van toezichthouders wordt verwacht dat ze waar mogelijk samenwerken, beschikbare gegevens delen en zo de rapportagelast (*reporting burden*) voor onder toezicht staande instellingen beheersen. Dit is verwoord in het principe van ‘co-ordination and consolidation [between] inspection functions’²³ en het principe van samenwerkend toezicht: ‘verschillende toezichthouders zullen niet dezelfde informatie vragen’.²⁴

Dit principe vraagt om een efficiënte informatiehuishouding en het hergebruik van gegevensverzamelingen die met een ander maar verenigbaar doel door andere toezichthouders zijn verzameld bij onder toezicht staande instellingen. Tussen de AFM en DNB is dit bijvoorbeeld geformaliseerd in een samenwerkingsconvenant.²⁵ Bestuurlijke afspraken zijn een randvoorwaarde om daadwerkelijke gegevensuitwisseling in de praktijk te realiseren. Dit stelt eisen aan de *interoperabiliteit* (dat wil zeggen de mogelijkheid om elkaars gegevens rechtstreeks te kunnen gebruiken en te koppelen) van data. Het gebruik van open standaarden en basisregistraties bevorderen deze uitwisselbaarheid.

Transparantie

Transparant toezicht behelst het verantwoording kunnen afleggen over gemaakte keuzes en behaalde resultaten. Om transparant te kunnen zijn moet inzicht gegeven kunnen worden in de wijze waarop de data zijn verwerkt die als basis hebben gediend voor gemaakte keuzes en behaalde resultaten representeren. Dit betekent dat de gegevensverwerking *traceerbaar* dient te zijn. Traceerbaarheid is de mogelijkheid om de historie, oorsprong en verwerking van data te verifiëren op basis van specifiek hiervoor geregistreerde meta-informatie.²⁶

Informatie-integratie

Onder informatie-integratie verstaat de OECD ‘[to] maximise information-sharing and optimal use of resources’.²⁷ Dit sluit aan bij het principe van coördina-

23. OECD 2014, p. 41.

24. Ministerie van Binnenlandse Zaken en Koninkrijksrelaties 2005, p. 22.

25. ‘Rectificatie Samenwerkingsconvenant tussen de Stichting Autoriteit Financiële Markten en De Nederlandsche Bank N.V.’, *Stcr*. 2016, 1863.

26. Traceerbaarheid is een voorwaarde voor *auditeerbaarheid*. Auditing is het uitvoeren van onderzoek naar opzet en werking van een proces of het resultaat dat dit proces voortbrengt. Een dergelijk onderzoek wordt ondersteund of kan soms alleen maar worden gerealiseerd door het bijhouden van een *audit trail*. De meta-informatie die ten behoeve van traceerbaarheid moet worden vastgelegd, kan hiervoor worden gebruikt.

27. OECD 2014, p. 51.

tie en samenwerking. Maar niet alleen het (kunnen) delen van informatie tussen instellingen speelt hier een rol, ook – of misschien wel juist – het delen van informatie binnen de toezichthouder zelf is hier van belang. Het is niet vanzelfsprekend dat gegevens die eenmaal zijn uitgevraagd, vrijelijk beschikbaar zijn voor andere belanghebbenden binnen de toezichtorganisatie. Niet zelden zijn er organisatorische, culturele of technische beperkingen die het delen van data binnen de organisatiegrenzen belemmeren. Het bevorderen van de *toegankelijkheid* van gegevens dient daarom expliciet aandacht te krijgen.

Naast de niet-functionele eisen die zijn afgeleid van de principes voor goed toezicht, zijn nog andere eisen van toepassing. Zo dient in het geval van persoonsgegevens de verwerking te voldoen aan de Algemene verordening gegevensbescherming (AVG). Ook het waarborgen van de *vertrouwelijkheid* van data, in de zin van beperking tot slechts gelegitimeerd gebruik, verscherpt de verantwoordelijkheden voor de toezichthouder in hoge mate.

Noodzaak van een datavoorziening

Uit voorgaande beschouwing is duidelijk geworden dat de opkomst van data een verscheidenheid aan kansen, verantwoordelijkheden, risico's en verwachtingen met zich mee brengt. Hoe evident de eisen aan de gegevensverwerking ook lijken te zijn, een juiste toepassing en beheersing ervan blijkt in de praktijk geen sinecure maar diep in te grijpen in de bedrijfsvoering van een toezichthouder. Duidelijk is dat het benutten van het datapotentieel waarbij tegelijkertijd aan externe verwachtingen en niet-functionele eisen wordt voldaan meer vergt dan de implementatie van nieuwe technologie of het aannemen van een aantal data-analisten. Het vereist daarentegen een datavoorziening²⁸ die de organisatorische *capability*²⁹ levert om systematisch data te benutten binnen de gestelde randvoorwaarden.

De vraag waarvoor DNB zich gesteld zag, is hoe gegeven genoemde ontwikkelingen in de wereld van data zo'n datavoorziening ingericht en vormgegeven moet worden. Dit inrichtingsvraagstuk staat centraal in de rest van dit artikel.

Centraal hierbij staat de balans tussen de benodigde organisatie en *governance* van zo'n datavoorziening enerzijds en de valorisatie (dat wil zeggen het benutten van waarde) van data anderzijds. Een lastige balans omdat de twee uitersten elkaar lijken tegen te spreken. Bij te veel organisatie en *governance* worden innovatie en experimenteren met data gemiddeld; te veel nadruk op valorisatie

kan ten koste gaan van niet-functionele eisen zoals interoperabiliteit, transparantie of vertrouwelijkheid.

Bij het inrichten van een datavoorziening – hoe gaan wij als toezichthouder gegevens verwerken – en het bepalen van de hiervoor genoemde balans zijn er drie uitdagingen. Ten eerste het realiseren van een voorziening waar innovaties met data op een verantwoorde manier kunnen plaatsvinden en waarbij de organisatie het vermogen³⁰ heeft ontwikkeld om deze innovaties structureel te operationaliseren binnen de toezichtprocessen. Ten tweede het belang om bij de realisatie van zo'n voorziening de eerdergenoemde niet-functionele eisen *by design*³¹ te implementeren. Ten slotte een houding binnen de organisatie die data ten principale ziet als grondstof die gedeeld moet worden binnen en buiten de organisatie, rekening houdend met wettelijke kaders. Een dergelijke houding is randvoorwaardelijk omdat data zich in de praktijk niet houden aan kunstmatige kaders zoals informatiesystemen, bedrijfsprocessen, organisatie-eenheden en zelfs rechtspersonen maar daar als het ware doorheen vloeien.³²

Visie op de voortbrenging van gegevens: kwadrantenmodel

Om het bestuurlijk en management kader in de uitwerking van het inrichtingsvraagstuk voor een datavoorziening handvatten te geven, gebruikt DNB het Kwadrantenmodel,³³ een begripvormend raamwerk rondom communicatie, organisatie, *governance*, beleid, investeringen, architectuur en prioriteit op het datadomein. Het is een eenvoudig model dat veel, met name publieke organisaties³⁴ in hun benadering van dit vraagstuk hebben overgenomen of dat door hen wordt toegepast.

28. Onder een voorziening verstaan we het geheel van beleid, processen, mensen en (technologische) middelen, ondergebracht in een organisatorische functie.

29. 'a firm's ability to perform repeatedly a productive task which relates (...) to a firm's capacity for creating value', R.M. Grant, 'Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration', *Organization Science* 1996, nr. 4, p. 359-467.

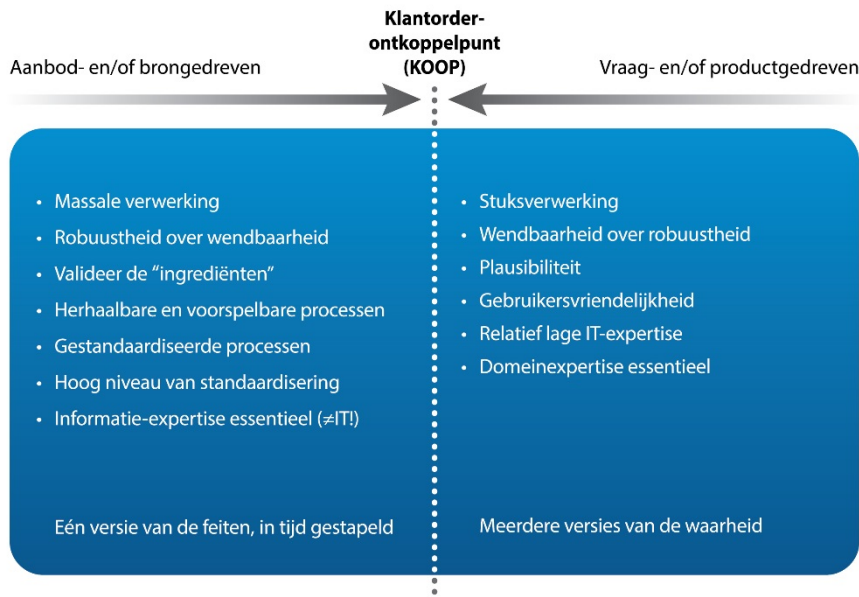
30. Sparrow spreekt in deze over 'the task of devising and constructing the organizational infrastructure necessary to turn the concept of a coherent compliancy strategy into an operational reality', M.K. Sparrow, *The Regulatory Craft*, Washington D.C.: Brookings Institution Press 2000.

31. Het 'by design'-principe is afgeleid van het specifiekere Privacy by Design-principe. Refererend aan art. 25 AVG wordt dit principe gedefinieerd als het treffen van passende technische en organisatorische maatregelen zowel bij de bepaling van de verwerkingsmiddelen als bij de verwerking zelf met als doel de gegevensbeschermingsbeginselen (...) op een doeltreffende manier uit te voeren en de nodige waarborgen in de verwerking in te bouwen ter naleving van de voorschriften (...). Deze manier van borging kan op vergelijkbare manier ook worden toegepast ter naleving van andere niet-functionele vereisten, zoals de kwaliteit(seisen) van data en de betekenis c.q. definities.

32. 'The wondrous and perilous properties of data and information in organizations', in: Th.C. Redman, *Data Driven*, Boston, Mass.: Harvard Business Press 2008.

33. Dit model is geïnspireerd op (1) The Toyota Way die haar oorsprong heeft in het werk van W. Edwards Deming, (2) het Cynefin Framework van Dave Snowden en (3) ontwikkelingen in data-architectuur van de afgelopen tien jaar met name op het gebied van feitmodellering, integratiemodellering en het ondersteunen van multi-realiteiten. 'Maak datamanagement bespreekbaar in de hele organisatie', in: *Future Bright. A data driven reality*, SAS Nederland 2015, p. 24-31, https://issuu.com/sasnederland/docs/futurebrightdatamanagement_nl.

34. Voorbeelden zijn de Autoriteit Financiële Markten (AFM), de gemeente Amsterdam en de Nederlandse Voedsel- & Warenautoriteit (NVWA).



70

Het model is geïnspireerd door de wijze waarop ruwe grondstoffen in de fysieke wereld worden verwerkt tot eindproducten. Eenzelfde uitdaging is er in de wereld van data. DNB ontvangt van verschillende partijen binnen en buiten haar organisatie steeds meer data. Van deze data (lees: ruwe grondstoffen) moeten verschillende informatieproducten worden gemaakt. Dit varieert van gebruikelijke informatieproducten als rapportages en dashboards om besluitvorming te verbeteren, datasets voor geavanceerde analyses, maar ook verplichte dataleveringen aan partijen binnen en buiten DNB. Meer ambitieus zijn informatieproducten die het resultaat zijn van algoritmes op data die een bedrijfsproces aansturen, bijvoorbeeld een verhoogd risico op een rechtspersoon waar nader onderzoek nodig is. De kern is dat data binnenkomen als ruwe grondstof en zodanig getransporteerd en verwerkt moeten worden dat de vele verschillende soorten afnemers er hun gevarieerde doelstellingen mee kunnen bereiken.

In deze logistieke beschouwing van gegevensvoortbrenging zijn er twee assen die vier kwadranten opleveren. Deze assen zijn het klantorderontkoppelpunt en de wijze van ontwikkeling.

Klantorderontkoppelpunt

In het proces waarin ruwe grondstoffen verwerkt worden tot eindproducten is het concept *klantorderontkoppelpunt (KOOP)* cruciaal: tot waar dringt de vraag van de afnemer door in het verwerkingsproces? Bij het kopen van een luxe motorjacht zal er pas geproduceerd gaan worden als de order van de klant geaccordeerd is (vraaggedreven). Het KOOP ligt in dat geval bij de ruwe grondstoffen, dus bij de start van het verwerkingsproces. Bij het kopen van lucifers daarentegen gaat de afnemer naar de supermarkt, haalt ze van het schap en rekent af (aanbodgedreven). Het KOOP ligt hier bij het eindproduct. Het KOOP wordt zo onderscheiden door

aanbodgedreven (*push*) en vraaggedreven (*pull*) zijde. Kenmerken van deze twee kanten zijn onder andere:

- *push* is aanbodgedreven, *pull* is vraaggedreven;
- *push* is gericht op massaverwerking, *pull* is gericht op stukverwerking;
- *push* is herhaalbaar, voorspelbaar en doorloopt een uniform proces;
- *push* stelt betrouwbaarheid boven wendbaarheid, *pull* stelt wendbaarheid boven betrouwbaarheid;
- *push* kent een grote mate van specificatie en standaardisatie, *pull* heeft dat per definitie minder;
- *push* kent een hoge automatiseringsgraad, *pull* een lagere.

In het verwerken van data wordt bovenstaande analogie overgenomen en ontstaan er twee gebieden die fundamenteel van elkaar verschillen (zie Figuur 1).

Aan de aanbodzijde (*push*, linkerkant) is er één versie van *feiten*. Vanuit het perspectief van de organisatie worden alle gegevens zoals aangeleverd door rapporteurs, technische systemen of andere instituties, beschouwd als feiten. Feiten kennen maar één versie, de versie zoals die is ontvangen, ongeacht eventuele onvolkomenheden die de data bevatten. Er is geen interpretatie nodig om deze gegevens op te slaan en te verwerken. Deze feiten zijn over de tijd heen gestapeld en worden in principe³⁵ niet verwijderd. Het met de data door de tijd reizen wordt daarmee mogelijk gemaakt. Allemaal niet-functionele eisen die passen in de principes van goed toezicht.

Aan de rechterkant (*pull*, vraagzijde) zijn er meerdere versies van de waarheid. Waarheden, ook wel realiteiten genoemd, zijn perspectieven op feiten. Verschillende gebruikers kunnen op basis van dezelfde feiten verschillende perspectieven hebben die allemaal binnen hun specifieke context geldig en juist zijn. Waarheden zijn

35. Tenzij specifieke wetgeving zoals bijvoorbeeld de AVG dit verplicht.

als het ware lenzen op de werkelijkheid. Met lenzen van verschillende sterkten en kleuren krijgen afnemers een beeld van de werkelijkheid die past binnen hun context. Dit kan bij DNB een toezichtperspectief zijn op banken, een macrostatistisch perspectief op basis van de betalingsbalans of een perspectief vanuit financiële stabiliteit.

Wijze van ontwikkelen

Een deel van de informatieproducten die gemaakt moeten worden, kent redelijk vastomlijnde eisen. Denk aan managementinformatie met betrekking tot bedrijfsvoering of het periodiek leveren van datasets aan de ECB, het CBS, EuroStat en het IMF. Het maken van deze informatieproducten kent een verloop dat te vergelijken is met ordentelijke systeemontwikkeling in de informatietechnologie. We noemen dit ook wel *systematisch ontwikkelen* (zie Figuur 2). Binnen DNB maar zeker ook daarbuiten is echter sprake van een niet te stoppen en potentieel waardeverhogende trend om innovatief en experimenteel om te kunnen gaan met data, vaak onder de noemer van *data science*. Onderzoekers, econometristen, statistici of andere kwantitatieve specialisten zoeken naar patronen, ‘*outliers*’ en inzichten die verstopt zijn in complexe en grote hoeveelheden data en die een indicatie geven waar nader onderzoek nodig is. De context waarin deze experts werken, moet een grote mate van vrijheid kennen. We noemen dit ook wel *opportunistisch ontwikkelen*.

Het onderscheid tussen deze twee vormen van ontwikkelen is cruciaal voor het datadomein omdat de twee vormen een fundamenteel andere inrichting en operationalisatie behoeven. Hieronder benoemen we enkele onderscheidende kenmerken:

- Er is sprake van functiescheiding tussen gebruiker en ontwikkelaar in de systematische ontwikkeling. Deze is niet of beperkt aanwezig bij opportunistische ontwikkeling. Hier is degene die de data verzamelt, duidt, prepareert, schoont, analyseert, interpreteert en er soms zelfs actie op onderneemt, een en dezelfde persoon.
- Systematische ontwikkeling wordt gekenmerkt door een defensieve *governance*, in tegenstelling tot de opportunistische ontwikkeling waar sprake is van offensieve *governance*. Defensieve *governance* betekent een planmatige, vastomlijnde en geformaliseerde manier van ontwikkelen, implementeren en beheren van gegevensverwerkingsprocessen. Denk hierbij aan het inrichten van systeemmonitoring, wijzigingsmanagement, OTAP³⁶-voortbrenging, segmentering van IT-omgevingen enzovoort. Bij een offensieve *governance* is dit veelal niet aanwezig en ook niet nodig. De beheersing van de gegevensverwerking is hier de volledige verantwoordelijkheid van de degene die de verwerking uitvoert.

- Systematische ontwikkeling tendert naar een centrale inrichting binnen een organisatie, in tegenstelling tot de opportunistische wijze van ontwikkeling die een grote mate van decentralisatie (soms tot de individuele onderzoeker) kent.
- De systematische wijze van ontwikkelen vormt zich rondom de niet-functionele eisen als duurzaamheid, betrouwbaarheid, schaalbaarheid en herhaalbaarheid. De opportunistische wijze van ontwikkeling is meer gericht op eisen als wendbaarheid, snelheid en experimenteren.
- De technische omgeving voor systematische voortbrenging is ingericht volgens standaarden die in de informatietechnologie gewoon zijn en bedoeld zijn voor langdurig, structureel en meervoudig gebruik. In de opportunistische omgeving daarentegen worden er feitelijk drie componenten aangeboden: data, computerkracht, analyse- en programmeergereedschappen. Dit soort omgevingen wordt vaak aangeduid met termen als datalab, *sandbox* of pilot-omgevingen³⁷ en kennen doorgaans een tijdelijk karakter voor de duur van een specifieke data-analyse.

De kwadranten nader belicht

De twee assen gecombineerd geven een viertal kwadranten (zie Figuur 3). Het is belangrijk te benadrukken dat de assen waarde vrij zijn. Kwadrant I is niet beter, mooier of belangrijker dan kwadrant IV. Hoe DNB de kwadranten invult, is contextafhankelijk. Voor toezichthouders en formele rapportages naar bijvoorbeeld de ECB of EIOPA is het van belang dat de gegevensverwerking transparant en auditeerbaar is (kwadrant I). Voor onderzoekers is het opportuun dat ze eigen data (kwadrant III) kunnen combineren met systematisch voortgebrachte data (kwadrant I) op zodanige wijze dat ze met gespecialiseerde analysegereedschappen (kwadrant IV) kunnen experimenteren.

• Kwadrant I: Feiten

In kwadrant I staan de gegevens zoals ze bij aanlevering zijn aangeboden. Er wordt gestreefd naar een minimale interpretatie van gegevens en maximale precisie en ondubbelzinnigheid. De automatiseringsgraad en de mate van beheersbaarheid zijn in dit kwadrant maximaal. Bij het doorvoeren van wijzigingen in de gegevensverwerking zal beheersbaarheid hoger worden gewaardeerd dan wendbaarheid. In de context van bijvoorbeeld DNB worden in kwadrant I hoge eisen gesteld aan de verwerking, met name ten aanzien van traceerbaarheid, beheersbaarheid, schaalbaarheid en betrouwbaarheid.

• Kwadrant II: Context

In dit kwadrant worden informatieproducten gemaakt en in beheer genomen die structurele behoeften van de gebruiker vertegenwoordigen. Deze informatieproducten

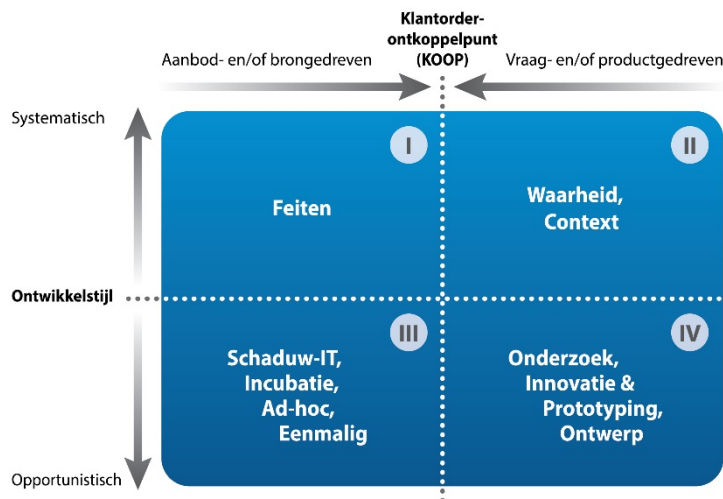
36. Ontwikkel, Test, Acceptatie en Productie is een gebruikelijke fasering die bij de ontwikkeling van informatiesystemen wordt doorlopen.

37. Zie bijvoorbeeld www.toezine.nl/artikel/228/ilt-verbetert-datagedreven-toezicht-met-datalab/.

Figuur 2 Wijzen van ontwikkeling



Figuur 3 DNB kwadrantenmodel



ten komen tot stand door een systematisch productieproces waarbij gebruik wordt gemaakt van de feiten in kwadrant I en waarbij vooraf duidelijk is verwoord wat de kwaliteitseisen zijn en op welke wijze de informatieproducten in de bedrijfsprocessen van waarde zijn. In dit kwadrant kunnen informatieproducten worden generaliseerd die op dezelfde feiten zijn gebaseerd maar als gevolg van een verschillend perspectief op die feiten van elkaar verschillen of elkaar zelfs tegenspreken.

• *Kwadrant III: Shadow-IT, ad hoc, eenmalig*

Elke organisatie heeft een kwadrant III. Denk aan alle Excelbestanden, MS Access databases en andere lokale gegevensdragers. Kenmerk van dit soort omgevingen is dat de enige persoon die de data begrijpt, vaak ook diegene is die de data er heeft neergezet. Het is ook het kwadrant waar het mogelijk moet zijn om data neer te zetten die een eenmalig of onvoorspelbaar karakter heb-

ben. Het is hierdoor ook niet nodig om deze data in kwadrant I te plaatsen. Belangrijk is verder dat de gegevens in kwadrant III onder beperkt georganiseerd beheer staan en dat de betreffende datakwaliteit hoogstens bekend is bij degene die de data er heeft neergezet. Let wel: ook in dit kwadrant staan de data wel degelijk onder *governance*. Alleen is het de gebruiker zelf die de niet-functionele *governance*-eisen moet handhaven.

• *Kwadrant IV: Research, innovatie, prototyping*

Kwadrant IV staat in het teken van het adagium ‘experimenteer, leer en verbeter’. Het is het kwadrant waar medewerkers vragen om (bij voorkeur goede) data, computerkracht en geavanceerde software om data-analyses uit te voeren. Het is het kwadrant waar veel vrijheid wordt geboden en veel wordt gevraagd van de verantwoordelijkheid van de betreffende medewerkers. Kwadrant IV wordt voor verschillende toepassingen

gebruikt. Veel voorkomend is prototyping: het iteratief ontwikkelen van een basisversie van een informatieproduct. Prototyping is bij uitstek geschikt wanneer een gebruiker een globaal idee heeft van zijn informatiebehoefte en deze graag in de vorm van een voorbeeld uitgewerkt wil zien. Samen met de gebruiker wordt dan vaak in een aantal iteraties het gewenste product ontwikkeld. Ad-hocanalyses zijn een andere toepassingsvorm. In feite zijn dit alle mogelijke soorten niet-standaardanalyses op reeds aanwezige data. Research is een andere toepassingsvorm, zoals onderzoek naar nieuwe risicomodellen voor kredieten, analytische voorspelmodellen voor solvabiliteit bij verzekeraars of algoritmes die *outliers* vroegtijdig identificeren. Daarnaast is kwadrant IV de aangewezen omgeving voor innovatie, zoals het uitproberen van nieuwe analytische technieken als *machine learning*. Ten slotte is kwadrant IV – vaak in combinatie met kwadrant III – de Haarlemmerolie van een datavoorziening. Voor situaties waarin er direct gereageerd moet worden zoals bij een calamiteit of andere niet van tevoren bedachte vormen van gebruik, biedt kwadrant IV de ruimte om naar eigen inzicht data te verwerken en analyses uit te voeren.

Van visie naar inrichting

Het kwadrantenmodel is door DNB gebruikt als referentiekader bij het doorvoeren van structurele vernieuwingen in de wijze waarop gegevens worden verwerkt voor toezicht en de macrostatistiek. Hierbij heeft het model niet alleen zijn toegevoegde waarde bewezen bij het (opnieuw) ordenen en inrichten van de (fysieke) gegevensverwerkingen maar ook bij het bepalen van de hiervoor gewenste organisatorische inrichting, besturing, benodigde kennis en competenties en noodzakelijke technologieën.³⁸ Deze invalshoeken komen hieronder achtereenvolgens aan de orde.

Keuzes in voortbrengingspatronen

Eerder is gemotiveerd dat een toezichthouder te maken heeft met diverse niet-functionele eisen als traceerbaarheid, betrouwbaarheid, toegankelijkheid en betrouwbaarheid. Het kwadrantenmodel maakt duidelijk hoe de *governance* van deze eisen verloopt. Aan de hand van het model kunnen nu keuzes gemaakt worden welke patronen van gegevensvoortbrenging wel en niet zijn toegestaan.

- Het systematische patroon: van KI naar KII. Dit patroon wordt gevolgd voor gegevensvoortbrengingen die structureel en periodiek moeten worden uitgevoerd en waarbij de niet-functionele eisen (geautomatiseerd) moeten worden gewaarborgd. Voorbeelden zijn periodieke doorleveringen van CRDIV en Solvency II-rapportages naar de ECB resp. EIOPA en dashboards voor het toezicht op de verzekeringssector.

38. Een beschouwing van benodigde technologieën is met opzet uit dit artikel weggelaten.

- Het opportunistische patroon: KIII naar KIV. Hier speelt de gegevensvoortbrenging zich vrijwel volledig af in de persoonlijke werkomgeving van een gebruiker of onderzoeksgroep. Voor een specifiek doel worden gegevenssets (handmatig) ontsloten en vervolgens gebruikt voor analyse. Een voorbeeld is een onderzoek van DNB naar de procyclische van initiële margevereisten van renteswaps (IRS), waarbij gebruik wordt gemaakt van informatie over de dagelijkse posities ontleend aan verschillende transactieregisters en geavanceerde analytische modellen.
- Het wendbare patroon: KI naar KIV: in dit patroon worden systematisch ontvangen en opgeslagen data vervolgens voor analytische doeleinden gebruikt. Zo worden de CRDIV-data nadat ze zijn ontvangen gevalideerd en opgeslagen, daarna ook beschikbaar gesteld voor analisten in KIV die ieder vanuit hun eigen werkomgeving ad-hocanalyses op deze dataset willen uitvoeren. In de praktijk is dit een veel gebruikt patroon dat door gebruikers en analisten hoog gewaardeerd wordt. Ze hoeven immers niet zelf voor de (periodieke) ontsluiting en validatie van de data te zorgen maar krijgen wel vrijheidsgraden in het zelf doorzoeken en analyseren van dit soort datasets.

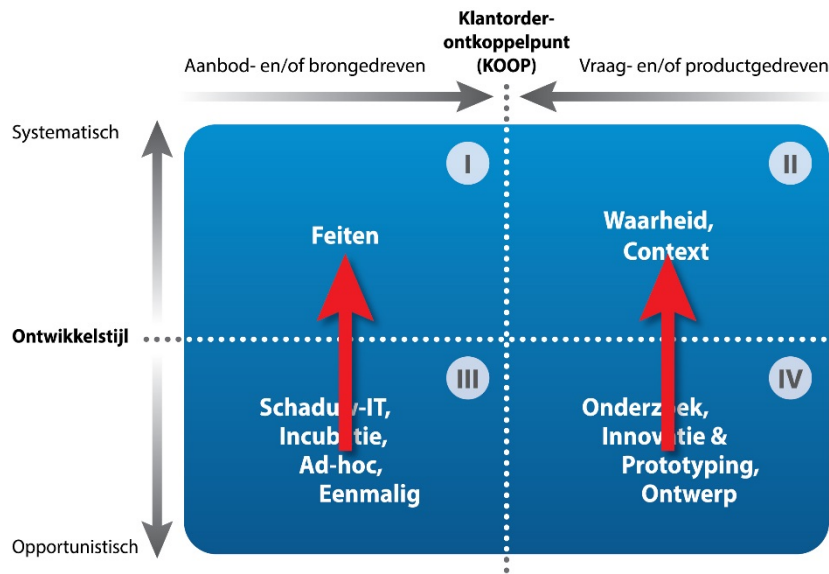
Vanuit het oogpunt van verantwoording en risicobeheersing is het daarentegen ongewenst dat opportunistisch voortgebrachte data (kwadrant III) worden gebruikt in formele informatieproducten (kwadrant II). Indien deze behoefte toch bestaat, is het nodig de gegevensontsluiting in kwadrant III te promoten naar kwadrant I, ofwel een degradatie van het informatieproduct in kwadrant II naar kwadrant IV (waarmee het onder verantwoordelijkheid komt van de data-analist of gebruiker die initieel het informatieproduct heeft ontwikkeld).

- *De ultieme questie: van kwadrant III/IV naar kwadrant I/II*

De grootste uitdaging bij het realiseren van datagedreven toepassingen is het structureel toepassen van bruikbare algoritmes, geslaagde experimenten en prototypen in reguliere toezichtprocessen. In termen van het kwadrantenmodel gaat het hierbij om de beweging van kwadrant III/IV naar kwadrant I/II (zie Figuur 4)

Van een algoritme dat ‘werkt op mijn laptop’ moet nu een permanent werkend en periodiek geautomatiseerd voortgebracht informatieproduct gemaakt worden dat schaalbaar is naar een grote(re) groep gebruikers en ‘*by design*’ voldoet aan de gestelde niet-functionele eisen.

Uit ervaringen in de praktijk blijkt dat hier een forse uitdaging ligt die voortkomt uit een brede kloof tussen kwadrant I/II en kwadrant III/IV. Die kloof wordt gevormd door verschillen in de manier van werken (systematisch versus opportunistisch), profiel van betrokken medewerkers (de ‘*suits*’ versus de ‘*hoodies*’), gebruikte technologieën en organisatorische positionering (centraal versus decentraal). Het is van belang dat leidingge-



venden erkennen dat alle kwadranten van cruciaal belang zijn en dat de randvoorwaarden worden gecreëerd voor optimale samenwerking en afstemming. Het ontwikkelen van een dergelijke *organisatorische competentie* blijkt in de praktijk een van de grootste uitdagingen in het werken met data te zijn en moet op vele vlakken gerealiseerd worden: cultuur, mens, proces en technologie. Dit moet gezien worden als een forse culturele transformatie waarbij nieuwe vormen van samenwerking tussen het datadomein en het toezichtdomein gevonden en opgezet moeten worden.

Keuze in organisatorische inrichting en governance

Bij de organisatorische inrichting van de datavoorziening kunnen verschillende functies worden onderscheiden. Zij verschillen wezenlijk van elkaar en dienen in enigerlei vorm belegd te zijn in een datagedreven organisatie.

- *Data Service Centre*

Het Data Service Centre (DSC) realiseert en onderhoudt de systematische en periodieke ontsluiting van reguliere databronnen. Het DSC richt daartoe, in nauw overleg met het (meest) betrokken toezichtdomein per gegevensbron (bijvoorbeeld een verplichte rapportage voor banken of een reguliere gegevensuitwisseling met een ketenpartner), het benodigde gegevensverwerkingsproces in dat de gecontroleerde voortbrenging van de data uitvoert. Deze verwerking wordt gekenmerkt door (geautomatiseerde) voortgangsbewaking, geautomatiseerde validatie op tijdige levering en kwaliteit van de data en de communicatie hierover met de desbetreffende gegevensleveranciers en rapporteurs. De zwaarte van de vereiste validaties wordt daarbij primair bepaald door de eisen van de gebruiker, de gegevenseigenaar en/of onderliggende wet- en regelgeving. Het DSC is daarmee het organisatieonderdeel dat de voortbrenging van gegevens in kwadrant I ondersteunt.

- *Business Intelligence Competency Centre*

Het Business Intelligence Competency Centre (BICC) verzorgt de systematische en periodieke uitnutting van de brondata in de vorm van gestandaardiseerde, desgewenst dynamische, informatieproducten (zoals rapportages en dashboards) ten behoeve van de reguliere analyseprocessen van de verschillende toezichtdomeinen. Ook dit betreft een gegevenswerkend proces maar dan gericht op het feitelijk gebruik van de data (*pull*, vraaggedreven). Het werk van het BICC wordt gekenmerkt door de vertaalslag die gemaakt moet worden om de data inzetbaar te maken voor reeds gestandaardiseerde werkprocessen. Dit impliceert dat het betrokken toezichtdomein over een eenduidig, samenhangend en stabiel beeld van de uit te voeren analyseprocessen dient te beschikken. Dit vergt vaak een intensief leerproces binnen toezicht om tot standaardisatie te komen van de gewenste, optimale werkwijzen. Juist voor dit doel worden vaak pilots of prototypes ontwikkeld in kwadrant III/IV (zie ook hieronder). Op basis van de vereisten die vanuit deze prototypes aan het licht komen, ontwikkelt het BICC structurele informatieproducten en gegevensleveringen. Het BICC is dus het organisatieonderdeel dat de voortbrenging van gegevens in kwadrant II ondersteunt.

- *Data-analyseafdelingen*

Data-analyseafdelingen zijn afdelingen, teams of soms zelfs individuele medewerkers die als taak hebben om data-analyses uit te voeren voor domeinspecifieke vraagstukken. Dit kunnen reguliere analyses betreffen zoals het opstellen van een marktbeeld van de toezichtpopulatie, maar het kan evenzeer gaan om ad-hoconderzoek naar de bruikbaarheid van nieuwe databronnen of de verkenning van een onderzoeksvraag met de inzet van bestaande datasets. Afgebeeld op het kwadrantenmodel zijn dit doorgaans gebruikers die opereren in kwadrant III en IV.

De verschijningsvorm van dit soort afdelingen of teams is veelzijdig en niet altijd direct herkenbaar binnen het gebruikelijke organogram. Richtinggevend is de mate waarin binnen een afdeling vormen van data-analyse plaatsvinden. Refererend aan het organogram van DNB betreft het daar bijvoorbeeld expertisecentra binnen verschillende toezichtdivisies, statistiekafdelingen binnen de divisie Statistiek, beleidsafdelingen, onderzoeksafdelingen en marktanalyse.

- *Data Management Office*

Het Data Management Office (DMO) heeft het formele mandaat om vanuit een organisatiebreed perspectief de regie te voeren op de informatiepositie en datavoorziening van de instelling als geheel. Hoofddoelstelling is dat alle geledingen van de organisatie op efficiënte en verantwoorde wijze optimaal gebruik kunnen maken van data ter ondersteuning van hun taken en verantwoordelijkheden. Centraal daarbij staat het realiseren van een optimale balans tussen enerzijds de noodzaak tot standaardisatie van technologieën, processen en spelregels en anderzijds het ondersteunen van de verscheidenheid aan gebruikerswensen zonder daarbij te vervallen in duplicering van gegevensvoortbrengingen of specifieke maatwerkoplossingen.

Naar hun aard vormen het DSC en het DMO sterk gecentraliseerde organisatieonderdelen omdat zij domeinoverstijgende taken vervullen. Het DSC is daarbij vooral gericht op operationele taken op datagebied, terwijl het DMO zich bezighoudt met de strategische en tactische inrichtingsvraagstukken. Het BICC is daarentegen niet noodzakelijkerwijs instellingsbreed georiënteerd. Het is immers goed denkbaar dat een specifiek gebruiksdomein dermate ruim en intensief gebruikmaakt van data in zijn werkprocessen dat de slagkracht van dat domein om meerwaarde uit data te halen sterk gediend is bij de inrichting van een eigen BICC. Vanzelfsprekend zal het DMO ook voor dergelijke 'locale' BICC's de instellingsbrede regievraagstukken verzorgen.

De regiefunctie van het DMO krijgt invulling door het opstellen van beleid en spelregels.

Hiermee wordt voorgeschreven hoe de gegevensverwerking ingericht dient te worden, waarbij maatwerkoplossingen worden beperkt en generieke, instellingsbrede oplossingen worden bevorderd voor de uiteenlopende gebruiksvarianten die binnen een organisatie voor het werken met data bestaan. Om deze *data-architectuur* (dat wil zeggen het ontwerp van de inrichting van de gegevensverwerking) '*fit for purpose*' te houden dient zij systematisch getoetst te worden aan externe ontwikkelingen en nieuwe gebruikerswensen tot uitbreiding c.q. aanpassing. In veel gevallen wordt daarvoor binnen een datagedreven organisatie een Change Advisory Board (CAB) opgericht waarbij gebruikers en data-architecten gezamenlijk proberen tot optimale oplossingen te komen. Bijzondere vermelding daarbij verdienen niet-functionele eisen met betrekking tot gegevensbeveiliging en de kwaliteitswaarborgen die gesteld dienen te worden indien data een systematische rol spelen in bedrijfspro-

cessen van de toezichthouder, waardoor de verantwoordingsvereisten navenant hoger komen te liggen.

Competenties van medewerkers

Aan de hand van het kwadrantenmodel is duidelijk gemaakt welke verschillende vormen van gegevensvoortbrenging onderscheiden kunnen worden. Deze vormen zijn ook in de organisatorische inrichting zichtbaar door middel van verschillende afdelingen die ieder specifieke ondersteuning bieden. In lijn hiermee zijn er ook verschillende en nieuwe soorten kennis en competenties nodig om hier invulling aan te geven. Algemeen kan gesteld worden dat de toezichthouder er belang bij heeft om zijn competenties in het werken met data – zijn *data literacy* – te vergroten. In de strategische agenda van sommige toezichthouders is dit reeds zichtbaar.³⁹ Meer specifiek worden binnen de verschillende kwadranten nieuwe functies zichtbaar. De *datamodelleur* specificeert gegevens en datasets door middel van modellen⁴⁰ en legt daarmee alle eigenschappen en niet-functionele eisen van de data en bijbehorende verwerkingen op een formele manier vast. De *data engineer* verzorgt op basis van deze specificaties de technische realisatie en periodieke operatie van de gegevensverwerking. De datamodelleur en -engineer vallen onder het Data Service Centre en werken hoofdzakelijk in kwadrant I. In kwadrant II staat de vraagzijde en daarmee een specifieke toezichtcontext centraal. In dit kwadrant is het dan ook essentieel om materiedeskundigheid te organiseren. Voor toezichthouders betekent dit dat ze zich de voor hen relevante databronnen eigen moeten maken qua betekenis, betrouwbaarheid en relevantie. Op basis van dit inzicht moeten de toezichthouders hun vraag formuleren in termen van gewenste informatieproducten. Ten slotte zullen ze praktische analysevaardigheden moeten aanleren om gerealiseerde informatieproducten in hun dagelijkse toezichtpraktijk ook daadwerkelijk te gebruiken. De materiedeskundigheid moet ook terugkomen bij de rollen die de informatieproducten in kwadrant II realiseren. Afhankelijk van gehanteerde functiebenamingen komen hier rollen terug als *informatieanalist*, *business intelligence specialist* of *applicatieontwikkelaar*. De gemene deler binnen deze rollen is dat kennis van de specifieke toezichtcontext, beschikbare datasets (in kwadrant I) en beschikbare technologische mogelijkheden nodig is alsmede de competentie om de geformuleerde vragen zowel functioneel als technisch te vertalen naar werkende oplossingen met inachtneming van niet-functionele eisen. In het scherp krijgen van zijn vraagstelling staat de toezichthouder niet alleen. Immers, mede voor dit doel is verwerking in kwadrant III en IV bedoeld. Met name in deze kwadranten is de *datawetenschapper* of *data scientist*⁴¹ gepositioneerd. Het profiel van deze relatief nieu-

39. Zie bijvoorbeeld de Agenda 2018 van de AFM, p. 36, <https://afm.nl/~/profmedia/files/afm/2018/agenda-2018-en.pdf>.

40. Zie bijvoorbeeld het logisch datamodel voor AnaCredit: www.dnb.nl/statistiek/digitaal-loket-rapportages/statistische-rapportages/banken/anacredit/index.jsp.

41. T.H. Davenport en D.J. Patil, 'Data scientist. The sexiest job in the world', *Harvard Business Review* 2012, nr. 10, p. 70-76.

we functie kent een mix van competenties op het gebied van analytische technieken en algoritmes, wetenschappelijke aanpak en programmeren. Het spreekwoordelijke schaap met vijf poten. Juist vanwege deze combinatie van vaardigheden bewijst deze functie bij uitstek haar toegevoegde waarde bij het werk in kwadrant III en IV. Snel en iteratief kunnen schakelen tussen ideeën en technische realisatie door middel van nieuwe gegevensvoortbrengingen en data-analyses is hier essentieel.

Conclusies

Het lijkt onomkeerbaar dat data een nieuw, permanent bedrijfsmiddel worden voor toezichthouders en organisaties in het algemeen. Dit brengt een groeiende behoefte aan ‘databewustzijn’ met zich mee. Ingeschatte risico’s worden in toenemende mate gebaseerd op basis van verscheidene (data)bronnen en slimmer wordende algoritmes die trachten de kennis van de expert te verbeteren. Hierover verantwoording kunnen afleggen vereist een gedisciplineerde omgang met data. Ervaring binnen verschillende organisaties laat zien dat het kwadrantenmodel in de praktijk behulpzaam is om met elkaar op een begrijpelijke wijze over de datahuishouding te kunnen praten vanuit een toezichtperspectief en stappen te zetten op weg naar een volwassen datavoorzieningsfunctie.

Het ligt dan ook voor de hand dat een volwassen datavoorziening als volwaardige bedrijfsfunctie wordt ingericht zoals ook gebruikelijk is voor andere bedrijfsmiddelen als medewerkers (human resource management) en kapitaal (financieel management). Anders dan andere organisaties hebben toezichthouders hierbij een extra uitdaging om deze inrichting vorm te geven binnen een set van noodzakelijke niet-functionele eisen. Bij het maken van afwegingen tussen verschillende eisen en belangen helpt het kwadrantenmodel om hier duiding aan te geven. De toenemende adoptie van het model binnen andere organisaties laat zien dat het model zijn nut bewijst.