

Whitepaper

Van exponentiële datagroei naar waardegedreven datastrategie

Martin van Mierloo
Hans Bossenbroek

1. Inleiding



In het stadje Ambalappuzha in Zuid-India ligt de Sri Krishna tempel. In deze tempel serveren ze al honderden jaren zoete Paal Paysam rijstpudding. Het kan goed zijn dat u hier nog nooit van heeft gehoord. Maar de legende achter deze rijstpudding kent u ongetwijfeld.

Lang geleden kreeg de koning van Ambalappuzha een pelgrim op bezoek. Hij daagde de pelgrim uit voor een potje schaak. De koning vroeg voor de wedstrijd welke prijs de pelgrim wilde hebben. De pelgrim vertelde dat een bescheiden prijs met een hoeveelheid rijst voldoende was. Op een schaakbord moest in het eerste vakje een rijstkorrel gelegd worden. Bij ieder volgend vakje moest het aantal rijstkorrels verdubbeld worden, totdat alle 64 vakjes gevuld waren. De koning stemde hier mee in.

Na een spannende schaakwedstrijd won de pelgrim. De koning liet enkele zakken rijst komen. Al ruim voordat de helft van het schaakbord gevuld was, kwam de koning tot het inzicht dat hij deze schuld nooit af kon lossen. De pelgrim onthulde toen zijn echte identiteit: hij bleek de God Khrisna te zijn. Hij vertelde de koning dat hij zijn schuld niet meteen hoefde af te lossen, maar dat ook over langere tijd kon doen. De koning besloot toen vanaf dat moment gratis Paal Paysam rijstpudding te serveren voor alle pelgrims die de Sri Khrisna tempel bezochten. Tot vandaag de dag krijgt u nog steeds een Paal Paysam rijstpudding als u de tempel bezoekt.

De exponentiele toename van rijst is te vergelijken met de exponentiele groei van data waar organisaties op dit moment mee te maken hebben¹⁾. Gedreven door nieuwe mogelijkheden en lagere kosten van opslag neemt de hoeveelheid data exponentieel toe. Toepassingen zoals video, Internet of Things, social media en Artificial Intelligence (AI) zorgen niet alleen voor de creatie van enorme hoeveelheden data, maar versterken ook de consumptie van immense hoeveelheden data.

Iedere organisatie bezit data. Toch slagen nog te weinig organisaties erin om data echt als strategisch asset in te zetten. Zonder planmatige aanpak blijft data een set nullen en enen zonder waarde. Sterker nog, zonder planmatige aanpak vormt de exponentiele groei van data eerder een bedreiging dan een kans.

In deze whitepaper onderbouwen we het belang van een planmatige aanpak van het omgaan met data en het ontwikkelen van een datastrategie. We omschrijven hoe u de dynamiek van data voor u kunt laten werken in plaats van tegen u. Hierdoor kunt u meer waarde halen uit uw data en daarmee nieuwe verdienmodellen ontwikkelen, beter innoveren en efficiënter opereren. Dit doen we aan de hand van de laatste theorieën op het gebied van data en aan de hand van een aantal voorbeelden.

2. De context van een datastrategie



Sinds de opkomst van het internet zijn organisaties door verschillende transformaties gegaan. Deze transformaties waren noodzakelijk om de concurrentie voor te blijven en te voldoen aan de verwachtingen van consumenten. Iedere transformatie heeft zijn oorsprong in nieuwe digitale mogelijkheden en bouwt voort op eerdere stappen.

2.1 De digitale transformatie

De eerste transformatie bestond uit het online brengen van de business. De tweede transformatie leidde daarna tot digitalisering van interne processen, vaak gedreven door mogelijkheden van technologie en kostenbesparingen.

De afgelopen jaren zijn veel organisaties gestart met een digitale transformatie¹⁾. Hierin staat niet alleen de verdere digitalisering centraal, maar ook de organisatievorm, digitale vaardigheden en digitale innovatie. Een digitale transformatie wordt vaak door het management gebracht als een offensieve tactiek om tot nieuwe waardecreatie te komen. In werkelijkheid is het meestal een noodzakelijke defensieve reactie op de innovatiesnelheid van startups, concurrenten en de macht van de platformbedrijven (Google, Amazon, Facebook etc.). Markten worden steeds onvoorspelbaarder door nieuwe toetreders en innovaties. Daarom moeten organisaties sneller innoveren en sneller nieuwe ideeën naar de markt brengen.

Organisaties slagen er zonder een digitale transformatie vaak onvoldoende in om snel andere vormen van waarde te creëren en vast te houden. Zeker binnen grote organisaties bestaan soms extreme verschillen van innovatiekracht tussen verschillende bedrijfsonderdelen²⁾. Een digitale transformatie zorgt ervoor dat de verschillen kleiner worden en de totale innovatiekracht vergroot wordt. Bijvoorbeeld door andere organisatievormen toe te passen, zoals

bijvoorbeeld het Spotify model, of vormen van zelforganisatie³⁾. Hierdoor worden barrières voor innovatie weggehaald. En ja, daarbij blijkt vaak dat 'het management' en de besturing het grootste obstakel vormen voor innovatie. Een andere bekende vorm is innovatie los van de organisatie neerzetten, ook wel spin-offs of spin-outs genoemd.

2.2 Na de digitale transformatie

Na een digitale transformatie zullen organisaties nog verder transformeren. Hier bestaan verschillende theorieën over. Enkele aspecten die hier waarschijnlijk in terug zullen komen:

- **Grotere rol voor ecosystemen en netwerken:** De fases 'online', 'digitalisering' en 'digitale transformatie' gingen voornamelijk over de organisaties zelf. De volgende transformatie zal waarschijnlijk gaan over de rol van de organisatie ten opzichte van haar ecosysteem. Dit bestaat o.a. uit leveranciers, klanten, kennis- en technologiepartners. De grootste versnelling in waardecreatie zal waarschijnlijk gebeuren door samen met partners in het ecosysteem te werken aan open innovatie. Hierbij zullen maatschappij en economie transformeren van waardeketens naar open en dynamische waardenetwerken.
- **Grotere rol voor duurzaamheid:** De maatschappelijke tendens naar meer duurzaamheid raakt iedere organisatie. Dit gebeurt op een directe manier door strengere regelgeving. Op een indirecte manier gebeurt dit door de reputatie van organisaties.
- **Vertrouwen (trust) wordt leidend:** vertrouwen binnen organisaties, in organisaties en tussen organisaties, zal centraal worden geplaatst in keuzes en strategie. Binnen vertrouwen vormt *veiligheid* - in letterlijke én emotionele zin - de belangrijkste pijler.

- **Purpose shift of purpose reboot:** het ultieme bestaansrecht van organisaties wordt steeds vaker op de proef gesteld, omdat oude bedrijfs- en verdienmodellen sneller achterhaald zijn. Daarom zullen organisaties steeds vaker en actiever hun missie bijstellen en aanscherpen. De *purpose* van een organisatie wordt steeds belangrijker om de juiste klanten, partners en talenten aan te trekken.
- **Van functies naar rollen.** Agile werken, DevOps en zelforganisatie vragen van medewerkers brede, generalistische kennis. Hoe groter de organisatie, hoe lastiger de kennisvergaring en -deling⁴⁾. Daarom zullen organisaties steeds sneller medewerkers met de juiste kennis moeten inzetten op de juiste plaats. Een logisch gevolg hiervan is dat formele functies verdwijnen en we meer van I-shaped en T-shaped naar X-shaped kennisprofielen gaan⁵⁾. Hierdoor stroomt kennis en kunde nog makkelijker door de organisatie en werkt het juiste talent op de juiste plek. Eventuele kennisgaten worden opgevuld met externe kennis. Daarom zal de flexibele schil van bedrijven toenemen. Tegelijkertijd is de flexibele schil ook onderdeel van het ecosysteem van een organisatie.

Tussen deze verschillende onderwerpen bestaan veel relaties. Bijvoorbeeld: om duurzaam onderdeel te zijn van een ecosysteem zal een organisatie te vertrouwen moeten zijn en zal de missie van een organisatie in het verlengde moeten liggen van haar partners.

Deze nieuwe vorm van transformeren zal, in tegenstelling tot een digitale transformatie, nog minder volgens een vast recept verlopen, omdat het ontdekken van duurzame waarde in een ecosysteem namelijk voor iedere organisatie anders zal zijn.

2.3 Het belang van waardecreatie

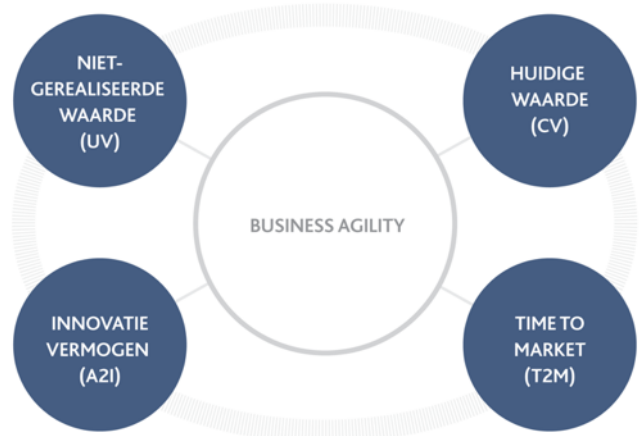
U zult zich misschien afvragen wat digitale transformatie en de hierboven geschetste veranderingen met data en een datastrategie te maken hebben. Deze veranderingen zullen veel invloed uitoefenen op de creatie van waarde door middel van innovatie, technologie en data. In de volgende hoofdstukken gaan we bijvoorbeeld in op de duurzaamheid van data, waardevermeerdering van data in ecosystemen en op veiligheid en vertrouwen in data. Deze veranderingen geven een andere dimensie aan het gebruik van data en zijn daarom belangrijk om mee te nemen in een datastrategie.

Deze transformaties hebben ook gevolgen voor uw IT-strategie. De IT-afdeling of -divisie wordt binnen veel

organisaties nog steeds gezien als een *cost center*. Deze manier van besturen paste goed bij de fase van digitalisering, maar gaat niet op in de fases daarna. In de volgende fases wordt de creatie van waarde als uitgangspunt belangrijker. De IT-afdeling moet daarom anders bestuurd en ingericht worden om een significante bijdrage te leveren aan de doelen van de organisatie.

Ken Schwaber (co-auteur van het Agile Manifesto) drukt de waarde van de organisatie uit langs vier assen⁶⁾:

- **Current value:** hoeveel waarde levert een organisatie nu voor medewerkers en klanten. Uitgedrukt in bijvoorbeeld klant- of medewerkerstevredenheid, omzet per medewerker of product-kosten-ratio.
- **Time to market:** hoe snel kan een organisatie nieuwe producten of diensten beschikbaar maken. Uitgedrukt in bijvoorbeeld: build- en release frequentie, tijd om te repareren, tijd van idee tot release, tijd van idee tot klantonderzoek etc.
- **Ability to innovate:** hoe goed is een organisatie in innovatie. Uitgedrukt in aantal gegenereerde ideeën en percentage R&D-tijd.
- **Unrealized value:** de waarde van de producten en diensten die nog gerealiseerd kan worden. Uitgedrukt in bijvoorbeeld percentage marktaandeel, of verschil tussen klantverwachting en realiteit.

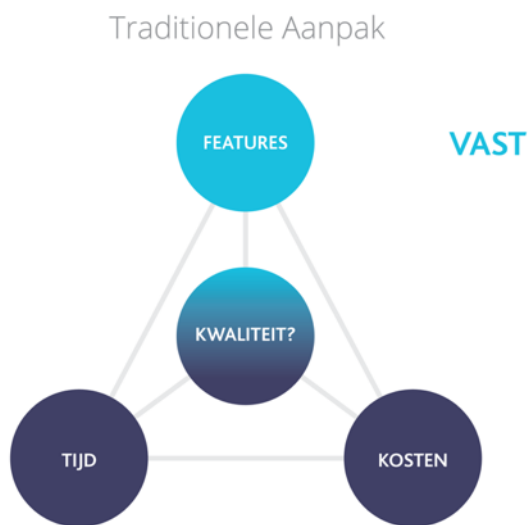


Deze uitgangspunten, inclusief onderliggende KPI's, kunnen in de breedte en in de diepte worden toegepast in een organisatie. Van organisatie- tot teamniveau en van development teams tot stafafdelingen. Door deze waarden te operationaliseren op verschillende niveaus bereikt een organisatie een veel grote wendbaarheid, zal er meer innovatie zijn en zal de waarde voor klanten vergroot worden.

2.4 Een post-agile aanpak

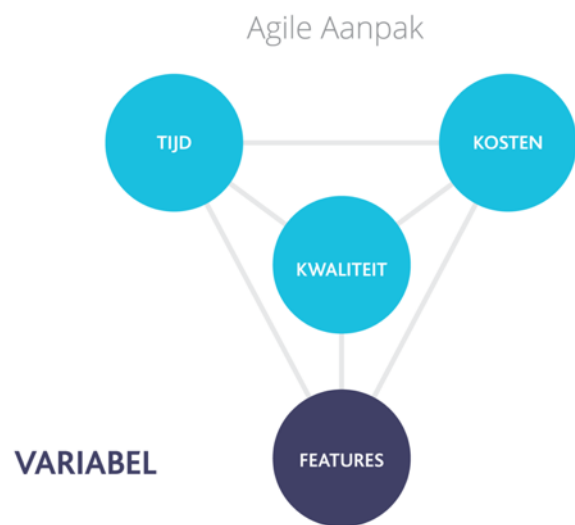
Bij traditionele softwareontwikkeling werden de functionaliteit en de kosten vooraf vastgesteld door middel van uitgebreide requirements en business cases. Als gevolg daarvan werd als eerste op de kwaliteit ingeboet. Dit zorgde ervoor dat veel IT-projecten niet alleen hun deadlines en budgetten overschreden, maar ook dat systemen vaak fragiel en incompleet waren.

Bij een Agile werkwijze staan tijd, kosten en kwaliteit vast. Features zijn de enige variabele. De features die de meeste (klant)waarde opleveren worden als eerste opgepakt. Hoewel dit een simpele verandering lijkt, zien we dat veel organisaties hiermee worstelen.

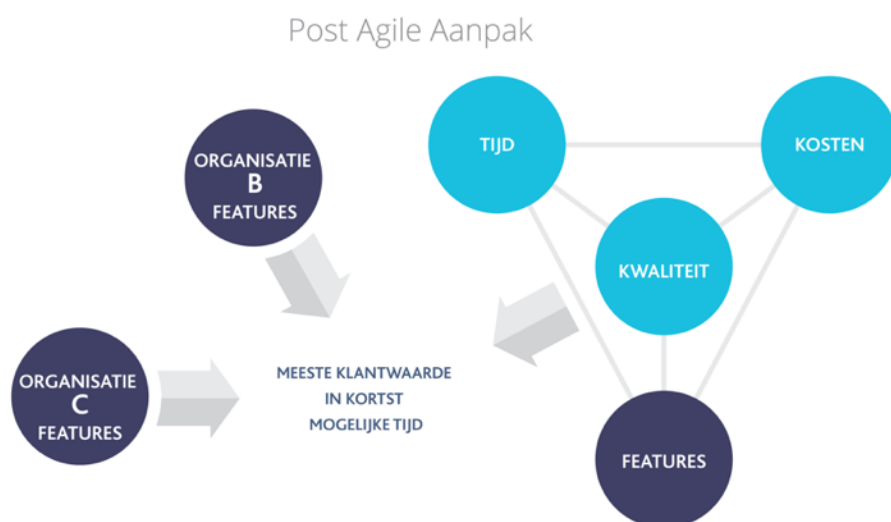


Deze verandering wordt vaak ingezet voor kleinschalige verbeteringen op teamniveau, maar minder frequent binnen de hele IT-organisatie en vrijwel nooit op organisatieniveau. Terwijl de praktijk juist laat zien dat dit nodig is.

Als we de voorgaande inzichten combineren met open innovatie, waarin een ecosysteem met partners, afnemers en leveranciers gezien wordt als deel van het geheel, dan wordt duidelijk hoe diepgaand de veranderingen zijn voor



organisaties. Deze zienswijze laat niet alleen het belang zien van eigen innovatiesnelheid en eigen focus op waardecreatie. Het toont ook het belang aan om dit synchroon te doen in een ecosysteem met partners, leveranciers en klanten. Hiermee wordt niet alleen gezamenlijk sneller waarde voor klanten ontwikkeld. Dit gebeurt op een duurzame wijze, waar meerdere partijen van profiteren.



2.5 Innovatie als leidend principe voor architectuur

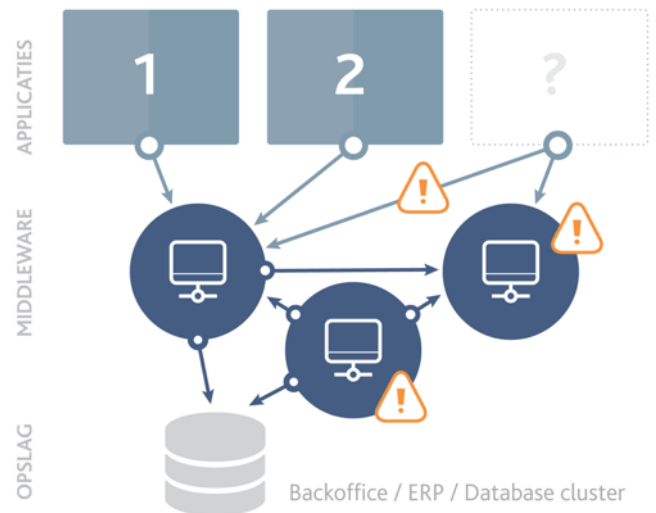
Veel enterprise architecturen zijn ontstaan vanuit een efficiëntie- en kostenperspectief in plaats van een waarde- of wendbaarheidsperspectief. Daardoor zijn deze architecturen gebaseerd op uitgangspunten als standaardisatie en centralisatie. Hierbij is integratie van systemen de belangrijkste uitdaging. Om maximaal invulling te geven aan waardecreatie, innovatie en *time to market*, zal een enterprise architectuur op fundamenteel andere wijze opgebouwd moeten worden:

- Presenteer een enterprise architectuur niet als doelarchitectuur of als kader, maar als richting. Want een architectuur is nooit af en zal gegarandeerd over 1 à 2 jaar achterhaald zijn. Benadruk de veranderlijkheid in plaats van het statische karakter van de architectuur.
- Hanteer wendbaarheid en modulariteit als belangrijkste principes. Dit zorgt ervoor dat maximaal invulling wordt gegeven aan innovatie. Het is een gegeven dat er in rap tempo nieuwe applicaties en services bij zullen komen. Kijk daarom hoe het bestaande landschap hier zo goed mogelijk mee om kan gaan. Kijk waar de meeste knooppunten en afhankelijkheden zitten. Probeer vervolgens systemen en services modulair te maken, zodat afhankelijkheden verkleind worden.

Een veelvoorkomend obstakel voor innovatie is de complexiteit van integraties. De grote hoeveelheid integraties in veel organisaties ontstaat doordat er veel wordt gewerkt met centrale applicaties en centrale vormen van dataopslag. Enterprise architecturen met een grote integratiecomponent (bijvoorbeeld Enterprise Service Bus of message brokers), kennen een aantal uitdagingen:

- Integratietechnologie omvat vaak grote hoeveelheden datatransformaties die specifiek zijn voor een domein. Daarmee is de herbruikbaarheid laag en zijn de kosten hoog.
- De onderhoudskosten van integratietechnologie zijn moeilijk toe te rekenen aan specifieke businessdoelen en daardoor lastig te zien als onderdeel van waarde.
- Het gebruik van integratietechnologie zorgt ervoor dat grenzen van systemen steeds onduidelijker worden. Steeds meer systemen praten 'verplicht' met een servicelandschap. Horen deze services en hun achterliggende applicaties of dataopslag nu tot een applicatie, of niet? Door deze afhankelijkheden ontstaat een steeds grotere spaghetti van services die moeilijk te ontvlechten is. Hierdoor wordt het IT-landschap van organisaties steeds onbeweeglijker en fragieler, wat weer tot hogere onderhoudskosten leidt.

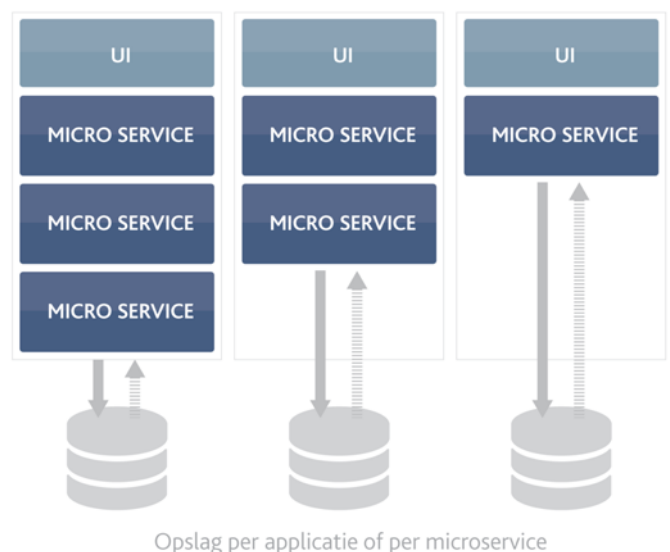
Klassieke architectuur



Veel enterprise architecturen kunnen door deze problemen slecht inspelen op de veranderlijkheid van de markt en de kansen die cloud-services bieden. Het toenemende belang van wendbaarheid voor diensten en data leidt tot andere uitgangspunten voor moderne enterprise architecturen. Bij deze architecturen staan modulariteit en data centraal.

In modulaire architecturen zijn systemen of applicaties opgebouwd uit onderdelen die verantwoordelijk zijn voor hun eigen functionaliteit en data. Deze functionaliteit en data worden ter beschikking gesteld via vooraf ontworpen services. Een veelgebruikte implementatie daarvan is het gebruik van microservices. Idealiter heeft iedere microservice de beschikking over eigen opslag die exact past bij het domein van de microservice. Hierdoor vermindert het aantal ongeplande afhankelijkheden en wordt de uiteindelijke wendbaarheid vergroot.

Modulaire Architectuur



Bij modulaire architecturen zijn datamodellen en dataopslag meer versnipperd⁷⁾. Daarom zijn aanvullende maatregelen nodig bij modulaire architecturen om grip te houden op de kwaliteit en management van data⁸⁾. Het goede nieuws is dat data en datamodellen meer gebonden zijn aan een domein. Dat maakt het eenvoudiger om bij dit domein zaken als een *dataglossary* of *master data management* toe te passen. Doe dit ook alleen waar het waarde toevoegt op het gebied van datakwaliteit of uitwisseling van data. En streef er ook niet naar om voor alle domeinen samen (of voor de hele organisatie) een *dataglossary* aan te leggen. Dit is een zeer tijdrovende klus en zal waarschijnlijk achterhaald zijn tegen de tijd dat de eerste versie er is.

De extreme innovatie in het clouddomein afgelopen jaren zorgt voor extra adaptie van modulaire architecturen. Immers, cloud services zijn in de basis ook sets van microservices met een lokale (domein)opslag. Om maximaal ruimte te bieden voor innovatie en maximaal waarde te halen uit de kansen die cloudtechnologie biedt, is een transitie naar een modulaire enterprise architectuur noodzakelijk.

2.6 De software commodity val

Afgelopen decennium zijn er twee evoluties in softwareontwikkeling geweest die de levenscyclus van software behoorlijk veranderd hebben:

- De productiviteit van softwareontwikkeling is enorm verhoogd door de opkomst van betere besturingssystemen, tools, frameworks, kennisdeling en andere organisatievormen (Scrum, DevOps).
- De opkomst van cloudtechnologie zorgt niet alleen voor lage opstart- en opslagkosten, maar biedt ook een rijke en centrale infrastructuur die bestaat uit kwalitatief hoogwaardige services. Cloud services bieden organisaties de oplossingen die de grenzen en structuur van iedere architectuur veranderen.

Deze twee evoluties versterken elkaar. Met name cloudtechnologie zal de komende jaren voor een verdere versnelling zorgen in softwareontwikkeling. Een gevolg hiervan is dat software steeds sneller verschuift van standaardproducten naar commodity. Softwareontwikkeling was 20 jaar geleden grotendeels maatwerk. Op dit moment wordt software steeds meer ontwikkeld volgens modulaire- en servicegeoriënteerde architecturen. De afgelopen jaren is dit door de opkomst van cloudtechnologie alleen maar versneld.

Dit betekent dat het belangrijker dan ooit is om scherpe keuzes te maken over welke software u zelf wilt ontwikkelen. Probeer dit niet alleen van het hier en nu te bekijken, maar probeer ook te kijken welke software binnen 1 à 2 jaar standaard beschikbaar zal zijn als cloud service, of als open source. Zodra dit gebeurt vermindert direct de waarde van uw software, terwijl de kosten van ontwikkeling, onderhoud en afscheid nemen vele malen hoger zijn dan gebruik van standaardsoftware. U kunt hiervoor gebruik maken van de verschillende fases, gebaseerd op de Wardley mapping methode⁹⁾:





3. Een ander perspectief op data

3.1 De financiële waarde van data

In 2017 publiceerde The Economist een artikel⁽¹⁾ waarin de auteurs onderbouwen waarom data inmiddels meer waard is dan olie. Sinds de opkomst van de benzinemotor is de toegang tot olie een drijvende kracht achter wereldwijde samenwerkingen én conflicten. Tech-giganten zoals Google, Facebook en Amazon zijn de nieuwe olieraffinaderijen. Waarbij ze niet alleen de olie winnen en verwerken, maar ook nog alle tankstations bezitten. De helft van alle online aankopen in de Verenigde Staten loopt bijvoorbeeld via een van de diensten van Amazon⁽²⁾. Daarbovenop komen nog de inkomsten van de e-commerce bedrijven die hun infrastructuur op de Amazon Cloud draaien.

Bedrijven zoals Google, Facebook en Amazon zijn extreem snel gegroeid, omdat ze slimme verdienmodellen hebben ontwikkeld op basis van hun data. Ze zetten dus niet alleen data in om kosten te verlagen, efficiënter te werken en gebruikers beter te helpen, maar slagen er ook in om hun data steeds opnieuw te kapitaliseren. In dat opzicht is data ook fundamenteel anders dan olie; data is anders voor iedere organisatie en daarmee nooit een commodity. En zelfs nadat data wordt geconsumeerd, behoudt het zijn waarde.

Veel organisaties besteden veel tijd aan hun business-strategie, maar weten nog onvoldoende data als strategisch

asset in te zetten. Onderzoek van o.a. Evans & Prince³⁾ laat zien dat het belang van data wel wordt gezien, maar dat er vaak geen duidelijke besturing en eigendom is. De meeste organisaties zien een CIO als bewaker van de data. Alleen is de CIO vaak niet verantwoordelijk voor data en wordt de waarde van data ook niet door de CIO gegenereerd, maar door de business.

In de jaren '90 ontstond de vraag of 'goodwill' ook in financiële waarde uitgedrukt kan worden. Afgelopen decennia zijn steeds meer organisaties goodwill ook op de winst- en verliesrekening gaan opnemen. Data wordt ook bij het merendeel van organisaties gezien als strategisch asset, dus economisch gezien hoort dit ook voorzien van een financiële waarde op de balans te staan. De verwachting is daarom dat data op termijn ook als asset op de winst- en verliesrekening wordt opgenomen. Dit betekent dat organisaties betrouwbaar moeten kunnen vaststellen wat de financiële waarde is van hun data. Is uw organisatie al in staat om dit te doen?

3.2 De waarde van data in de tijd

Data heeft te maken met een waardevermindering in de tijd. De meest waardevolle data is data die direct in het primaire proces ingezet kan worden om keuzes te maken, voorspellingen te doen of klanten direct te helpen.

Ondanks het feit dat realtime- en *actionable* data het meest waardevol is, ligt de focus van veel organisaties vaak op reactieve en historische data. Er wordt veel tijd en geld besteed aan *datawarehouses* en *data lakes*. Deze zijn in de meeste gevallen niet geschikt voor realtime analyse en dataverwerking. Hun waarde is vaak beperkt, omdat ze ver van het primaire proces afstaan.

Ook op het gebied van AI-toepassingen worstelen veel organisaties met toepassing van AI in het primaire proces.

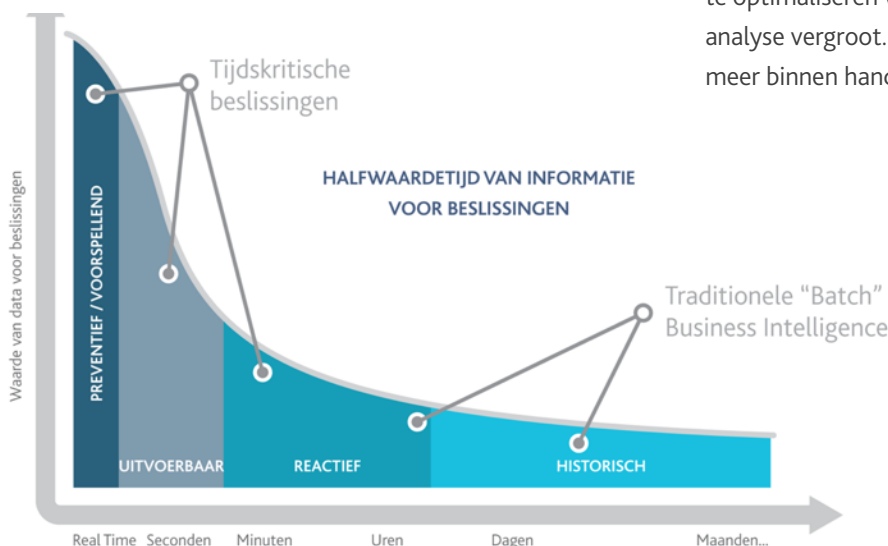
De kernoorzaak blijkt te zitten in de vele data-transformaties die nodig zijn om data geschikt te maken. Er zijn verschillende onderzoeken⁴⁾ die aantonen dat tot 80% van de data-analyse besteed wordt aan ophalen en geschikt maken van data. De uiteindelijke waarde wordt in 20% van de tijd bepaald. Deze datatransformaties zijn niet alleen bewerkelijk, maar ze laten zich ook lastig opschalen over grote datavolumes en omgevingen met veel verkeer.

AI inzetten om tot nieuwe inzichten te komen op basis van historische data is mooi, maar levert voornamelijk binnen de organisatie waarde op. AI moet ook onderdeel worden van het primaire proces als klanten en leveranciers hiervan willen profiteren.

Dit vraagt om een andere zienswijze, waarbij de inzet van realtime data in combinatie met AI-toepassingen meer centraal komen te staan. In een datastrategie moet nagedacht worden over het operationaliseren van realtime data, maar ook over het creëren van meerwaarde voor historische data. Dit kan door antwoord te geven op vragen zoals:

- Welke data is nodig om het primaire proces te verbeteren in termen van klantwaarde, efficiëntie en kosten?
- Welke AI-toepassingen kunnen de waarde van data vergroten, zodat het primaire proces beter of sneller verloopt?
- Zou door uw organisatie gegenereerde data andere organisaties of ketenpartners kunnen helpen in hun proces? Moeten ze realtime kunnen beschikken over deze data, of mag het ook reactief zijn?

Veel organisaties investeren momenteel in het neerzetten van DataOps. DataOps is een benadering die lijkt op DevOps, maar dan gericht op het stroomlijnen van het data-voortbrengingsproces. Door systemen, processen en mensen te optimaliseren wordt snelheid van datatransformatie en analyse vergroot. Daarmee komt de inzet van realtime data meer binnen handbereik.



Bron: Perishable Insights, Mike Gualtieri, Forrester

3.3 De duurzaamheid van data

Al sinds het begin van dit millennium is er een exponentiële ontwikkeling van de creatie van data. Deze wordt mogelijk gemaakt door technologische ontwikkelingen zoals big data en cloudtechnologie en goedkoper wordende opslagkosten. Deze hebben de weg vrijgemaakt voor hoogwaardige toepassingen op het gebied van video, big data en AI.

Deze exponentiële groei leidt tot een aantal belangrijke vragen: wat betekent exponentiële datagroei voor uw organisatie? En wanneer is genoeg genoeg? Is meer data ook altijd beter? En wat is de duurzaamheid van grote hoeveelheden data? Op basis van de huidige groeicijfers moeten we in het jaar 2100 meer data opslaan dan het aantal atomen op aarde⁵⁾. Dit is fysiek vrijwel onmogelijk, dus er zal een afvlakking of kantelpunt komen.

Tegelijkertijd geldt bij veel organisaties nog als norm dat meer data altijd beter is. De grote cloud-leveranciers verleiden organisaties massaal om tegen lage prijzen data op te slaan. Als gevolg hiervan neemt de energieconsumptie van datacenters ook fors toe, met grote gevolgen voor het milieu⁶⁾.

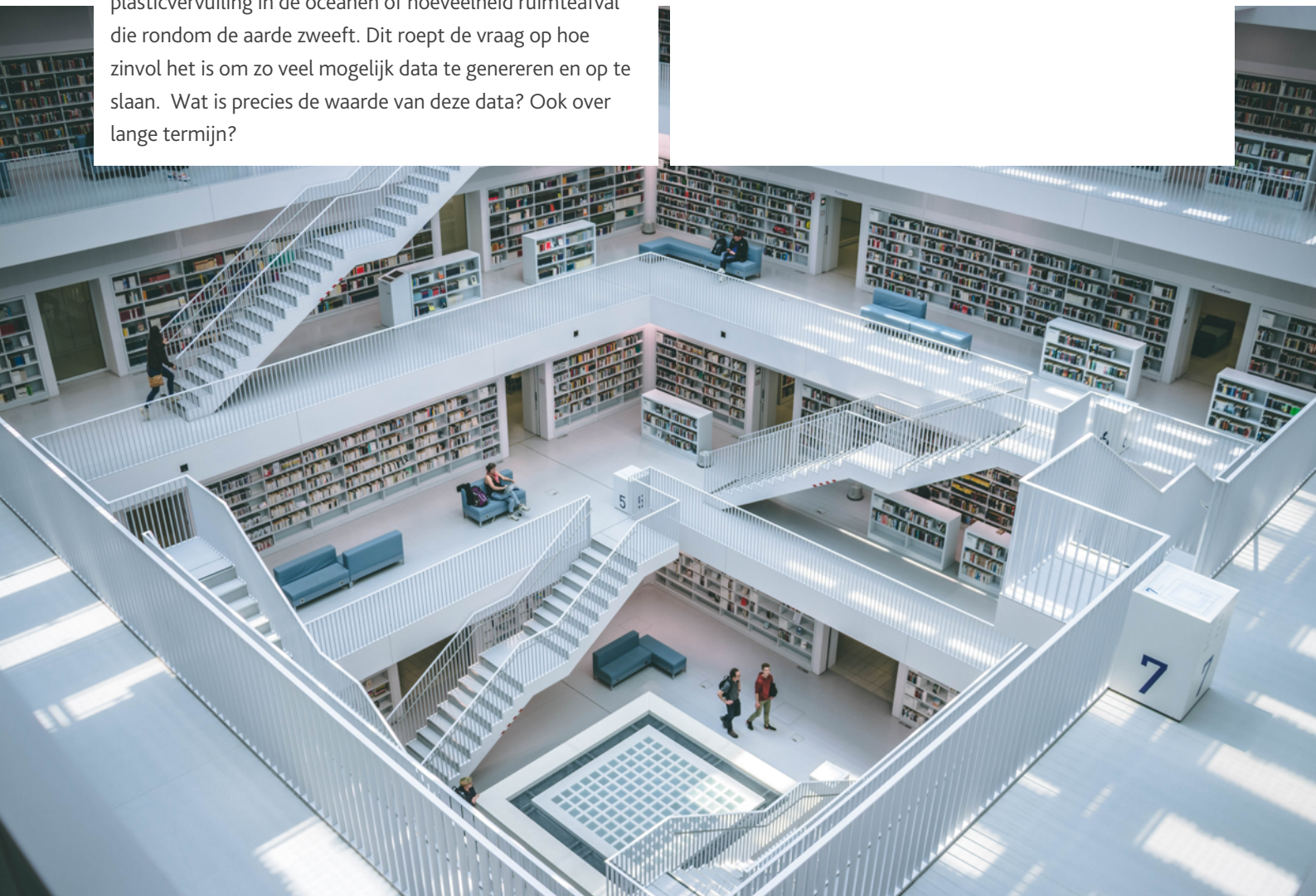
De exponentiële toename van data zal zonder duidelijke visie leiden tot situaties die te vergelijken zijn met de plasticvervuiling in de oceanen of hoeveelheid ruimteafval die rondom de aarde zweeft. Dit roept de vraag op hoe zinvol het is om zo veel mogelijk data te genereren en op te slaan. Wat is precies de waarde van deze data? Ook over lange termijn?

De complexiteit hierbij is dat de waarde van data vooraf niet altijd duidelijk is. En aangezien de kosten laag zijn om veel data te verzamelen en op te slaan ligt het voor de hand om dit te doen. Bijvoorbeeld: alle klikdata van alle bezoekers op een high-traffic website opslaan zal ongetwijfeld tot enkele nuttige inzichten leiden. Maar ook tot 99,9% overtollige data die zonder maatregelen ergens voor lange termijn blijft staan.

Daarom is het belangrijk om planmatig na te denken over de duurzaamheid van data. Idealiter begint dit bij de generatie van data. Stel hierbij vragen als:

- Hoeveel waarde vertegenwoordigt deze data? Bij voorkeur financieel uitgedrukt.
- Op welk moment verliest de data zoveel waarde dat archiveren of vernietigen beter is? Na X seconden, uren, dagen, maanden?
- Zijn er processen of procedures ingesteld om de data weer op te ruimen, archiveren of vernietigen?
- Welke andere organisaties of ketenpartners kunnen wellicht nog profiteren van de data?

Gecombineerd met de halfwaardetijd van data zal actief opschonen van data niet alleen leiden tot meer waardevolle data, maar ook tot betere veiligheid van data en een grotere duurzaamheid.





4. Structureel de waarde van data vergroten

4.1 Een gecombineerde aanpak

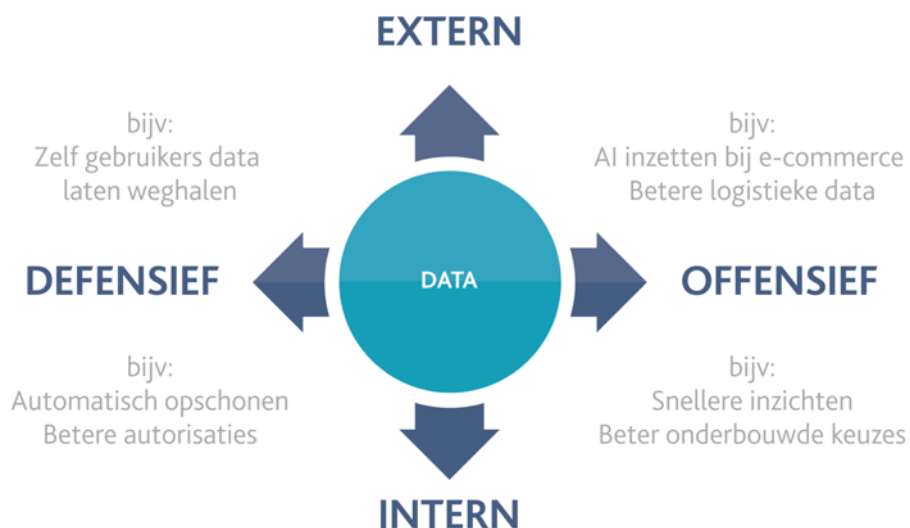
De vorige hoofdstukken omschrijven verschillende ontwikkelingen en aandachtspunten die een ander perspectief brengen op data.

Samengevat:

- Gedreven door toenemende concurrentie en veranderlijkheid moeten IT-afdelingen transformeren van kostengerichte- naar waardegerichte dienstverlening. De enterprise architectuur, gebruikte technologieën, data en organisatie moeten volledig in dienst staan van waardecreatie.
- Software veroudert in een hoog tempo en kent net als data een beperkte houdbaarheidsdatum. De halfwaardetijd van software en data zal alleen maar korter worden en vraagt daarom om een andere aanpak.

- Omdat vertrouwen en reputatie belangrijker en tegelijkertijd vluchtiger zijn dan ooit, zijn bescherming van data en persoonsgegevens cruciaal. Een datalek kan de reputatie enorm beschadigen. De beveiliging van data is fundamenteel anders dan de beveiliging van software: in het verleden werd data beschermd door de softwareapplicaties die de data creëerden, nu zal data apart beveiligd moeten worden omdat deze op andere manieren (her)gebruikt gaat worden.

Hoe maakt u als CIO de juiste keuzes rondom inzet van data en hoe doet u dit op een duurzame en kostenefficiënte manier? Een manier om hiernaar te kijken is om in uw datastrategie onderscheid te maken tussen offensieve tactieken en defensieve tactieken. Deze kunnen weer onderverdeeld worden in interne maatregelen en externe maatregelen:



4.2 De veiligheid en logistiek van data

Data is zo belangrijk geworden dat steeds betere maatregelen nodig zijn om data te beschermen tegen ongewenst misbruik. Dit heeft niet alleen te maken met toenemende zorgen rondom privacy en reputatieschade, maar ook met de toenemende uitwisseling van data tussen organisaties. Zeker nu organisaties steeds meer van ketens naar ecosystemen gaan wordt het nog belangrijker om expliciet vast te stellen wie bij welke data mag, hoe de data beveiligd is en wanneer data verwijderd of geanonimiseerd moet worden. Door uiteenlopende wetgeving tussen landen is dit voor internationale ondernemingen een behoorlijke uitdaging geworden. Ook cloud-infrastructuren leveren nieuwe vraagstukken op op het gebied van toegang, locatie en bewaartermijnen van data.

De defensieve kant van data was jarenlang het domein van databasebeheerders en van de afdeling die verantwoordelijk is voor informatiebeveiliging. Afgelopen jaren is dit domein door o.a. privacywetgeving (GDPR, AVG) veel complexer geworden. Nu kunnen ook eigenaren van persoonsgegevens een beroep doen op recht op vergetelheid of dataportabiliteit. Er zijn nieuwe rollen ontstaan zoals de Data Protection Officer. Verantwoordelijkheid en eigendom kunnen niet meer alleen maar bij de IT-afdeling liggen, maar zullen ook in de business verankerd moeten worden.

Met de opkomst van cloudtechnologie en modulaire architecturen volstaan algemene gecentraliseerde toegangso oplossingen niet meer. Data kan potentieel door tientallen services, systemen en organisaties bewegen. Daarom zal data steeds meer bij de bron voorzien worden van autorisatie-, encryptie- en privacyregels. Concreet betekent dit dat bij het modelleren van data uitgedrukt moet worden of deze data of attributen privacygevoelig zijn en wie geautoriseerd is. Een krachtige manier om autorisaties uit te drukken is het gebruik van Attribute Based Access Control (ABAC). Hierbij kunnen autorisaties uitgedrukt worden in eenvoudige maar krachtige expressies. Ook kunnen eventuele bewaar- en opschoonregels worden vastgelegd bij de creatie van data.

Door privacy, toegang en bewaartermijn te zien als integrale attributen van data (oftewel metadata) zullen de kwaliteit en veiligheid van data toenemen. Een ander aspect van goede metadata is de traceerbaarheid van data. Het zal normaal worden om data te waarmerken met een uniek kenmerk (bijv. UUID of EAN) zodat data beter getraceerd kan worden, zowel binnen als buiten de

organisatie. In de logistieke sector en de retailsector is dit al jarenlang gebruikelijk. Hierdoor zal de traceerbaarheid en daarmee de veiligheid van data toenemen.

Het is daarnaast van belang om te beseffen dat een veilige en gesloten wereld een illusie is. Dit heeft geleid tot de *zero trust principes*: door niemand te vertrouwen komt er een andere basis te liggen onder de vertrouwensrelaties tussen personen, systemen en organisaties. Steeds meer applicaties en vormen van cloudopslag zullen deze vormen van autorisatie gaan ondersteunen.

*Dance like nobody's watching,
encrypt like everybody is!*

*Werner Vogels
CTO Amazon*

4.3 De meerwaarde van data

De offensieve kant van data bestaat uit het planmatig en gericht ontwikkelen van data, waardoor de waarde toeneemt. Er zijn verschillende manieren waarop dit mogelijk is.

De eerste manier is om te kijken welke data nodig is om klanten beter te helpen of om processen efficiënter te laten lopen. Dit klinkt triviaal, maar toch kijken onvoldoende bedrijven op deze manier naar hun databehoeftes. Nog te vaak wordt data gemodelleerd vanuit de doelen van een systeem en te weinig vanuit het oogpunt van waardecreatie voorbij het systeem. Dus start met het in kaart brengen van de databehoeftes uit een systeemoverstijgend perspectief. Doe dit samen met een zo breed mogelijke vertegenwoordiging van IT en business. Kijk in de eerste ronde naar de 'wat' en niet naar de 'hoe'. Probeer alle ideeën vast te leggen en parkeer geen ideeën omdat de data er niet is, of omdat er technologische belemmeringen zijn. Selecteer de meest belovende en meest urgente databehoeften en ga vervolgens kijken naar hoe de data verkregen of gegenereerd kan worden. Er zullen ongetwijfeld lastig te realiseren behoeftes tussen zitten. Juist op deze punten kunt u al het talent van de IT-afdeling en de laatste AI- en cloudmogelijkheden inzetten om het toch op te lossen. Er is meer mogelijk dan u verwacht!

De tweede manier is actief data combineren tot nieuwe inzichten én deze actionable maken. Het combineren van data gebeurt natuurlijk al lang en wordt mogelijk gemaakt door oplossingen zoals datawarehouses en data lakes. Maar zoals in hoofdstuk 3 is omschreven slagen maar weinig organisaties erin om (grote) hoeveelheden data te combineren in hun primaire proces en op basis hiervan real-time klanten te helpen of processen te optimaliseren. Daag uw teams en leveranciers daarom uit om met oplossingen te komen voor dit vraagstuk.

Voor de derde manier gaat u buiten uw organisatie aan de slag. U gaat namelijk kijken welke partners, leveranciers of onbekende partijen misschien iets over hebben voor uw data. U zult ontdekken dat er altijd partijen geïnteresseerd zijn in bepaalde vormen van data. Dit kan voor u secundaire data zijn, die voor een andere partij in het primaire proces zit. Of vice versa: uw primaire data kan een ander wellicht op een ondersteunende manier helpen. Uiteraard moet uw data anoniem en gestructureerd zijn en moet u zich houden aan wet- en regelgeving.

Het loont ook om te onderzoeken welke data al beschikbaar is in open datasets. In steeds meer branches komen gestandaardiseerde datasets beschikbaar onder een open source licentie. Zoals Steward Brand al in 1984 onderbouwde: *Information wants to be free*. Data die voor grote groepen waardevol is, zal op termijn allemaal gratis en open toegankelijk zijn.

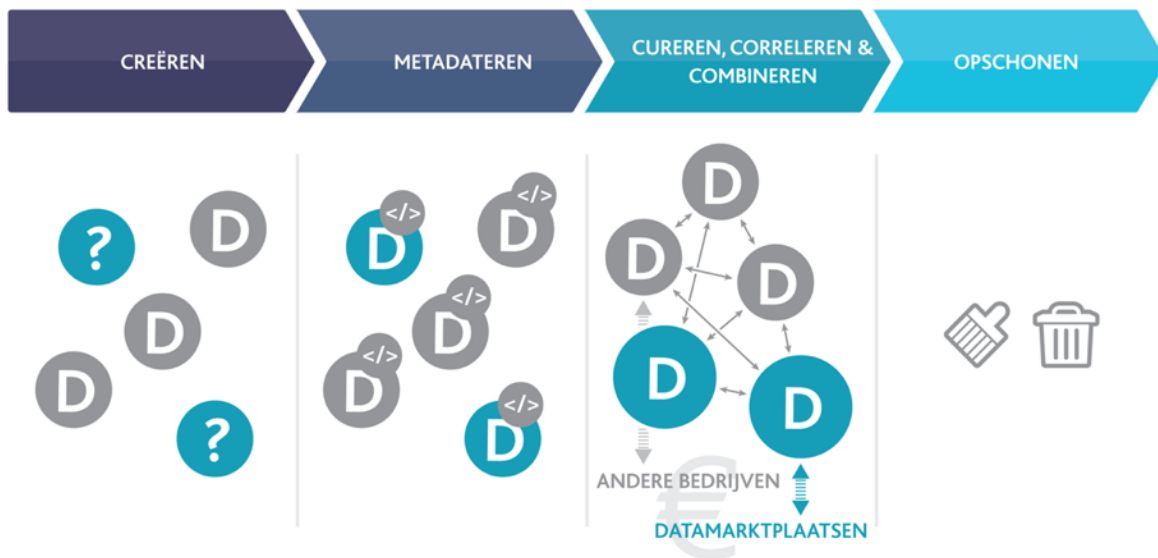
De opkomst van datamarktplaatsen

Afgelopen jaren zijn tientallen marktplaatsen voor data ontstaan. Op deze websites vinden vraag en aanbod van data elkaar. Data wordt er eenmalig of structureel verhandeld. Het is nog een jonge markt, dus volop in ontwikkeling. Er is nog geen standaard prijsmodel voor data is, dus data wordt altijd bij opbod en na onderhandeling verkocht.

Veel datamarktplaatsen specialiseren zich in bepaalde vormen van data:

- Internet of Things: verschillende vormen van sensordata. Bijv. <https://data.iota.org>
- AI data: goed geannoteerde en geclassificeerde datasets. Bijv. <https://bonseyes.com>
- Financiële data: financiële organisatie- en klantdata. Bijv. <https://www.quandl.com>

Datamarktplaatsen zijn een logische evolutie van de data-economie. Dus het loont om u hierin te verdiepen en te kijken welke mogelijkheden datamarktplaatsen kunnen bieden voor uw organisatie.



4.4 Planmatige aanpak

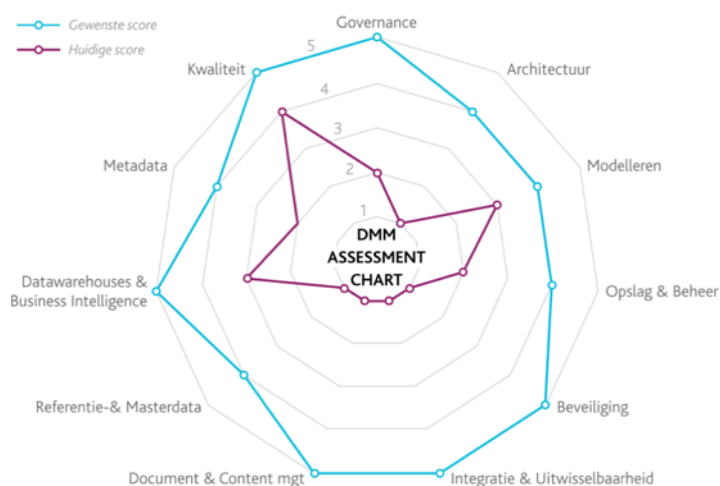
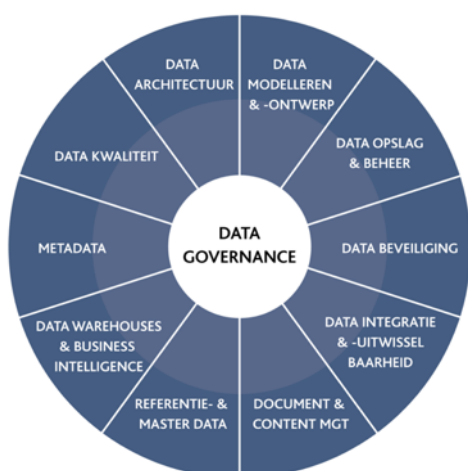
Er zijn verschillende modellen die kunnen dienen als startpunt voor een datastrategie. Een veelgebruikt model is het datamanagementframework van de Datamanagement Association¹⁾, ook wel afgekort tot DMBOK2. Dit model bestaat uit 10 aspecten van datamanagement die samen leiden tot een data governance structuur.

Dit model belicht zowel de organisatorische, proces- als systeemkant van data. Onderdeel van de DMBOK2 is een volwassenheidsmodel voor datamanagement. Op ieder van de 10 aspecten van het datamanagementmodel wordt gekeken naar de volwassenheid van de Planning (P), Control (C), Development (D) en Operations (O). De uitkomst van dit assessment is een sterdiagram waarin de huidige score en gewenste score geplotted kan worden.

Hoewel dit een gedegen en bewezen model is, is het naar onze mening nog te veel als defensief model ingestoken.

Het is een prima handleiding hoe u de goudvoorraad goed kunt bewaken (of desgewenst een Fort Knox kunt bouwen), maar gaat minder over hoe u goud vindt en goud kunt delven. Daarom adviseren we organisaties ook altijd om gericht op zoek te gaan naar meerwaarde van data om daarmee de bedrijfsdoelstellingen te bereiken. Dit start bij het doelbewust creëren van waardevolle nieuwe data. Door slimme metadatering zal data eenvoudiger te transformeren en hergebruiken zijn. Ook zal geavanceerde metadatering leiden tot hogere datakwaliteit, betere privacy en betere bescherming van data.

De groei naar volwassenheid op datagebied zou daarom niet alleen defensieve acties moeten bestaan, maar ook uit offensieve acties. Door dit planmatig aan te pakken zal ook uw organisatie meer waarde halen uit uw data en daarmee sneller, innovatiever en succesvoller zijn in de hedendaagse data-economie.



Afsluiting

Deze whitepaper is onderdeel van een serie whitepapers en blogs met het onderwerp 'data' als rode draad. Vanuit Luminis geloven wij dat goed gebruik van data leidt tot voorsprong binnen bedrijven. Daarom besteden wij veel tijd aan het praten met organisaties over data. De inzichten die wij uit deze gesprekken halen delen wij graag met u en andere organisaties.

Wilt u ook in gesprek met ons over data? Of wilt u meer weten over dataplatforms of datastrategieën? Kunt u hulp gebruiken bij uw datastrategie? Of heeft u een andere zienswijze? Wij komen graag in contact met u.

Over de auteurs

Hans Bossenbroek is oprichter en CEO van Luminis. Door continu voorop te lopen in technologie is Luminis onder zijn leiding gegroeid naar een organisatie van 200 medewerkers. Contact: hans.bossenbroek@luminis.eu

Martin van Mierloo is product marketing manager bij Luminis Technologies. Martin heeft ruim 20 jaar ervaring met de ontwikkeling van digitale strategieën voor B2B en B2C organisaties. Contact: martin.vanmierloo@luminis.eu

Over Luminis

Luminis is een software- en technologiebedrijf met vestigingen in Amsterdam, Apeldoorn, Arnhem, Eindhoven, Rotterdam en de UK. Vanuit deze vestigingen werken 200 professionals aan technisch hoogwaardige oplossingen voor klanten. Luminis levert diensten op het gebied van cloud-technologie, IT consultancy, dataOps en design. Daarnaast heeft Luminis een eigen data-as-a-service cloudplatform ontwikkeld: InformationGrid.

Voor meer informatie
<https://www.luminis.eu>
info@luminis.eu
+31 (0)88 58 64 600

Hoofdstuk 1

1. Reinsel, D., Gantz, H. Rydning, J. (2018) - *The Digitization of the World* - <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>

Hoofdstuk 2

1. Martin, JF. (2018)- *Unlocking success in digital transformations* — McKinsey&Company
2. Herzog, P., Leker, J. (2010) - *Open and closed innovation - Different innovation cultures for different strategies* — International Journal of Technology Management. 52. 322-343.
3. Laloux, F. (2015) - *Reinventing organizations*
4. Serenko, A., Bontis, N., Hardie T. (2017) - *Organizational size and knowledge flow: a proposed theoretical link* - Journal of Intellectual Capital 23 Oct 2017
5. McGowan, H. (2019) - *The X-shaped learner* - <https://www.thinkbeyond.co.nz/blog/x-shaped-learner/>
6. Schwaber, K. (2019) – *Evidence based management guide* – Scrum.org
7. Fowler, M. (2011) – *Polyglot persistence* - <https://martinfowler.com/bliki/PolyglotPersistence.html>
8. Fowler, M. (2019) – *How to move beyond a monolithic datalake to a distributed data mesh* - <https://martinfowler.com/articles/data-monolith-to-mesh.html>
9. Wardley, S. (2016) – *Wardley maps*

Hoofdstuk 3

1. <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
2. <https://www.cnbc.com/2018/07/12/amazon-to-take-almost-50-percent-of-us-e-commerce-market-by-years-end.html>
3. Evans, N., Price, J. (2018) - *Death by a thousand cuts: behaviour and attitudes that inhibit enterprise information asset management* - <http://www.informationr.net/ir/23-1/paper779.html>
4. <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/>
5. Kouloupoulos, T., Friend, D. (2018) – *The bottomless cloud* - <https://wasabi.com/bottomless-cloud/>
6. <https://fortune.com/2019/09/18/internet-cloud-server-data-center-energy-consumption-renewable-coal/>

Hoofdstuk 4

1. Cupoli, P., Earley, S., Henderson, D. (2014) – *DAMA-DMBOK2 Framework* – DAMA international