

Onafhankelijk consultant
Informatie Management
Ronald Damhof ontwikkelde
het Datakwadrantenmodel

“Maak datamanagement bespreekbaar in de hele organisatie”

Ronald
Damhof

ONAFHANKELIJK
CONSULTANT
INFORMATIE
MANAGEMENT

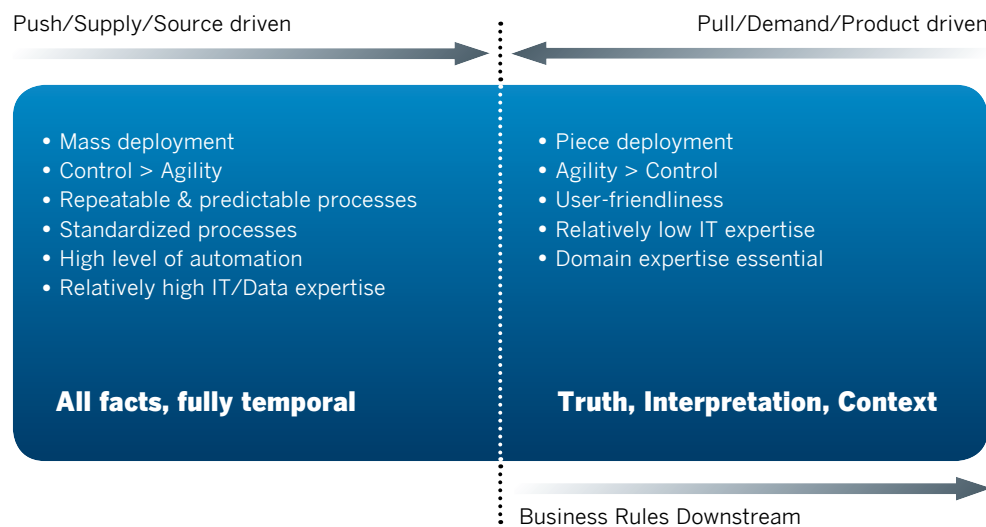
Het vakgebied datamanagement staat bol van het jargon. De meeste business managers hebben geen idee wat al die termen betekenen, laat staan dat ze hen helpen te snappen welke data nu precies welke waarde hebben en hoe ze hiermee moeten omgaan. Om een bedrijfsbrede discussie over data mogelijk te maken, ontwikkelde Ronald Damhof het Datakwadrantenmodel.

Damhof werkt als onafhankelijk consultant Informatie Management voor grote organisaties zoals Ahold, De Nederlandsche Bank, de Belastingdienst, Alliander en vele organisaties in de financiële sector en de gezondheidszorg. Het zijn data-intensieve organisaties die steeds beter beseffen dat de kwaliteit van hun werk in toenemende mate wordt bepaald door de kwaliteit van hun data. Maar hoe kom je vanuit dat besef tot een goede datastrategie? Een strategie die iedereen in de organisatie snapt, van de bestuurder in de boardroom tot aan de engineer bij ICT? Om het onderwerp datamanagement bespreekbaar te maken, ontwikkelde Damhof een kwadranten-model.

To push or to pull?

Damhof begint met het uitleggen van een begrip dat iedereen in het voortgezet onderwijs wel voorbij heeft zien komen: het klantorderontkoppelpunt. Dit heeft betrekking op de mate waarin de vraag impact heeft op het productieproces. Als voorbeeld noemt hij het maken van een luxueus jacht. Het productieproces begint pas als de order van de klant bekend is. Het klantorderontkoppelpunt ligt bij de start van het productieproces. Daartegenover staan bijvoorbeeld lucifers. Als een klant lucifers wil hebben, gaat hij naar de supermarkt en koopt hij ze. Als hij zwarte lucifers wil, heeft hij pech gehad. Het klantorderontkoppelpunt ligt helemaal aan het eind van het productieproces. De productie van een auto echter heeft een standaardgedeelte en een maatwerkgedeelte. De klant kan nog aangeven dat hij een specifieke kleur wil, leren bekleding, andere velgen et cetera. Het klantorderontkoppelpunt ligt ergens in midden van het productieproces. Damhof: "Ook met het maken van een rapport, dashboard of analytische omgeving is er sprake van een klantorderontkoppelpunt dat ergens in dat midden ligt."

The Data Push Pull Point



Hoe kom je tot een datastrategie?

Het klantorderontkoppelpunt splitst het productieproces in twee stukken: een push- en een pull-kant, ook wel een aanbodgedreven deel en een vraaggedreven deel. Pushsystemen zijn erop gericht schaalvoordelen te halen bij toename van volume en vraag, waarbij de (data)kwaliteit van het product gewaarborgd blijft. Aan de andere kant zijn er pullsystemen die vraaggedreven zijn. Verschillende type gebruikers willen vanuit hun eigen expertise en context de data kneden tot 'hun' product, hun waarheid.

Opportunistic of systematisch ontwikkelen?

Op de Y-as projecteert Damhof de dimensie ontwikkelstijl. "Daarmee bedoel ik: hoe ontwikkel je een informatieproduct? Dat kun je systematisch doen; dan zijn de gebruiker en de ontwikkelaar twee verschillende personen en pas je een defensieve governance toe, gericht op controle en compliance. Alles wat engineers hebben geleerd om software op gedegen wijze voort te brengen wordt hier in praktijk gebracht. Je ziet dit vaak bij gecentraliseerde, bedrijfsbreed toegepaste data, bijvoorbeeld financiële data en data die aan toezichthouders wordt gerapporteerd." Je kunt ook een opportunistische ontwikkelstijl gebruiken. "In dat geval zijn de ontwikkelaar en de gebruiker vaak één en dezelfde persoon. Denk aan de data scientist die wil innoveren met data, die analytische modellen wil maken en wil beproeven. Denk ook aan situaties waar snelheid van levering essentieel is. De governance is hierbij offensief, dat wil zeggen dat de focus ligt op flexibiliteit en aanpasbaarheid."

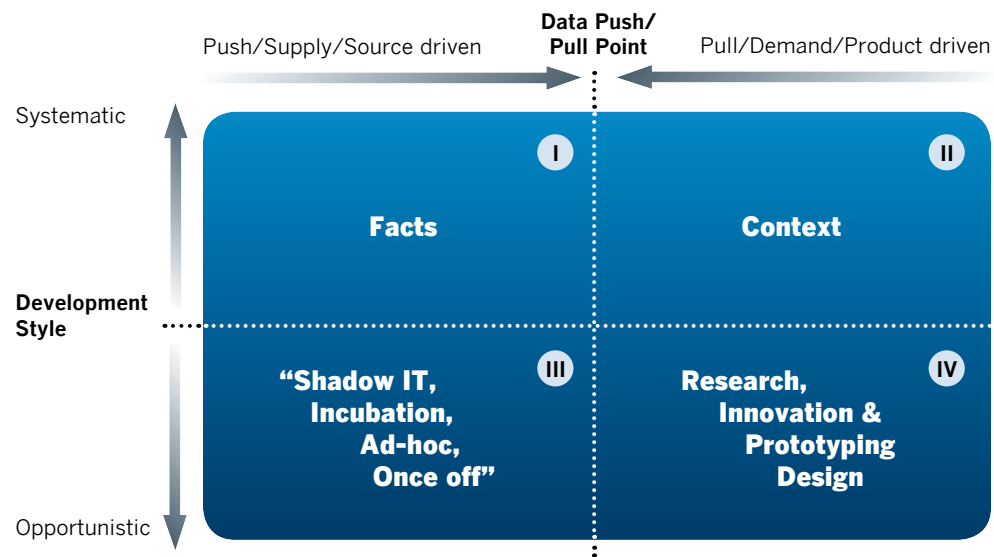
The Development Style



Datakwadrantenmodel

Gecombineerd leveren deze twee dimensies het volgende plaatje op.

A Data Deployment Quadrant



Damhof: "In kwadrant I vindt je de keiharde feiten. Deze data kan in haar volle ruwe omvang op begrijpelijke wijze ter beschikking worden gesteld aan de kwadranten II en IV. Data in kwadrant I worden geproduceerd door sterk gestandaardiseerde systemen en processen, zodat ze volstrekt voorspelbaar en repeteerbaar zijn."

Daar diagonaal tegenover, in kwadrant IV, staan data die worden gekenmerkt door innovatie en prototyping. "In dit kwadrant werken de data scientists, die eigenlijk maar drie dingen vragen: data, computerkracht en coole software." Steeds vaker worden er aparte afdelingen als Innovation Labs opgezet om data scientist volop te laten experimenteren en analyseren met de data. Met als doel innovatie. "Je hebt dit type datamanagement nodig om goede ideeën op te doen en te testen. Als iets werkt, dan is het vervolgens zaak om dit concept vanuit het vierde kwadrant omhoog te krijgen naar het tweede kwadrant, want pas als je de data systematisch kunt genereren en analyseren dan kun je er ook schaalvoordelen mee behalen in je business. Dan kun je ze bedrijfsbreed gaan toepassen."

Damhof gaat verder: "Te vaak spreek ik data scientists die hele gave inzichten verkrijgen in een soort sandbox omgeving. Maar ze vergeten of zijn niet in staat die inzichten ten gelde te maken in een productiesituatie. Ze krijgen hun inzichten maar niet van kwadrant IV naar kwadrant II. Hier komt governance om de hoek kijken." En daar ligt voor veel organisaties de grote uitdaging, weet Damhof. "Als je managers dit model uitlegt en vraagt waar hun prioriteit ligt, dan zeggen ze allemaal dat ze eerst het fundament op orde moeten maken, het eerste kwadrant. Maar vraag je waar ze hun geld nu in



"Nu organisaties steeds meer data genereren, kunnen ze hier niet langer slordig mee omgaan. Dit is dé tijd om datamanagement en de daarbij behorende governance goed op poten te zetten. Het Datakwadrantenmodel helpt daarbij"

Ronald Damhof

investeren, waar ze nu innoveren, dan is het vaak in het vierde kwadrant. Dat ze zich bezighouden met deze meer experimentele en exploratieve vorm van datamanagement is mooi, maar dat kan pas als je fundament klopt. Anders heb je een hypermodern toilet dat niet is aangesloten op het riool, dat wordt dus een hele vieze bende."

Vraag de gemiddelde data scientist wat hem het meeste tijd kost en zijn antwoord zal zijn om de data op het juiste kwalitatieve niveau te krijgen. Laat dit nu net het doel van kwadrant I zijn. "Een data scientist met krachtige analytische software, veel computerkracht én hoogwaardige data gaat pas echt een verschil maken."

Betrouwbaarheid versus flexibiliteit

"Managers willen nog wel eens roepen dat systemen betrouwbaar en flexibel moeten zijn, maar dit zijn communicerende vaten. En heel betrouwbaar en robuust systeem is minder flexibel. En bij een extreem flexibel systeem moeten de eisen ten aanzien van de betrouwbaarheid worden verlaagd", vertelt Damhof. "Het Datakwadrantenmodel maakt dit voor managers inzichtelijk. In kwadrant I gaat betrouwbaarheid boven flexibiliteit en in kwadrant II en IV gaat flexibiliteit boven betrouwbaarheid." Er zijn dan ook nogal wat verschillende soorten expertise en competenties nodig om data optimaal te kunnen uitnutten.

Expertise en competenties

Je ziet vaak dat organisaties van één persoon expertise en competenties vragen die het gehele kwadrant afdekken. Die mensen zijn er niet. Medewerkers in kwadrant I hebben een engineering profiel. Het zijn informatie- en data engineers, opgeleid in data-architectuur en data modellering. "Let op, dit is dus niet het klassieke IT profiel. Dit zijn engineers die model-gedreven kunnen ontwikkelen en de noodzaak van conceptueel en logisch modelleren goed begrijpen." Deze expertise is zeer schaars. Daar staan de kwadranten II en IV tegenover. "Daar zijn mensen nodig met een expertise in het betreffende business domein aangevuld met Business Intelligence-competenties en/of analytische competenties."

Feiten en waarheid

Damhof noemt kwadrant I van het model ook wel 'the single version of the facts'. Die feiten worden vervolgens ter beschikking gesteld van medewerkers in kwadrant II en IV. Zij kunnen daarmee hun

“Binnen organisaties gaat het vaak over ‘the single version of the truth’, maar die bestaat helemaal niet. Er is een ‘single version of the facts’ en er zijn meerdere ‘truths’”

Ronald Damhof

eigen waarheden maken. Omdat in deze rechterhelft van het model dezelfde feiten worden gebruikt om meerdere waarheden te creëren – afhankelijk van de context en de achtergrond van de gebruiker van de data – noemt Damhof deze helft wel ‘the multiple version of the truth’. Bedenk daarbij ook dat de ‘truth’ over tijd nogal eens verandert. Damhof: “Je hoort bedrijven vaak praten over ‘the single version of the truth’, maar die bestaat helemaal niet. Het hangt van de context, jouw zienswijze, achtergrondkennis en ervaringen af hoe je bepaalde feiten interpreteert.”

Kwadrant III

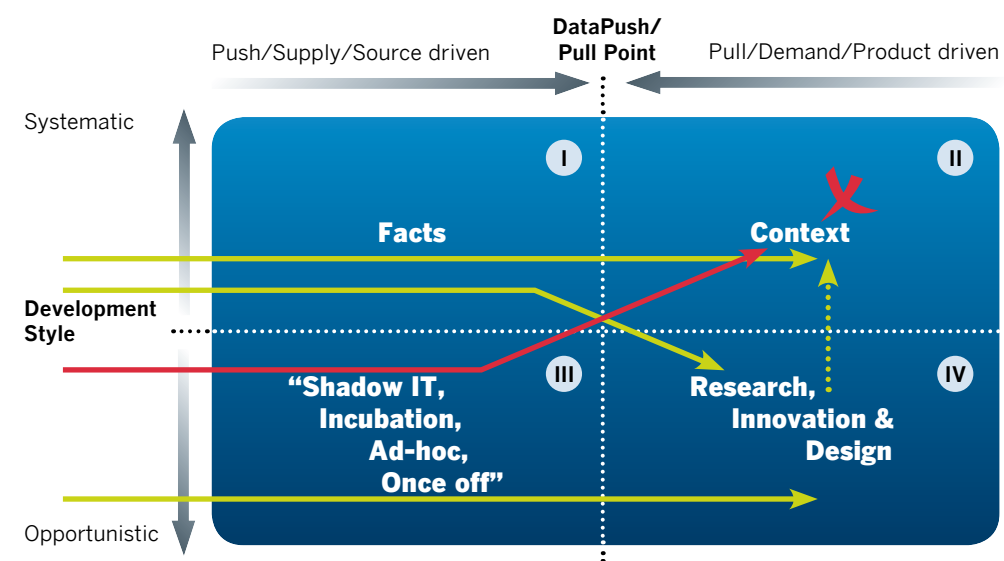
Het kwadrant dat weinig genoemd is, maar wel ongelooflijk belangrijk is, is kwadrant III. Het is het kwadrant van databronnen die niet onder governance staan. Een ad-hoc download die je doet van een Open Data Provider, een lijstje in Excel dat je wilt gebruiken, een set van contra informatie die je op een cd hebt ontvangen. Damhof: “Sterker nog, misschien wil je in dit kwadrant wel governed data uit kwadrant I combineren met een eigen dataset in kwadrant IV, prima!”

De reis door de kwadranten

Om waarde uit data te halen kun je verschillende bewegingen maken in het model. Je kunt van een op feiten gebaseerd datamanagement toegroeien naar een model waarin ook de context belangrijk is (van kwadrant I naar II). “Eigenlijk is dit de klassieke reis van ‘ontsluit data en maak een informatieproduct’, zegt Damhof. Dit werkt echter vaak slecht omdat dit traject uitgaat van bekende eisen en wensen van de gebruiker. “En dat weet die gebruiker echt niet van te voren.” Veel organisaties kiezen voor een meer agile gedreven vorm. Van kwadrant I naar kwadrant IV naar kwadrant II. Laat de medewerkers in kwadrant IV op iteratieve wijze met de data in kwadrant I/III komen tot een informatieproduct. En alleen als het van belang is om dit onder beheer te brengen, dan promoveer je het product naar kwadrant II.

Het is ook mogelijk om vanuit kwadrant III te bewegen naar IV. “Je hebt je eigen datasetjes en je wilt wat proberen? Prima...”, zo zegt Damhof. De enige beweging die een organisatie nooit en te nimmer mag maken is die van kwadrant III naar kwadrant II. “Want in dat geval gebruik je data waarvan je niet helemaal zeker bent omdat er geen goede governance op wordt toegepast op een manier die wel om die governance vraagt. Een voorbeeld is een compliance rapport voor de toezichthouder dat je wilt maken met data waar geen governance op staat. Dat moet je niet willen.”

How we produce, process variants



Maak datamanagement bespreekbaar

Damhof merkt in zijn dagelijkse praktijk dat zijn datakwadrantenmodel organisaties helpt om te praten over datamanagement. “Bij mijn huidige klant, De Nederlandsche Bank, hoor ik tegenwoordig regelmatig de kreten: ‘ik wil dit dataproduct graag van kwadrant IV naar kwadrant II brengen’, ‘we moeten eerst de data in kwadrant I zetten’, ‘degene die de data aanlevert is toch echt verantwoordelijk voor de data in kwadrant I’, ‘ik wil wat ruimte om data in kwadrant III tijdelijk op te slaan’. Iedereen weet dan wat dat betekent. Dat is nieuw, op die manier heeft de organisatie nog nooit over data nagedacht. En dat geldt eigenlijk voor vrijwel ieder data-intensief bedrijf. Organisaties praten al heel lang over ‘data als een asset’, maar in de praktijk gaan ze op een heel ongestructureerde manier met data om, waardoor ze die asset nooit ten gelde maken. Nu organisaties steeds meer data genereren, kunnen ze niet langer zo slordig ermee omgaan. Dit is dé tijd om datamanagement goed op poten te zetten. Het datakwadrantenmodel helpt daarbij.” ■