

Van data en informatie, naar kennis en wijsheid.

# Data Democratie, Data Mesh en Data Fabric

## Een drieluik voor data gedreven werken



Van data en informatie, naar kennis en wijsheid.

## Een 3-luik: Data democratie, data mesh en de data fabric architectuur.

In de moderne, data gedreven wereld is het vermogen om data om te zetten in concrete waarde cruciaal voor succes. Organisaties die hun data effectief inzetten, behalen niet alleen een concurrentievoordeel, maar optimaliseren ook hun bedrijfsprocessen en stimuleren groei. DIKW Intelligence helpt MKB-ondernemingen bij het ontwikkelen en implementeren van een datastrategie op maat, waarmee het volledige potentieel van data binnen en buiten de organisatie wordt ontsloten.

### Trends in Analytics

Sinds een paar jaar, met de opkomst van clouddiensten voor analytics en de ambitie om data gedreven te werken, krijgen drie nieuwe (architectuur) concepten meer bekendheid. Data Democratie, Data Mesh Architectuur en Data Fabric Architectuur. De afgelopen decennia zijn voornamelijk 2 en 3-tier oplossingen toegepast ook wel bekend als respectievelijk Kimball en Inmon of Data Vault. Zie kader (DWH 101).

Wat kunnen deze nieuwe trends in het data en analytics landschap betekenen voor organisaties? Hoe verhouden deze zich tot elkaar?

Voordat we in gaan op het beantwoorden van deze vragen een korte uitleg van de concepten. Een zoekopdracht op internet levert al snel vele beschrijvingen op.

Dit whitepaper geeft een beknopt overzicht van deze 3 met de belangrijkste kenmerken en inzicht wat dit voor uw organisatie kan betekenen.

#### Trends in Data & Analytics



# “Verwacht geen ander resultaat als je hetzelfde blijft doen”

## Data Autocratie

Een van de meest gehoorde klachten is het gebrek aan tijdige toegang tot betrouwbare data en de middelen om deze te verwerken tot data producten.

Van oudsher is het maken van nieuwe data producten een IT-feestje, vaak centraal georganiseerd en gedirigeerd. Een backlog, één team van data engineers, -scientists en -analisten die de business vraag vaak uit 2de hand vernemen, via een business informatie analist, alsof ze zelf niet kunnen praten met degene die de vraag stelt. En omdat het centraal georganiseerd is, wordt het ondersteunde systeem een Data Warehouse ook centraal ingericht.

In het kort een autocratie!

Dit heeft uiteraard voor- en nadelen, zoals.

### Voordelen:

- Afdwingen van uniforme data definities, wat de consistentie van rapportages ten goede komt.
- Data beschikbaar op een plek, voorkomen van gesleep met data door de organisatie. Wat inefficiënt is en in het kader van AVG regelgeving ook risico's met zich meebrengt.

### Nadelen:

- Zoals geschetst vormt de centrale organisatie vaak een bottleneck zeker bij toenemende vraag. Ook al wordt het team groter, dan vergt dat weer veel onderlinge afstemming.
- Veelal een one-site fits all oplossing, elke vraag complex of eenvoudig wordt op dezelfde manier opgelost.
- Het DWH team staat op afstand van de bronnen en soms ook tot de vraag. Er gaat veel tijd verloren in het leren kennen van de brondata en vervolgens aansluiten op de vraag tot een product.



## DATA DENIAL

You don't trust data, you find ways to undermine it and generally avoid

## DATA INDIFFERENT

You might see some data, but don't care what it says and have no need for it. You

## DATA INFORMED

You use data sometimes, when it supports your opinions or decisions you have already made.

## DATA DRIVEN

You seek to find clarity by searching for the data first, it informs and shapes the decisions you make.

### 1. Data Democratie

Data democratie is niet zozeer een systeem architectuur maar een manier van organiseren. Het adresseert de toenemende vraag naar het zelfstandig kunnen uitvoeren van data analyses, met als ultiem doel de **data-gedreven** strategie van een organisatie te realiseren.

Data-democratie bevordert bottom-up ideeën omdat geïnformeerde medewerkers beter bijdragen aan strategie en processen

In een data democratie heeft iedereen, dus ook niet data-specialisten, het recht om zelfstandig data producten te maken, maar net als in een democratische samenleving zijn er ook plichten.

Hieronder de basis spelregels:

1. Vrije toegang voor iedereen: Alle medewerkers krijgen toegang tot bedrijfsdata op basis van rol en noodzaak, zonder poortwachters of complexe aanvragen; data is het nieuwe "gelijk speelveld".
2. "You create it you own it." Als vervaardiger van een data product ben je ook verantwoordelijk voor het onderhoud, vastleggen van definities, registratie, tijdig beschikbaar stellen, gebruikers ondersteunen e.d..
3. Hergebruik van bestaande data producten, maak geen kopieën of vele varianten op hetzelfde data product.
4. Data geletterdheid verplicht: Organisatie investeert in opleidingen zodat iedereen data kan lezen, vertrouwen en toepassen in eigen werk. Iedereen neemt ook actief deel aan dit proces van leren en experimenteren.



# Welke architectuur past hier nu het beste bij?

## 2. Data Mesh Architectuur

De data mesh architectuur is gebaseerd op het idee dat data een product is en niet bijzaak van een proces in een systeem. Volgens dit idee wordt het tijdig leveren van kwalitatief betrouwbare data voor analyses en rapportages ook onderdeel van die bronsystemen. We noemen dit shift-left. Dus het CRM systeem is ook verantwoordelijk voor het maken van datasets voor klant-analyses en bijvoorbeeld ERP over productiviteit.

Het combineren van 2 of meerdere bronnen is de verantwoordelijkheid van de afnemer, deze maakt in feite weer een nieuwe data set. Die als zodanig weer beschikbaar wordt gesteld voor andere gebruikers. Volgens het “you create it, you own it” principe.

De kernprincipes:

1. Domeineigendom - Domeinteams nemen volledige verantwoordelijkheid voor hun data. Analytische en operationele data worden georganiseerd rond domeincontexten en eigendom verschuift van centrale IT naar domeinteams.
2. Data als product - Data wordt behandeld als product met externe afnemers. Domeinteams leveren hoogwaardige data die voldoet aan behoeften van andere domeinen, als openbare API.
3. Self-service platform - Centraal platformteam biedt domeinonafhankelijke tools voor het bouwen, uitvoeren en onderhouden van inter-operabele dataproducten. Domeinteams werken naadloos zelf.
4. Federatief beheer - Standaardisatie door beheergroep zorgt voor interoperabiliteit van alle dataproducten binnen organisatieregels. Creëert compliant data-ecosysteem.

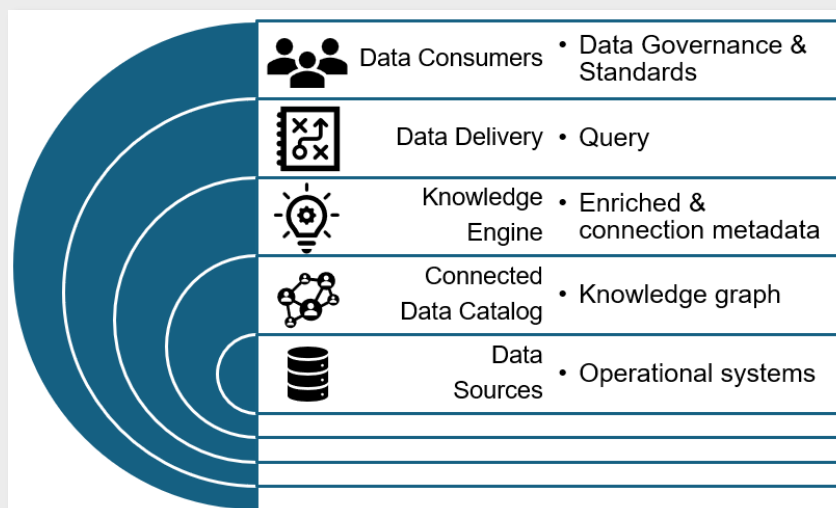
### 3. Data Fabric Architectuur

In tegenstelling tot het Data Mesh architectuur concept, wat meer organisatorisch van karakter is, is de Fabric architectuur veel meer een platform.

In het kort omschreven als: een metadata gedreven, centrale datamanagement laag die verspreide data logisch samenbrengt en als één omgeving presenteert (data federatie concept). Deze laag levert geïntegreerde services voor integratie, governance, kwaliteit, security en self service toegang over alle platformen heen.

De kernprincipes

1. **Gecentraliseerd datamanagement:** één logisch governancelaag (policies, security, kwaliteit) bovenop alle bronnen, ongeacht technologie of locatie (on-prem, multi-cloud, SaaS).
2. **Logische in plaats van fysieke centralisatie:** data mag in bronplatformen blijven; de fabric verbindt ze via een uniforme abstractielaag, vaak met virtuele views en slim ingerichte pipelines.
3. **Metadata- en kennisgraaf-gedreven:** uitgebreide catalogus, lineage en knowledge graph sturen integratie, data-discovery en automatisering (bijvoorbeeld het genereren van pipelines of policies).





## De kracht van de combinatie!

Op zichzelf staand zijn de besproken concepten geen vernieuwde ideeën, eerder een logische evolutie van architectuur, bestaande uit: organisatie, processen en technologieën.

Duidelijk is dat beide architecturen de complexiteit van data-management willen verminderen en de toegang tot data verbeteren. Tevens benadrukken beide dat data behandeld moet worden als een product, met kwaliteit, metadata, en governance.

Beide zijn geen specifieke technologieën, maar conceptuele raamwerken die je moet implementeren om data democratie succesvol te maken. Beide zijn ontworpen om te schalen in complexe, gedistribueerde omgevingen.

**Data Fabric** is geschikt als je een complexe, legacy omgeving hebt en centrale integratie en automatisering wilt.

**Data Mesh** is geschikt als je organisatorische verandering wilt en domeinteam meer verantwoordelijkheid wilt geven.

In de praktijk ziet DIKW meerwaarde in hybride oplossingen, waarbij Data Mesh principes worden gecombineerd met Data Fabric technologie voor integratie en metadata-beheer.

Dus samengenomen een sterke combinatie die de belofte van een data gedreven strategie kunnen waarmaken. Met extra nadruk op culturele shift naar organisatie brede data-adoptie in plaats van alleen technische architectuur.

# DWH 101

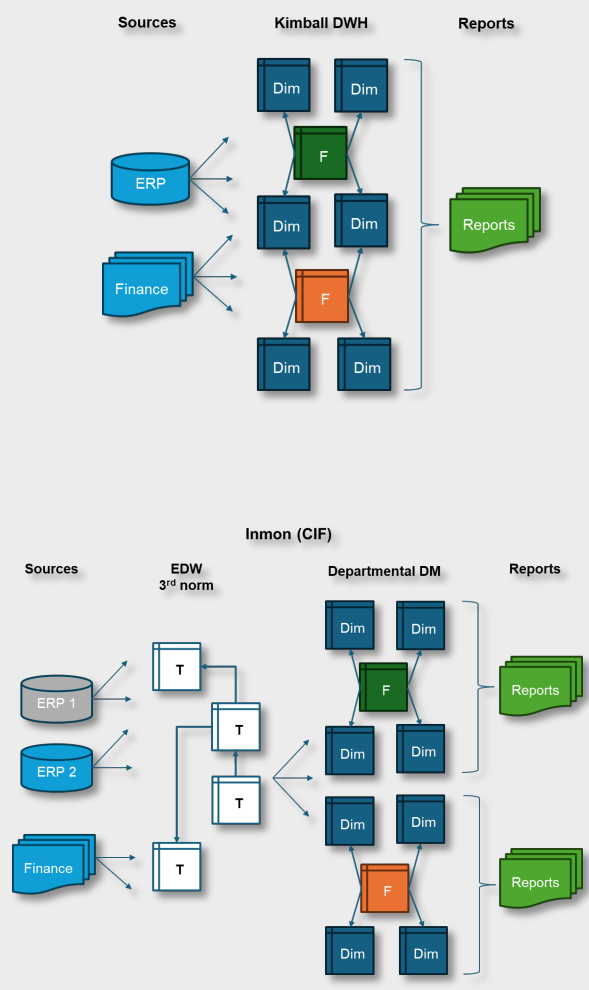
## Kimball, Inmon en Data Vault

### Achtergrond

Een eenvoudige vraag aan Perplexity levert het volgende op.  
*“Leg mij uit wat de verschillen en overeenkomsten zijn tussen de volgende 3 data architectuur concepten: respectievelijk 1) Kimball en 2) Inmon en 3) Data Vault?”*

Kimball, Inmon en Data Vault zijn alle drie benaderingen voor datawarehousing, maar ze verschillen vooral in modellering, volgorde van bouwen en waar je integratie en historie borgt. Ze hebben als overeenkomst dat ze enterprise-breed willen integreren en historie willen bewaren, maar met andere accenten en bouwblokken.

- Kern van elk concept
- **Kimball:** Dimensioneel model (sterren-/snowflake-schema) direct in de datawarehouselaag, georganiseerd rond businessprocessen met feit- en dimensietabellen en conformed dimensions. Gericht op snelle, incrementele levering per proces (busarchitectuur).
  - **Inmon:** Genormaliseerde enterprise data warehouse (3NF) als centrale “single version of the truth”, waaruit daarna dimensionele data marts worden gevoed. Architectuur bekend als Corporate Information Factory (CIF).
  - **Data Vault:** Modellering met Hubs, Links en Satellites, expliciet ontworpen voor integratie van veel bronnen en volledige historisering in een robuuste, auditable EDW kern. Dimensionele data marts worden meestal bovenop de Vault gebouwd (bv. als “Business Vault”).

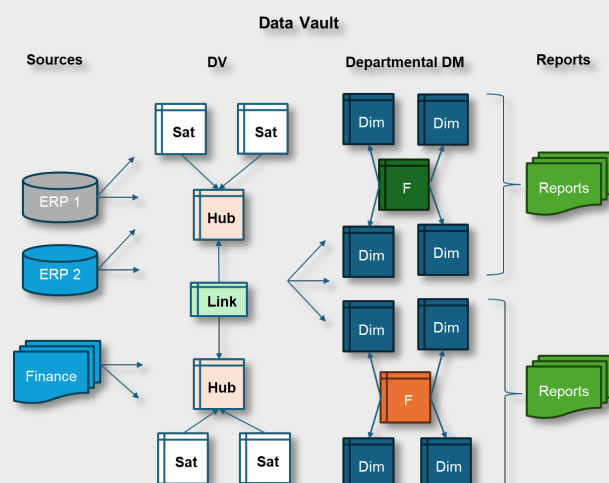


## DWH 101

### Kimball, Inmon en Data Vault

#### Overeenkomsten tussen de drie:

- Richten zich op een enterprise datawarehouse als centrale informatievoorziening, niet alleen losse marts.
- Ondersteunen historisering, zij het met andere technieken (SCD in Kimball, tijdgestempelde records in 3NF, Satellites in Data Vault).
- Gebruiken een meerlagige architectuur met bron → integratielaag → afnemerslaag (marts/semantic layer), al is de logische vorm per aanpak anders.



#### In de praktijk:

Gedreven vanuit een IT-perspectief is veelal een hybride “best-of-all-worlds” van de drie gemaakt, waardoor een complexe oplossing in modellering en ETL is ontstaan. Hierdoor zijn impact-analyses, wijzigingen en incident-afhandeling traag en duur.

- Inmon en Data Vault voegen veel tabellen en joins toe en vragen relatief veel upfront-ontwerp en integratiewerk; zonder scherpe afbakening van scope blijft het EDW jaren “in aanbouw” voordat de business echte inzichten krijgt.
- Terwijl Kimball veronderstelt dat er duidelijke proces-eigenaars en governance rond conformed dimensions zijn; als dat ontbreekt, ontstaan losse data marts, inconsistenties en hoge TCO.



## Adviseren, Implementeren en waarde creëren.

De informatievoorziening van uw organisatie omvat verschillende disciplines, zoals het inrichten van een gegevenspakhuis (of data warehouse), het inrichten van management- en stuurinformatie (rapportages en performance-management) en het inrichten van een analyse omgeving voor het 'minen' van uw data. De beschikbaarheid van informatie neemt exponentieel toe. De overvloed aan informatie vereist een goed door-dachte, continue en systematische aanpak, behandeling en analyse van deze informatie.

Data gedreven besluitvorming speelt in toenemende mate een belangrijke rol in de bedrijfsvoering. Organisations die in staat zijn om systematisch data en informatie om te zetten in kennis en wijsheid, zullen sneller en effectiever tot betere besluitvorming komen

In al deze disciplines kunt u DIKW Intelligence om advies vragen en ervan verzekerd zijn dat wij u kunnen helpen met het realiseren van uw doelstelling op ons vakgebied.

Dit alles met als doel om bedrijfsmatig, competitief en/of strategisch voordeel te behalen.

DIKW Intelligence  
Wattbaan 1 – 3439 ML Nieuwegein  
Tel. 030-3078743 – [info@dikw.com](mailto:info@dikw.com) – [www.dikw.com](http://www.dikw.com)