

Modelleren met BIReady en Data Vault concept

# Logisch business-model bepaalt generatie

Pieter Rambags

Met de komst van het datawarehouse-modelleringsconcept Data Vault en het datawarehouse management-systeem BIReady, zijn in de wereld die Business Intelligence heet interessante mogelijkheden ontstaan om projecten qua uitvoering anders op te zetten.

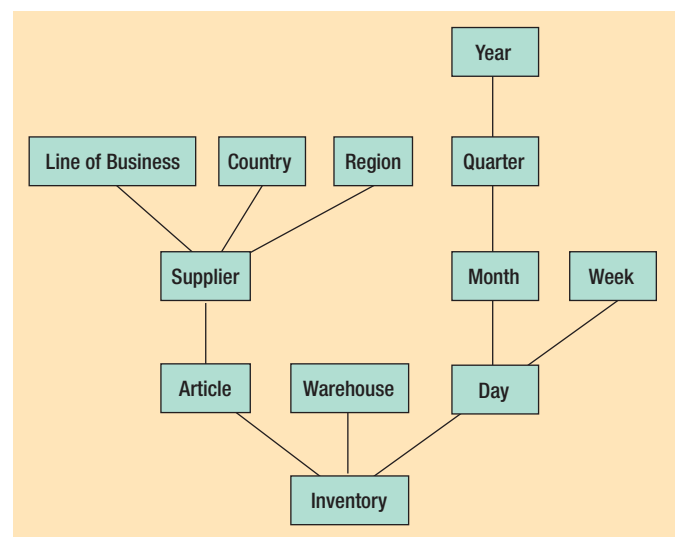
Waar voorheen ongeveer 80 procent van de realisatie van het datawarehouse zich afspeelde onder technenuten, ver van de eindgebruiker, en bestond uit het ontwikkelen van complexe datamodellen en laadscripsts, kan juist deze activiteit door toepassing van Data Vault en inzet van BIReady worden versneld. Hierdoor kan de aandacht van de consultants verschuiven naar het achterhalen en vastleggen van de business requirements en het integreren van de oplossing in de bedrijfsarchitectuur.

## Achtergrond

Iedere goede BI-consultant vraagt zich regelmatig af waarom zogenaamde Data Warehouse Management Systemen (DWHMS) alleen weggelegd lijken te zijn voor de allergrootste ondernemingen. Een DWHMS is een systeem waarbij het modelleren van het datawarehouse door gereedschap wordt overgenomen. Deze systemen genereren naast datawarehouse-modellen ook vaak elementaire laadscripsts om de data uit de bronnen over te hevelen naar het datawarehouse. Ondernemingen die zich deze gereedschappen niet kunnen veroorloven moeten langdurige trajecten met consultancy-bedrijven afleggen om de ontwikkeling hiervan ter hand te nemen. En daar blijft het dan niet bij. Na afsluiting van het project is er een dermate complexe omgeving gecreëerd, dat ook beheer ervan vervolgens vaak terecht komt bij deze consultancy-bedrijven.

Tot voor kort waren er enkele grotere spelers op de markt actief zoals SAP met haar oplossing Business Warehouse (BW) en Kalido. Eerstgenoemde wordt over het algemeen ingezet in grote projecten bij grote klanten, die de operationele processen al hebben geautomatiseerd, gebruikmakend van SAP R/3. Op basis van een vooraf gedefinieerde, dwingende methodologie wordt de implementatie ter hand genomen. Kalido wordt eveneens ingezet bij grote ondernemingen in grote projecten en in complexe

omgevingen. Voor deze systemen geldt dat ze bijzonder restrictief zijn ten aanzien van de inrichting. Daarnaast zijn deze varianten buitengewoon duur en complex. Dit zorgt ervoor dat het zowel voor klanten als consultants een enorme investering vergt om er voordelen mee te behalen. Hiermee wordt dus voornamelijk de markt van grote (multinationale) ondernemingen bediend, waar investeringen in datawarehousing al gauw in de miljoenen euro's kunnen en mogen lopen. Tot slot houden deze omgevingen er een eigen manier van modelleren op na. Omdat het 'black box' oplossingen betreft zou gesteld kunnen worden dat dit aspect niet zo relevant is. De realiteit wijst echter uit dat (met name in een SAP BW georiënteerde omgeving) juist deze interne modellering voor vele hoofdbrekers zorgt, vooral wanneer sprake dient te zijn van tactische en strategische managementrapportage. Aanpassing blijkt dan niet zo eenvoudig.



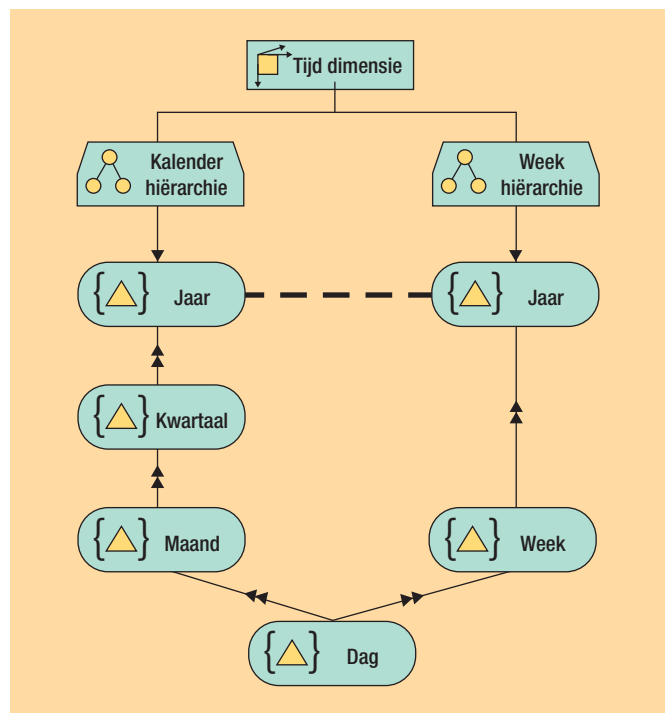
Afbeelding 1: Voorbeeld logisch model: ERD input.

## Data Vault

Op dit moment lijkt een best practice op het gebied van datawarehousing het modelleren op basis van het Data Vault concept, ontwikkeld door Dan Linstedt. Dit concept maakt het mogelijk een ont koppeling te realiseren tussen bron en datawarehouse, waarbij aspecten als traceability en auditability alsook het vasthouden van historie worden geregeld. Vanuit Data Vault kunnen traditionele sterschema's worden gegenereerd, die geschikt zijn voor (eindgebruikers)rapportage en analyse.

## Er wordt een aanzienlijke versnelling bereikt in de realisatie van een datawarehouse

Het concept Data Vault bevat een recept voor de implementatie van een logisch business-model, een model zoals dat daadwerkelijk met de eindgebruikers is afgestemd en dat weergeeft welke informatie nodig is en op welke wijze deze informatie geanalyseerd zou moeten worden. Verder is het modelleren op basis van Data Vault dermate generiek dat het bij correct gebruik eenvoudig is om het datawarehouse uit te breiden met nieuwe aandachtsgebieden. Dit laatste zorgt bij grotere datawarehouse-implementaties voor een versnelling van de ontwikkeling van vervolgfases. Vooral als in eerdere fasen is geïnvesteerd in de ontwikkeling van toolkits of generatoren die generatie van het datawarehouse-model voor rekening nemen op basis van input van logische, bedrijfsspecifieke structuren. Nadeel blijft echter het gegeven dat het datawarehouse nog handmatig moet worden



Afbeelding 2: Voorbeeld logisch model: Adapt-schema van tijddimensie.

uitgebreid of aangepast en dat ook de extractie-, transformatie- en laadprocessen (ETL) zelf moeten worden bedacht en ontwikkeld.

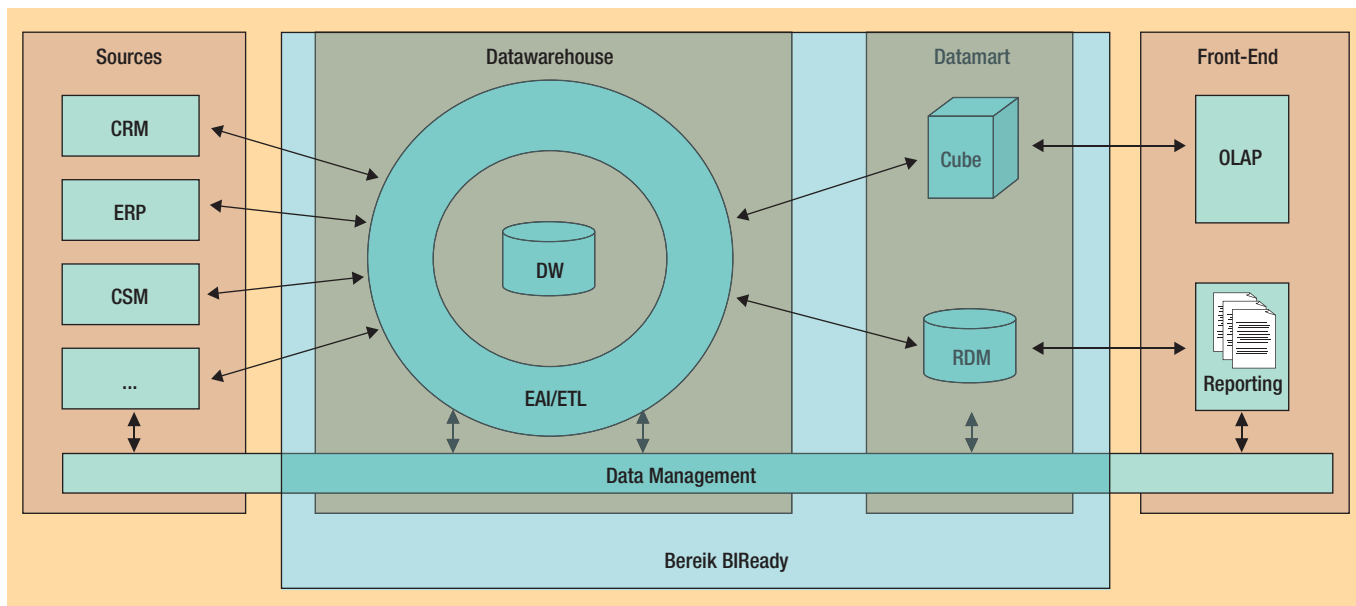
## BIReady

Al weer ruim een jaar geleden is door dr. Harm van der Lek het product BIReady geïntroduceerd. Het heeft reeds gedemonstreerd dat het mogelijk is een datawarehouse-model met bijbehorend ETL-processen te genereren puur en alleen op basis van een logisch business-model en mapping van dit business-model naar brondata. Hiermee wordt een aanzienlijke versnelling bereikt in de realisatie van een datawarehouse. Dat het product in tegenstelling tot eerdergenoemde oplossingen als SAP en Kalido ook geschikt is voor het midden- en kleinbedrijf, blijkt uit het gehanteerde licentiemodel dat rekening houdt met de complexiteit van het business-model en dus met de complexiteit van de onderneming.

Het concept van BIReady is eenvoudig te bevatten. Dat komt deels doordat de complexe algoritmie onzichtbaar voor de consultant is verwerkt in het product. Door de gebruiker dient een helder business-model te worden gedefinieerd. Dit business-model is een ERD-diagram (Entity Relationship Diagram) dat het volgende weergeeft:

- welke meetgegevens gerapporteerd dienen te worden (bijvoorbeeld omzet, kosten, KPI's);
- welke beschrijvende elementen daarbij zichtbaar moeten worden gemaakt (de dimensies zoals bijvoorbeeld product, klant, tijd, geografie etcetera);
- welke navigatiepaden toegankelijk moeten zijn (de hiërarchieën of drill-down paden);
- hoe de gebruiker wenst om te gaan met historie (de mogelijkheid om te rapporteren op basis van historische verbanden in de data).

Let wel: het betreft hier een zuiver logisch (!) model dat een licht werpt op de behoefte van de eindgebruiker (met een logisch model wordt een techniekonafhankelijke beschrijving van de werkelijkheid vanuit gezichtspunt van de ondernemer bedoeld). In eerste instantie hoeft dit model geen rekening te houden met technische beperkingen zoals: het wellicht niet beschikbaar zijn van bepaalde gegevens; beperkingen van front-end producten. Dit logische model wordt aan BIReady gegeven, gevoed als het ware, waarna in BIReady grafisch de vereiste modelattributen kunnen worden gekoppeld aan de brondata-attributen, zodoende definiërend waar later de gegevens vandaan moeten komen. Zodra dit gedaan is bouwt BIReady op basis van de spreekwoordelijke 'druk op de knop' zowel het datawarehouse-schema als de bijbehorende procesinstructies die het laadproces voor hun rekening nemen. Waar nu de meeste tijd in gaat zitten is het achterhalen van de gebruikersbehoeften en het maken van het business-model, inclusief de genoemde mappings. Het creëren van het datawarehouse-model met bijbehorende laadprocessen is nog slechts een kwestie van seconden! Het gegenereerde datawarehouse-model voorziet voorts in extra functies ter facilitering



**Afbeelding 3:** Referentiemodel BI-architectuur.

van de vereisten met betrekking tot het mogelijk maken van het vasthouden en onderhouden van diverse vormen van historische data.

Omdat management-informatie onderhevig is aan voortschrijdend inzicht en wijzigingen van het logische business-model derhalve eerder regel dan uitzondering zijn, voorziet de adjust-functie in BIReady in het reorganiseren van het datawarehouse-model, zodat met behoud van reeds eerder opgeslagen gegevens het onderliggende model geschikt wordt gemaakt voor de veranderde situatie.

### Proof of concept

Gezien het innovatieve karakter van BIReady is het begrijpelijk dat veel organisaties na introductie ruim een jaar geleden ervoor terugdeinsden als eerste een implementatie te overwegen. Het betrof toen nog een jonge firma met een jong product, iets waar bedrijven zich over het algemeen niet aan wagen. Doordat Nippur, in de zoektocht naar best practices, de toegevoegde waarde van BIReady wist te waarderen, kon een van haar klanten bereid gevonden worden een proof of concept uit te voeren. Nippur had reeds een studie naar het business-model en de bijbehorende sourcing van het datawarehouse afgerond, waardoor in dit geval snel gestart kon worden met de inrichting van BIReady en de evaluatie van vooraf gedefinieerde acceptatie-criteria.

Opmerkelijk aan dit traject was dat de aandacht inderdaad meer verplaatst werd naar het overleg met de gebruikersorganisatie om een zo eenduidig en volledig mogelijke definitie van het business-model vast te kunnen stellen. Het voeden van BIReady en de generatie van het datawarehouse-model en bijbehorende proces-instructies ging, ondanks de learning curve die de consultants uiteraard moesten doorlopen, redelijk soepel. Na twee weken kon deze proof of concept worden afgerond en besloot de klant tot de aanschaf en verdere implementatie.

Na driekwart jaar is gebleken dat hetgeen de proof of concept al beloofde is waargemaakt. In de consultancy-opdracht is vooral aandacht naar de business case, de infrastructurele en meer algemeen architectuurtechnische eisen en de logische requirements gegaan. De daadwerkelijke ontwikkeling van het datawarehouse is door BIReady uitstekend verwezenlijkt en zorgt niet voor verrassingen in de planningen.

## Na twee weken besloot de klant tot aanschaf en verdere implementatie

Tijdens de consultancy-fase kon kort worden geschakeld met BIReady waar het vereiste of gewenste verbeteringen van het product betrof. Vanwege de goede en kennisintensieve relatie tussen beide ondernemingen was een half woord genoeg om issues snel van tafel te krijgen. Dit heeft er toe geleid dat het product daardoor verder aan kwaliteit heeft gewonnen. Natuurlijk blijft er nog wat te wensen over. Zo is (nog) niet voorzien in aspecten als: ondersteuning van meerdere ontwikkelaars; professionele scheiding van ontwikkeling, acceptatie en productie. Echter deze beperkingen kunnen deels omzeild worden door procedurele afspraken te maken in de ontwikkelteams. Het product heeft inmiddels voldoende bewezen open te zijn waardoor, indien noodzakelijk, voor en na BIReady in de informatielogistiek ingegrepen kan worden om gewenste doelen te realiseren.

### Pieter Rambags

Drs. P. Rambags (pieter.rambags@nippur.nl) is managing partner van Nippur.