

CDM waarborgt scheiding bedrijfsregels en data

# Business rules in toekomstvast datawarehouse

Erik Fransen, Antoine Stelma, Carlos Kruiner

**In DB/M nummer 5 van dit jaar is inzicht gegeven in de ontwikkelmethodiek CDM (Centennium Datawarehouse Methodiek) waarmee toekomstvast datawarehouses worden gemodelleerd en automatisch gegenereerd. Dit vervolgartikel onderzoekt de relatie tussen Business rules en Business Intelligence.**

De ontwikkelmethodiek CDM is opgebouwd uit drie pijlers voor Structureren, Modelleren en Genereren van het datawarehouse: respectievelijk BI-referentiearchitectuur, Data Vault en Template Based Development. Zie DB/M 5, pag. 22 voor de BI-referentiearchitectuur. Business rules zijn binnen de toepassing van Business Intelligence onder andere nodig voor het inhoudelijk definiëren van de Key Performance Indicatoren (KPI) van de organisatie. Veelal leidt het tot stand komen van de KPI's tot het ontwikkelen van een datawarehouse (soms met analytische doeleinden) waarin de business rules meehelpen om de dimensies, hiërarchieën en meetwaarden te definiëren. Toch worden deze business rules meestal niet expliciet gedefinieerd en/of in samenhang opgesteld en centraal toegankelijk opgeslagen. In het beste geval zijn business rules een restproduct van trajecten zoals masterdata-, change-, kennis- of datakwaliteitmanagement. Business rules worden al sinds jaar en dag toegepast bij het ontwikkelen van BI-rapporten en -analyses, maar omdat deze vaak verborgen (lees: impliciet) zijn, verweven zijn in (complexe) softwareprocedures en/of niet gestructureerd zijn opgesteld en vastgelegd, is het beheren van de regels complex. Het ontbreken van expliciete en beheersbaarheid leidt tot de volgende uitdagingen:

- Stuur- en verantwoordingsinformatie is nauwelijks of niet tijdig voor handen. De complexiteit van de business rules in BI-systemen zorgt ervoor dat beslissingen niet kunnen worden genomen omdat procesgegevens niet op een juiste manier naar het BI-systeem doorkomen. Denk hierbij aan wijzigingen als gevolg van marktbevingen, reorganisatie en fusies, business process re-engineering en het conformeren aan nieuwe wetgevingen door de business;
- Het verkrijgen van een overzicht van alle relevante business rules is onmogelijk vanwege de vele locaties waar ze zijn toegepast;
- De gebruikers kunnen dan ook niet de relaties en patronen

tussen de verschillende business rules identificeren en kunnen hierdoor niet de impact bepalen bij wijzigingen;

- De business rules zijn hierdoor implementatie- c.q. platformafhankelijk gemaakt. Dit betekent dat gebruikers afhankelijk zijn van de mogelijkheden van het platform en de grillen van de bijbehorende leverancier.

Business rules in BI zijn beslissingsondersteunend en worden vaak als volgt toegepast: het inhoudelijk definiëren van de KPI's van de organisatie; het vergelijken van procesresultaten en what-if scenario's; de verrijking, integratie en aggregatie van transactiegerichte procesgegevens.

Het belang van business rules is voor het management: verbeterd inzicht in de besluitvormingsprocessen doordat beslissingsprocessen zichtbaar en inzichtelijk worden; vergrote flexibiliteit in bedrijfs- en informatieprocessen zorgt voor een betere reactie (lees verkorte doorlooptijd) van de implementatie van een gewijzigde strategie; lagere Total Cost of Ownership (TCO) als geheel vanwege lagere kosten in het beheer van de business rules.

Het belang van business rules is voor de business: grotere transparantie en toegankelijkheid van de business rules voor de gebruikers; beleidsregels, wetgeving en contractuele verplichtingen worden geïnterpreteerd op een betrouwbare, eenduidige en herhaalbare wijze; binnen gestelde kaders eenvoudig aanpasbaar met als gevolg een verkorte doorlooptijd bij wijzigingen, dus snellere opleveringen.

Het belang van business rules is voor business analisten: het voorkomen van dubbelingen in business rules; het bepalen van de onderlinge relaties tussen business rules; het vergelijken van berekeningen volgens verschillende versies van dezelfde business rule in de tijd vanwege versiebeheer.

Het belang van business rules is voor ontwikkeling en beheer: vereenvoudigd ontwerp bij ontwikkeling van BI-systemen

(datamodellen, ETL-stromen, semantische lagen en informatie producten); uitgaande van het feit dat business rules wijzigen, wordt de flexibiliteit van de uitvoerende BI-modules vergroot bij het gebruik van een aparte business rule-repository; hergebruik van business rules; separaat testtraject van de business rule-module en daarmee ook vereenvoudigd testtraject van de ETL; lagere TCO bij ontwikkeling en beheer van alle betrokken softwaremodules in de dataverwerking en presentatie van informatieproducten.

## Business rules en het datawarehouse

De BI-referentiearchitectuur kent vijf verticale lagen en vijf horizontale lagen. De vijf verticale lagen zijn beschreven in DB/M 5 2010. In dit artikel gaan we verder in op de laag 'Beheer Business Rules', waarbinnen de business rules worden gedefinieerd. We hanteren overigens de definitie van de business rules zoals die is gedefinieerd door de Business Rules Group ([www.businessrulesgroup.org](http://www.businessrulesgroup.org)): "een business rule is een bewering welke (een deel van) de business beschrijft of beperkt". De beschrijving en/of beperking in de vorm van een bewering via een regel is in feite een modellering van wat zich afspeelt in de business processen en dus bij uitstek geschikt om opgenomen te worden in de BI-referentiearchitectuur. Binnen CDM wordt de architectuur in combinatie met Data Vault gebruikt om de uitvoering van business processen optimaal te weerspiegelen in het datawarehouse. Het datawarehouse wordt derhalve een centrale locatie waarin het verloop van alle business processen terugvindbaar is: ook de beweringen, afleidingen en beperkingen die im- of expliciet in de business worden toegepast, worden zichtbaar en historisch opgeslagen.

Naast het expliciet ontwerpen van het datamodel (binnen CDM via Data Vault) voor de opslag van de (historische) afspiegeling van het verloop van de business processen, het definiëren en vastleggen van alle metadata (context van de data), kiezen we binnen CDM ook voor het expliciet beschrijven en vastleggen van business rules. De business rules worden daardoor niet

alleen expliciet maar ook zodanig gedefinieerd en gepositioneerd dat deze eenvoudig onderhoudbaar zijn door gebruikers. De business rules worden bij voorkeur declaratief gedefinieerd (via een verzameling van beschrijvende beweringen of beperkingen). Het daadwerkelijk uitvoeren en toepassen van de business rules wordt uitgevoerd via een algoritme. Het algoritme kan ontwikkeld worden in een procedurele programmeertaal, waarbij gewaarborgd moet zijn dat er geen directe afhankelijkheid is tussen de regels en de procedure. De regels veranderen namelijk door de tijd heen: in het kader van beheersbaarheid dient er een strikte scheiding te zijn tussen de regels en het toepassingsalgoritme.

De zogenaamde business rules engines zijn hier expliciet voor ontwikkeld: in deze softwaretoepassingen worden business rules strikt gescheiden, zowel functioneel als technisch, van het toepassingsalgoritme. Op deze wijze wordt het mogelijk om te kiezen tussen toepassingsalgoritmen zonder de business rules aan te passen. Een voorbeeld is het toepassingsalgoritme Forward Chaining: op basis van de geselecteerde data worden regels 'afgevuurd' die vervolgens leiden tot conclusies. Deze conclusies kunnen ook weer voorwaarden zijn van een regel, waardoor die regel wordt afgevuurd, net zo lang totdat een eindsituatie is bereikt. Deze vorm van inferentie is bij uitstek geschikt om vanaf de brondata regels toe te passen die zowel worden toegepast in het datawarehouse (registratielaag) alsook in de datamarts (publicatielaag).

Daarbij dient onderscheid gemaakt te worden tussen toepassing van de business rules in batch en real-time. De batchverwerking van business rules betekent niets anders dan het toepassen van alle geldende regels bij het (nachtelijk) vullen (in het geval van CDM genereren) van het datawarehouse en de afgeleide datamarts. Real-time toepassing van business rules is mogelijk indien in het bronsysteem een nieuwe klant wordt opgevoerd, via het voorportaal in het datawarehouse wordt ingeladen en vervolgens op basis van een aantal kenmerken, door middel van de geldende business rules, wordt gecategoriseerd. Deze toegekende cate-

Type business rule	Omschrijving	Voorbeeld	Fase toepassing
Expressie	Vervolgconclusies die getrokken worden n.a.v. primaire business rules die niet worden vastgelegd in bronsystemen. Business rules met deelresultaten die niet in het brondomein worden vastgelegd maar waar wel een informatiebehoefte voor is.	ALS StatusFactuur = 'ja' EN Betaalgedrag = 'goed' DAN geaccepteerd = 'ja'	3
Verzameling	Een specifieke verzameling/set aan gegevens samengesteld door middel van een view, select-statement of tabel.	<TABEL> SELECT [<expressie>, [<expressie>]...] FROM <TABEL> <VERZAMELING>	1, 2, 3
Berekening	Geeft een waarde voor elke resultaatrij.	SUM, CASE WHEN, IF ... StatusFactuur = CASE WHEN datum_voldaan IS NULL THEN 'ja' ELSE 'nee' END	3
Filter	Filtert een verzameling op basis van twee gelijksoortige expressies.	StatusFactuur = 'ja'	3, 4, 5
Integratie	Combineert twee gegevensverzamelingen op basis van twee gelijksoortige expressies.	"SAME AS" link binnen Data Vault	2
Hiërarchieën	Denk hierbij aan de hiërarchische structuren in dimensies in de presentatielagen (datamarts) van datawarehouses of kubussen en/of draaitabellen.	<ul style="list-style-type: none"> <li>• JAAR <ul style="list-style-type: none"> <li>• MAAND <ul style="list-style-type: none"> <li>• DAG <ul style="list-style-type: none"> <li>• UUR</li> </ul> </li> </ul> </li> </ul> </li> </ul>	3, 4, 5

Afbeelding 1: Typen business rules in de BI-referentiearchitectuur

gorie wordt dan direct opgeslagen in het datawarehouse bij de betreffende klant.

Een ander toepassingsalgoritme is Backward Chaining: hier wordt uitgegaan van een hypothese die getoetst moet worden op basis van de business rules. Deze vorm zien we terug bij analytische toepassingen in BI zoals toekenning van kredieten of verzekeringen. In een cross-selling proces zou bijvoorbeeld getoetst kunnen worden of een klant in aanmerking komt voor een korting op een nieuwe verzekering. De business rules engine gaat nu redeneren vanaf de conclusie 'Nieuwe Verzekering met Korting = Ja' en gaat vervolgens zelfstandig op zoek naar de voorwaarden die hierbij horen, op basis van de verzameling regels die bekend is. Op deze manier worden andere voorwaarden waaraan voldaan moet worden gevonden en getoetst (via een prompt naar de gebruiker of via data in het datawarehouse of datamart), bijvoorbeeld 'aantal huidige verzekeringen GROTER DAN 4'. Op deze wijze kan snel en effectief aan cross-selling worden gedaan op basis van business rules in het datawarehouse.

Business rules worden primair toegepast tussen de volgende verticale lagen (uitgaande van de BI-referentiearchitectuur):

1. Van Bron naar Voorportaal. Geen toepassing van 'echte' business rules. Slechts de standaard databaseconversies zoals datumconversie, tekensetconversie conform de afspraken binnen Data Vault worden hier toegepast;
2. Van Voorportaal naar Registratielaag. De focus in deze stap ligt op kwaliteitscontrole; zijn de business keys bekend, zijn er andere afwijkingen in de data vast te stellen, kruiscontroles aantallen dataregels, succesvolle afloop van integratie van data enzovoort. Business rules worden hier slechts toegepast op het punt van integratie: welke regels hanteert de business bij het combineren van data uit verschillende bronnen? Deze regels, om bijvoorbeeld klanten uit diverse bronnen te verbinden in het datawarehouse, worden gedefinieerd via business rules, temeer als er sprake kan zijn van relatief complexe integratieregels;
3. Van Registratielaag naar Publicatielaag. De Registratielaag vormt de kern van het datawarehouse: hierin worden alle gegevens binnen de context van de onderliggende business processen historisch en feitelijk vastgelegd. De Publicatielaag is het venster van de gebruikers: via deze laag kijken zij naar

de business, op hun eigen wijze en in context. In de registratielaag bevindt zich in het algemeen de bulk van de business rules die worden toegepast en waarvan de resultaten zichtbaar worden in de publicatielaag. Voorbeelden zijn complexe berekeningen, categorisatie van data, kwalificatie van data in risicogroepen, selectie van data op basis van doelgroepregels, regels voor laden van dimensies, hiërarchieën en feiten;

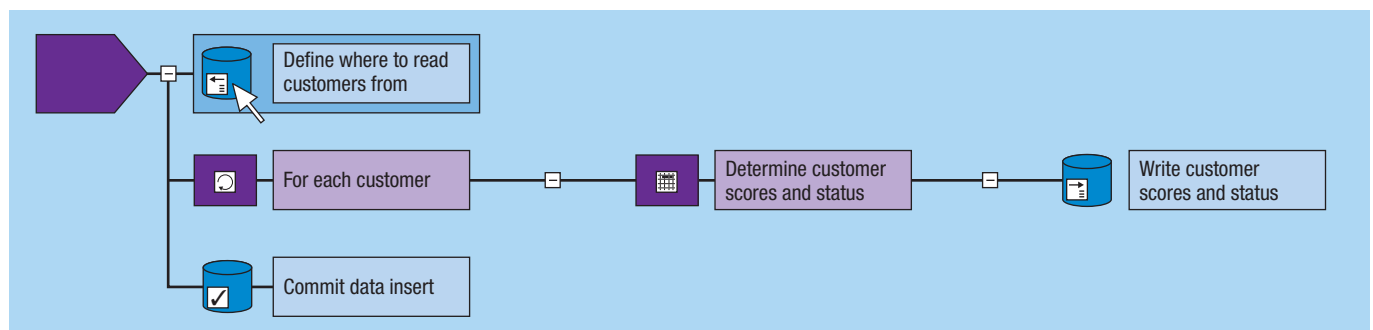
4. Van Publicatielaag naar Semantische laag. Bij voorkeur worden er geen business rules toegepast tussen deze twee lagen: de semantische laag vormt uitsluitend de technische toegangspoort van de rapportage- en/of analyse-omgeving tot de publicatielaag;
5. Van Semantische laag naar Distributielaag. De Distributielaag verzorgt de rapportages en analyses voor de eindgebruiker. De overgang van Semantische laag naar de Distributielaag kenmerkt zich bij voorkeur door het ontbreken van business rules, aangezien deze reeds in laag 1, 2 of 3 zijn gedefinieerd. Slechts expliciete datafilters worden toegepast in laag 5.

### Technische implementatie

Het ligt voor de hand om de business rules binnen de BI-referentiearchitectuur zowel functioneel als technisch op een centrale locatie op te slaan, we doen dat tenslotte ook al voor onze bedrijfsbrede data, namelijk via opslag in het datawarehouse. Uiteraard willen we daarbij maximale flexibiliteit in de toepassing. Daarbij hebben zowel de data als de metadata als de business rules een verschillende betekenis op een verschillend moment. Eenvoudig gezegd: we willen data van vorig jaar ook tegen de geldige business rules van nu kunnen berekenen. In deze 'gouden driehoek' beschrijven metadata de context van business rules en de data.

Business rule metadata beschrijven onder andere de eigenaar, wie de laatste wijziging heeft gedaan, voor welke doelgroep de regels van toepassing zijn en waar in de BI-referentiearchitectuur de business rules toegepast worden.

Het zwaartepunt van de business rules ligt in de overgang van de Registratielaag naar de Publicatielaag (positie 3). De opslag van de business rules in deze fase vindt plaats in de Registratielaag, waarbij Data Vault als modelleringsprincipe wordt gehanteerd. De voordelen hiervan zijn evident: er is één centraal opslagpunt van alle business rules die van toepassing zijn tijdens de over-



Afbeelding 2: Procesmodel (Bron: [www.visual-rules.com](http://www.visual-rules.com)).

gang van Registratie- naar Publicatielaag; de business rules worden opgeslagen in een Data Vault: alle historie van de business rules wordt vastgelegd en auditeerbaarheid is gegarandeerd; door middel van toepassing van Hub en Satellieten worden business rules uniek geïdentificeerd en compleet beschreven via metadata; de business rules gebruiken bestaande Hubs en Satellieten in de Data Vault om de gedefinieerde voorwaarden (StatusFactuur = 'Ja' EN Betaalgedrag = 'goed') te toetsen. Afleidingen (bijvoorbeeld Geaccepteerd is 'ja') worden opgeslagen in de Publicatielaag en zijn volledig regenereerbaar. Historie van toepassing van business rules is op deze wijze altijd gewaarborgd. Door de flexibele methodiek kan Data Vault op verschillende manieren ondersteuning bieden aan het gevraagde. Er zijn meerdere mogelijkheden om de data te koppelen aan de business rules.

## Kenmerken

Door middel van toepassingsalgoritmen worden de business rules 'afgevuurd' en worden conclusies afgeleid (berekening, classificatie enzovoort). Naast het ontwikkelen van een dergelijk toepassingsalgoritme met de gebruikelijke talen zoals SQL of C#, waarbij de ontwikkelaar expliciet de onafhankelijkheid tussen business rules en het algoritme moet waarborgen, zijn er business rules engines beschikbaar die in staat zijn vaak meerdere, bewezen toepassingsalgoritmes in te zetten. Dit betekent dat de focus meer kan liggen op het analyseren, ontwerpen en relatief eenvoudig beheren van business rules in plaats van het ontwikkelen van een eigen engine met alle nadelen van dien.

Een typische business rules engine heeft de volgende kenmerken (zie <http://searchsystemschannel.techtarget.com>):

- Business Rule Repository: een database waarin business rules worden opgeslagen zoals gedefinieerd door de gebruikers;
- Business Rule Editor: een intuïtieve front-end applicatie met een user interface waarmee business rules gedefinieerd, ontworpen, gedocumenteerd en beheerd kunnen worden;
- Query & Reporting Component: gebruikers en beheerders rapporteren op bestaande regels;
- Rules Engine Execution Core: de software waarmee regels worden gevonden en afgevuurd.

## Voorbeeld

Visual Rules ([www.visual-rules.com](http://www.visual-rules.com)) is een business rule engine met een Rule Modeler, waarin regels worden gedefinieerd en getest, vervolgens worden opgeslagen in een repository en daarna worden afgevuurd door de Execution Server. De inputdata voor de regels (voorwaarden) en output van de regels (consequenties) bevinden zich in het datawarehouse alsook in transactiebases.

Afbeelding 2 is een voorbeeld van een grafische weergave van een procesmodel gestuurd door business rules in Visual Rules: voor elke klant in het datawarehouse wordt op basis van het aantal bonuskaartpunten bepaald wat de Klantscore wordt en de bijbehorende Status (Premium, Inactive, Normal), zie afbeelding 3. De beslistabel is een standaardfunctionaliteit waarmee eenvoudi-

Bonus Card	Bonus Points	Customer Scoring	Customer Status
Gold	> 100	6	Premium
	= 0	0	Inactive
	Else	3	Normal
Silver	> 100	5	Premium
	= 0	0	Inactive
	Else	2	Normal
Else	> 100	4	Premium
	= 0	0	Inactive
	Else	1	Normal

Afbeelding 3: Beslistabel (Bron: [www.visual-rules.com](http://www.visual-rules.com)).

ge maar ook complexere verzamelingen van (samenhangende) regels gedefinieerd worden. De rule engine raadpleegt het datawarehouse, bepaalt via de beslistabel voor alle klanten de score en status en slaat deze op, bijvoorbeeld in de datamart zodat deze beslissing direct beschikbaar komt voor rapportage of als input kan dienen voor een front-end applicatie.

## Toekomstvast datawarehouse

In dit artikel hebben we de positie en toegevoegde waarde van business rules in een BI-architectuur beschreven. De BI-referentiearchitectuur garandeert dat de diverse business rules goed binnen de architectuur worden gepositioneerd. De definitie en het beheer van de business rules wordt daardoor minder complex, want transparanter. De business rules kunnen in het datawarehouse worden opgeslagen, door middel van de toepassing van Data Vault is historische opslag en auditeerbaarheid van business rules verzekerd. Een business rule engine vervangt de traditioneel zelf ontwikkelde programmatuur door middel van een softwarecomponent waarin verschillende vormen van inferentie standaard zijn opgenomen. De gebruiker definieert slechts de business rules (in de vorm van formules, berekeningen, beslistabellen, regels) en de engine verzorgt het executeren van de regels op het juiste moment. Een ontwikkeling die recht doet aan het toekomstvaste datawarehouse: niet alleen de datamodellen en metadata en masterdata worden separaat doch in samenhang beheerd, ook de business rules worden transparant, eenduidig en in relatie tot de andere componenten van de architectuur beheerd. De opslag van de business rules in de Data Vault garandeert binnen CDM de (her)genereerbaarheid van datawarehouse en datamarts met behoud van alle business rules. Door gebruik te maken van een business rule engine en repository wordt het mogelijk om niet alleen de business rules eenvoudig te beheren maar worden de juiste regels ook nog eens automatisch geëxecuteerd op de wijze die past bij het beslissingsproces. Inzet van complexe ETL-tools of ingewikkeld en foutgevoelig programmeerwerk om business rules onder te brengen in het datawarehouse zullen daarom tot het verleden gaan behoren.

**Erik Fransen** is senior business consultant, **Antoine Stelma** is lead BI architect en **Carlos Kruijer** is senior BI consultant, allen werkzaam bij Centennium BI Expertisehuis.