

Resultaten van het Nationaal Data Warehouse Onderzoek 2008

Organisaties maken pas op de plaats

Sjoerd Janssen en Hans Lamboo

Het volwassenheidsniveau van datawarehouse-omgevingen in Nederland is ten opzichte van 2005 niet gestegen. Dat is de meest opvallende conclusie van het Nationaal Data Warehouse Onderzoek 2008 (DWO 2008).

Net als in 2005 onderzochten Database Magazine en Atos Origin de staat van de Nederlandse datawarehouse-omgevingen. 92 respondenten uit meer dan 70 verschillende organisaties namen tussen 5 mei en 3 juli deel aan dit onderzoek. Een overgrote meerderheid van deze respondenten is betrokken bij de ontwikkeling en/of organisatie van de datawarehouse-omgeving. Zij vulden de vragen gericht op de volwassenheid van de datawarehouse-omgeving in. De overige acht respondenten maken gebruik van de informatie waar de datawarehousing-omgeving in voorziet. Zij vulden de vragen in die gericht zijn op acceptatie en de ondersteuning die de datawarehouse-omgeving hen biedt.

Het volwassenheidsmodel

Het Cost and Life cycle Management (CLM) model van Atos Origin is gebruikt om te bepalen wat de volwassenheid van de datawarehouse-omgevingen in Nederland is, zie afbeelding 1. Dit model kijkt naar de volgende aspecten:

- Ambitie, wat wil de organisatie met datawarehousing en BI bereiken;

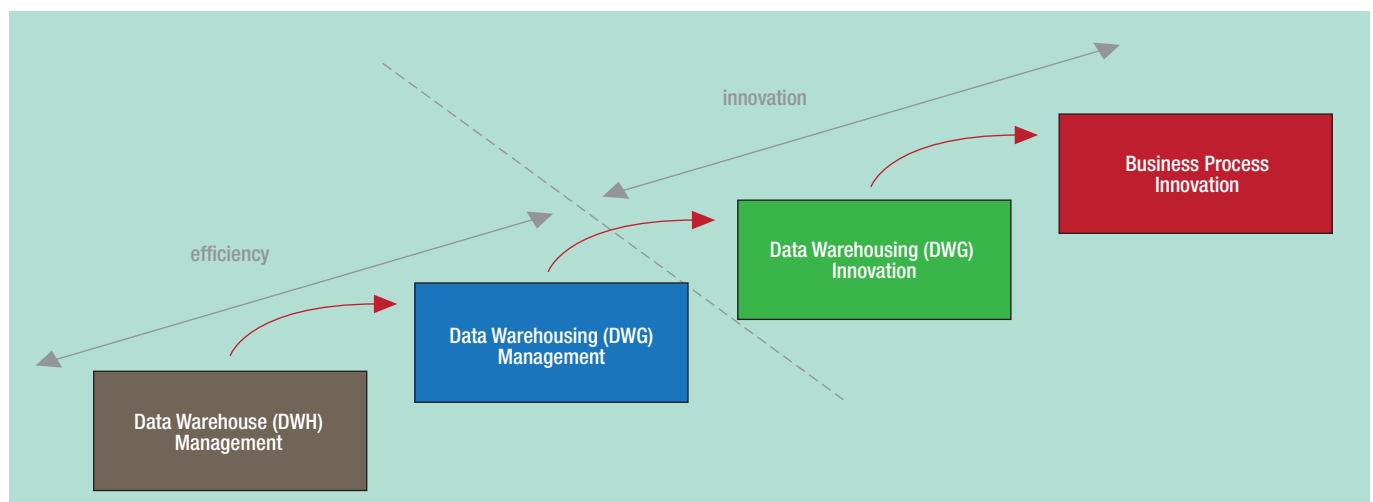
- Reikwijdte, wie maakt er gebruik van het datawarehouse;
- Toepassingen, hoe wordt het datawarehouse gebruikt (rapportage, analyse, datamining etcetera);
- Governance, in welke mate zijn datawarehouse-processen gedefinieerd en werkt de organisatie op deze wijze;
- Technologie, welke tools worden gebruikt;
- Financiering, hoe worden datawarehousing en BI gefinancierd?

Het lijkt dat organisaties de ambities in de loop der jaren naar beneden bijstellen

Het model kent vier niveaus.

Data Warehouse (DWH) Management. De focus van de personen die betrokken zijn bij de ontwikkeling en organisatie van de datawarehouse-omgeving ligt bij dit niveau vooral op de techniek. Datawarehousing wordt hier gezien als een project en een onderwerp dat bij IT thuishoort. Het is dus (nog) geen business onderwerp.

Data Warehousing (DWG) Management. Zowel de business als IT is op dit niveau betrokken bij de datawarehouse-omgeving.



Afbeelding 1: CLM volwassenheidsniveaus.



Meerdere afdelingen en gebruikersgroepen gebruiken de datawarehouse-omgeving. Communicatie en coördinatie tussen IT en de business, maar ook tussen de verschillende business afdelingen staat dan ook centraal.

Data Warehousing (DWG) Innovation. De focus van de datawarehouse-organisatie verschuift op dit niveau naar het gebruik en de acceptatie van het datawarehouse. Er is dan ook veel aandacht voor het op het juiste tijdstip in het juiste formaat en aan de juiste persoon aanleveren van de juiste informatie. De datawarehouse-omgeving ondersteunt het maken van voorspellingen op basis van de beschikbare gegevens.

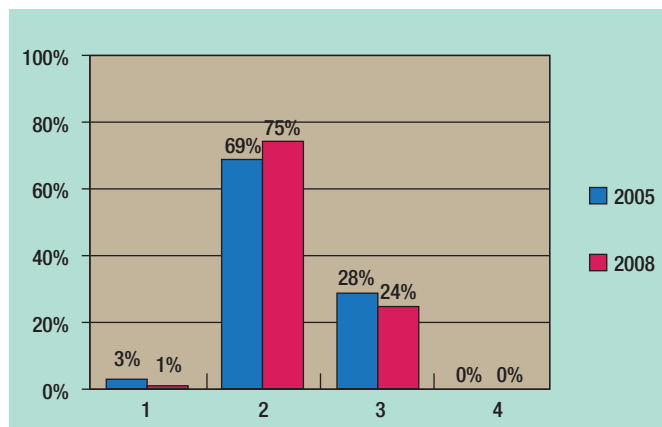
Business Process Innovation. Op dit niveau zijn datawarehousing en BI overal aanwezig en onderdeel van de omgeving waarin wordt gewerkt. De informatie uit het datawarehouse vormt hier een bron voor organisatieverandering. De organisatie wisselt toegesneden informatie uit met leveranciers, partners en klanten.

De belangrijkste resultaten

Uit het onderzoek blijkt dat 75 procent van de datawarehouse-omgevingen zich qua volwassenheid op het niveau van Data Warehousing Management (niveau 2) bevindt. Daarbij valt op dat het totaalbeeld nauwelijks afwijkt van dat in 2005.

Het percentage dat zich op het niveau van Data Warehousing Innovation (niveau 3) bevindt is zelfs iets afgenomen, zie de grafiek in afbeelding 2.

Het doel van datawarehousing is het ondersteunen van de organisatie met relevante informatie op basis waarvan zij beslissingen



Afbeelding 2: Verdeling naar volwassenheidsniveaus.

kan nemen. De behoeften van de organisatie zijn dus leidend in het bepalen van de gewenste omgeving. Het niveau waarop de datawarehouse-omgeving zich bevindt is dus niet goed of slecht, maar moet passen bij de situatie waarin de betreffende organisatie zich bevindt. Waar het om gaat is dat een organisatie en datawarehouse-omgeving optimaal op elkaar zijn afgestemd. Dit kan natuurlijk op elk niveau.

Ambitie naar beneden bijgesteld

Zo is het van belang dat de ambitie ten aanzien van de datawarehouse-omgeving aansluit op de ambities van de organisatie. Ongeveer de helft van de deelnemers geeft aan dat verbetering van de efficiëntie van de informatievoorziening het voornaamste doel is. In 2005 werd uitbreiding van de informatievoorziening nog als voornaamste doel gezien.

Opvallend is dat de helft van de deelnemers met een datawarehouse-omgeving van maximaal vijf jaar oud aangeeft dat het ondersteunen van een nieuwe manier van werken en besluitvorming het voornaamste functionele doel is. Minder dan 20 procent van de deelnemers met een datawarehouse-omgeving van meer dan vijf jaar oud noemt dit nog als belangrijkste doel.

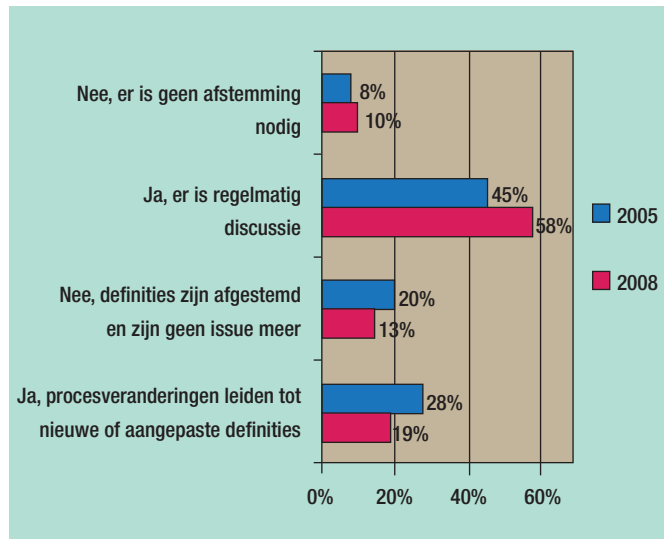
Specifieke datawarehouse databases nemen net als in 2005 maar een klein percentage in

Wanneer we dit vergelijken met 2005, toen minder 'oude' datawarehouse-omgevingen bestonden, komen we tot de conclusie dat het lijkt dat organisaties de ambities van de datawarehouse-omgeving in de loop der jaren naar beneden bijstellen.

Meer gebruikers, meer discussie

Of de datawarehouse-omgeving veel wordt gebruikt is in veel gevallen niet bekend. Van de respondenten geeft 55 procent namelijk aan het gebruik niet te meten. Van de respondenten die aangeven het gebruik wel te meten geeft maar een derde aan de resultaten hiervan structureel te gebruiken bij de verbetering van de omgeving. Dit is een gemiste kans.

De meerderheid van de respondenten geeft aan dat de datawarehouse-omgeving verschillende afdelingen met brede gebruikersgroepen of zelfs de gehele organisatie ondersteunt. Er is een bredere gebruikersondersteuning dan in 2005. Dit leidt echter wel tot veel discussie. Van de respondenten geeft 77 procent aan dat er discussies zijn over de definities van de gerapporteerde data, zie afbeelding 3. Wat daarbij niet helpt is dat in veel gevallen (68 procent) het eigenaarschap van de data niet is toegekend of dat er geen actieve invulling aan dit eigenaarschap wordt gegeven.



Afbeelding 3: Definities van gerapporteerde data.

Nuttige maar incomplete informatie

Ondanks dat er veel discussie is, geeft de meerderheid van de eindgebruikers toch aan dat de datawarehouse-applicatie nuttig is voor hun werk. Daarnaast geeft een meerderheid van deze respondenten aan regelmatig andere systemen voor management-informatie te gebruiken. Aangezien er helaas maar acht eindgebruikers hebben deelgenomen aan het onderzoek kunnen we hier niet al te veel conclusies aan verbinden. Wel is het zo dat de overige 84 respondenten van het onderzoek ook aangeven dat het datawarehouse niet als enige bron voor management-informatie wordt gezien.

Financiering op organisatieniveau, eigenaarschap niet

Bijna twee keer zoveel respondenten als in 2005 geven aan dat het datawarehouse als een algemene dienst wordt gezien en dat de financiering dan ook op organisatieniveau plaatsvindt. Daarbij zijn de kosten vaak medebepalend voor de verdere ontwikkeling van het datawarehouse en sturen organisaties vooral op zichtbare resultaten.

Hoewel de financiering op organisatieniveau goed geregeld lijkt te zijn, is het opmerkelijk dat meer dan de helft van de respondenten aangeeft dat het eigenaarschap van de datawarehouse-omgeving niet (mede) aan de business is toebedeeld. In 2005 gaf de meerderheid van de respondenten (64 procent) nog aan dat het eigenaarschap (mede) aan de business was toegekend.

Tools: Microsoft wint terrein

Respondenten noemen Oracle het vaakst als belangrijkste database, maar Microsoft wint hier terrein. Het aandeel van Oracle daalt van 63 procent naar 43 procent terwijl dat van Microsoft stijgt van 18 procent naar 29 procent. Bij de nieuwere datawarehouses is in bijna evenveel gevallen Microsoft als Oracle het datawarehouse database management systeem. Specifieke datawarehouse databases zoals Sybase IQ en Teradata nemen net als in 2005 maar een klein percentage in. Dit terwijl de datawarehouses wel in omvang toenemen. Een kwart van de datawarehouses is nu meer dan 2 TB groot. Het is niet verwonderlijk dat de omvang een sterke relatie heeft met de leeftijd van de datawarehouse-omgeving. Van de datawarehouses die meer dan tien jaar geleden zijn ontstaan heeft 67 procent een omvang van meer dan 2 TB. Dit percentage ligt bij de datawarehouses die maximaal vijf jaar oud zijn maar op 10 procent.

Om deze datawarehouses van data te voorzien gebruikt bijna een op de vijf respondenten Informatica als belangrijkste ETL-tool. In 2005 gaf bijna een op de drie respondenten dit aan. Informatica verliest dus terrein. Ook op ETL-gebied heeft Microsoft met de introductie van Microsoft Integration Services terrein veroverd. Van de respondenten geeft 12 procent aan deze ETL-tool te gebruiken. Dit gaat overigens niet ten koste van de voorloper van Microsoft. Het gebruik van Microsoft Data Transformation Services is met 7 procent immers nauwelijks teruggelopen.

Bij de rapportage- en analyse-tools worden BO (23 procent) en Cognos (17 procent) net als in 2005 het vaakst genoemd. Overigens werd er toen apart naar OLAP- en rapportage-tools gevraagd, waar dit nu in een vraag is gecombineerd. De meeste

Positie	DBMS	ETL	Front-end
1	Oracle 43% 2005: 63%	Microsoft 19% 2005: 8%	Business Objects 23% 2005: Rapportage 22%, OLAP 15%
2	Microsoft 29% 2005: 18%	Informatica 18% 2005: 31%	Cognos 17% 2005: Rapportage 17%, OLAP 3%
3	IBM 10% 2005: 11%	Oracle 17% 2005: 22%	Microsoft* 15% 2005: Rapportage 4%, OLAP 1%

* Microsoft Excel is hier buiten beschouwing gelaten. Aandeel van Microsoft Excel in 2008: 16%, in 2005 Rapportage 13% en OLAP 3%

Afbeelding 4: Positie van de BI-leveranciers in het DWO.

pakketten bieden nu namelijk zowel OLAP- als rapportage-functionaliteit. BO en Cognos worden op de voet gevolgd door Excel (16 procent), maar dit wordt slechts door 2 respondenten als enige rapportage- en analyse-tool genoemd. In de meeste gevallen wordt Excel dus gebruikt in combinatie met andere rapportage- en analyse-tools. Het gebruiken van meerdere rapportage- en analyse-tools komt overigens veel vaker voor. Van de respondenten geeft 56 procent aan meer dan een tool te gebruiken.

Afstemming van definities noodzakelijk

De meerderheid van de datawarehouse-omgevingen bevindt zich net als in 2005 op het niveau van Data Warehousing Management. Steeds meer gebruikers en afdelingen worden door het datawarehouse ondersteund. Dat brengt echter ook de nodige problemen met zich mee. Zo ontstaat er veel discussie over de definities van de gerapporteerde data. Het is dan ook belangrijk dat er overeenstemming over deze definities wordt bereikt. Het lijkt simpel: gegevenselementen die hetzelfde zijn moeten door verschillende gebruikers en systemen met dezelfde term worden aangeduid. Gegevenselementen die van elkaar afwijken, moeten ook afwijkend worden aangeduid. Let wel, het afstemmen van definities is natuurlijk een business aangelegenheid.

Stem af op gebruik

Dat de datawarehouse-omgeving meer gebruikers ondersteunt, betekent nog niet dat er ook meer en effectief gebruik wordt gemaakt van de datawarehouse-omgeving. Het gebruik wordt immers vaak niet gemeten. Hoe weten we dan of een eenmaal ontwikkeld overzicht nog steeds bruikbaar is voor de organisatie? In de steeds sneller veranderende markten kan het overzicht dat een half jaar geleden nog bruikbaar was nu totaal ongeschikt zijn. Statistieken over het gebruik kunnen hierbij helpen. Zo kan worden nagegaan welke overzichten niet of juist wel relevant zijn. Wanneer organisaties, op basis van deze gegevens, analyses

maken kunnen ze deze gebruiken bij het optimaliseren van de datawarehouse-omgeving. De datawarehouse-omgeving kan dan beter worden afgestemd op het daadwerkelijk gebruik.

Terug naar de Business

Het functionele doel van de datawarehouse-omgeving is naar beneden bijgesteld. Dit hangt wellicht samen met de constatering dat het eigenaarschap van de datawarehouse-omgeving weer meer bij IT is belegd. Het verbeteren van de efficiency is een doel dat goed past bij een IT-afdeling. Willen organisaties meer rendement uit hun omgeving halen dan is het van belang dat er meer business gedreven wordt gewerkt en het eigenaarschap dan ook daar wordt ingevuld.

Bepaal de prioriteit

Binnen een datawarehouse-omgeving zijn op een moment wel 10 á 15 uitdagingen te benoemen. Per organisatie verschillen deze. Een volwassenheidsmodel, zoals in dit artikel beschreven, kan helpen om te bepalen welke aspecten binnen de datawarehouse-omgeving het meest achterblijven. Daarmee wordt duidelijk welke uitdagingen als eerste moeten worden opgepakt.

Tot Slot

De volledige resultaten van het Nationaal Data Warehouse Onderzoek 2008 verschijnen in een boek dat Array Publications, de uitgever van Database Magazine, eind september uitgeeft. Wie verder wil discussiëren over de resultaten van dit onderzoek zien we graag terug tijdens de sessie op het Datawarehousing en Business Intelligence Congres.

Sjoerd Janssen en Hans Lamboo

Sjoerd Janssen (Sjoerd.Janssen@AtosOrigin.com) is consultant BI/DWH bij Atos Origin en geeft leiding aan Data Warehousing Cost & Lifecycle Management (CLM).

Hans Lamboo is hoofdredacteur van Database Magazine.

