

De IT-only invalshoek is voor de meeste organisaties verleden tijd

Tussen 36 en 120 minuten BI per dag

Paul van der Linden en Hans Lamboo

Tussen 15 maart en 15 mei heeft Database Magazine het Nationaal Data Warehouse Onderzoek (DWO) 2005 gehouden. Het is de tweede keer dat DBM zo'n onderzoek heeft uitgevoerd.

De eerste keer, in 2000, lag de focus op het uitvoeren van datawarehouse-projecten. Wat zijn de ambities, de kosten en de valkuilen waar rekening mee moet worden gehouden. Deze keer heeft de aandacht gelegen bij aspecten die betrekking hebben op het beheren en onderhouden van datawarehouse-omgevingen. Maar ook acceptatie en gebruik door eindgebruikers (business) hebben in het DWO 2005 een plaats gekregen. De veranderde aandacht heeft er alles mee te maken dat de meeste organisaties inmiddels al over een datawarehouse-omgeving beschikken en daarmee de eerste perikelen ('hoe doe ik dat nou?') inmiddels zijn ontstegen. Maar is dat inderdaad zo? Beschikken de Nederlandse organisaties inmiddels over een geolied datawarehouse-apparaat en krijgen IT'ers van de business de bijbehorende waardering?

De survey

Gedurende twee maanden hebben op www.dbm.nl twee vragenlijsten gestaan die betrekking hadden op technische vragen (voor IT) en acceptatievragen (business). Behalve een aantal inleidende vragen over grootte van bedrijf, branche etcetera ging het hierbij om 28 technische vragen en 25 acceptatievragen. Deelname aan het DWO is geheel anoniem geweest. Uiteindelijk hebben 65 personen de technische vragen ingevuld en zijn er 22 respondenten aan de acceptatievragen. Kijkend naar andere onderzoeken die zich richten op datawarehousing en/of BI is dit een zeer goede respons.

Onderzoeksmodel: volwassenheidsniveaus

Om de positie te bepalen waarin Nederlandse organisaties zich op datawarehouse- en BI-niveau bevinden is gebruik gemaakt van het Cost & Lifecycle Model (CLM) van Atos Origin. Atos Origin is ook de partij geweest die samen met DBM het onderzoek heeft opgezet en uitgevoerd.

Het doel van datawarehousing is om de organisatie te voorzien van relevante managementinformatie. Alles wat zich bevindt tussen de operationele systemen die de data aanleveren en het daadwerkelijk gebruik van managementinformatie is de dataware-

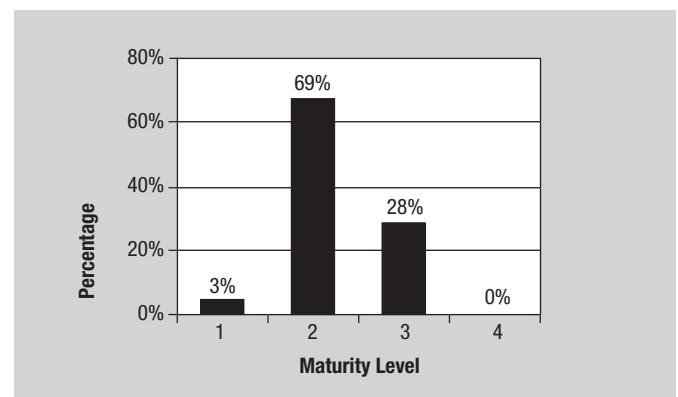
house-omgeving. Om te bepalen hoe goed organisaties erin slagen om de gevraagde managementinformatie te leveren moet dus het samenwerkend geheel wat data warehouse-omgeving heet onder de loep worden gelegd.

Om het datawarehousing landschap in kaart te brengen zijn vier plateaus gedefinieerd.

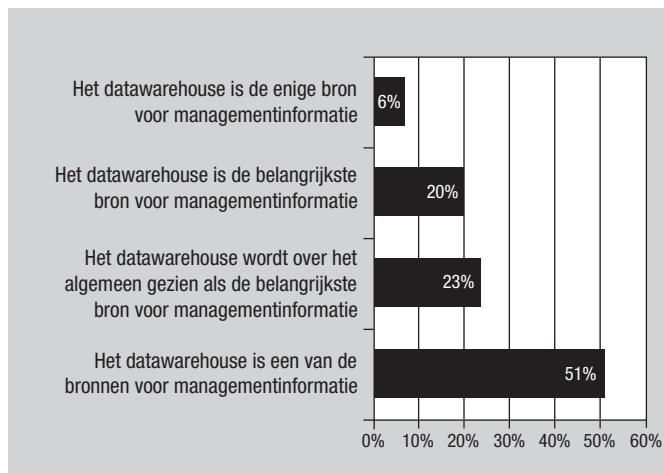
Data Warehouse (DWH) Management: de nadruk ligt op het datawarehouse in technische zin. Het data warehouse wordt veelal gezien als een project. Veel inspanning is nodig om een herhaalbare werkwijze op te bouwen. Aanleiding is veelal het nastreven van een effectievere en efficiëntere rapportage. Financiering gebeurt in de meeste gevallen op projectbasis (incidenteel).

Data Warehousing (DWG) Management: uitbreidingen en wijzigingen aan de datawarehouse-omgeving gaan via een gedefinieerde aanpak. Het datawarehouse wordt nu gezien als een proces waardoor aandacht ontstaat voor beheer, organisatie-aspecten, rollen en procesinrichting. Het aantal organisatie-onderdelen die gebruik maken van het datawarehouse neemt toe waardoor afstemming over bijvoorbeeld de betekenis van data noodzakelijk wordt. Betere analysemogelijkheden ontstaan en opportunity spotting wordt mogelijk. Het datawarehouse wordt gezien als een cost centre.

Data Warehousing Innovation: de nadruk is inmiddels verschoven naar het gebruik en de acceptatie van het datawarehouse. De meeste organisatie-onderdelen maken voor een groot deel van hun werkzaamheden gebruik van datawarehousing. Afspraken



Afbeelding 1: Verdeling naar Maturity Level.



Afbeelding 2: Hoe ziet de organisatie het datawarehouse?

worden zoveel mogelijk gemeenschappelijk (boven processen) afgesproken. Aandacht voor interprocesverbetering ontstaat (opportunity spotting over processen heen). Datawarehousing wordt gezien als een profit centre.

Business Process Innovation: datawarehousing als bron voor organisatieverandering. Informatie uit het data warehouse staat hier centraal en suggereert kansen en signaleert bedreigingen. Afspraken worden op organisatieniveau gemaakt. Aandacht strekt zich uit tot de extended enterprise.

Organisaties die zich op een specifiek plateau bevinden kunnen binnen dat kader evolutionair groeien. Tussen de verschillende plateaus doen zich echter crises voor. Er is extra inspanning nodig om hier doorheen te breken en een volgend plateau te bereiken. Ook is het zo dat een uitdaging niet op elk niveau even belangrijk is. Organisaties die een veelheid van datawarehouse issues hebben moeten zich er dus van bewust zijn welke van deze issues, gegeven hun DWH-plateau, nu relevant zijn en welke issues voorlopig niet geadresseerd hoeven te worden.

Het plateau waarop datawarehousing zich bevindt is niet goed of slecht. Het wordt bepaald door de situatie waarin de betreffende organisatie zich bevindt. Waar het om gaat is dat een organisatie en de datawarehouse-omgeving optimaal op elkaar zijn afgestemd. Dat kan uiteraard op elk niveau.

Onderzoeksmodel: aspecten

Het maturity level wordt onder andere bepaald door te kijken naar de volgende aspecten:

- Reikwijdte, welk deel van de organisatie (onderdelen, gebruikerstypen) maakt gebruik van het datawarehouse;
- Toepassingen, soort gebruik van het datawarehouse (rapportage, OLAP, datamining etcetera);
- Governance, in welke mate en hoe is het beheer geregeld;
- Technologie, welke tools worden gebruikt;
- Financiering, hoe worden datawarehousing en BI gefinancierd;
- Ambitie, wat is het doel van datawarehousing?

Overigens zullen de meeste organisaties aspecten laten zien die thuishoren bij verschillende maturity levels. Dit is normaal. Het gaat er juist om dat duidelijk wordt waar de datawarehousing van een organisatie zich als geheel bevindt. Indien de meeste indicatoren wijzen op een level 2 organisatie (DWG Management) is het zaak om eerst level 1 problemen te adresseren, vervolgens zich te concentreren op de level 2 aspecten en level 3 en 4 voorlopig even met rust te laten.

In dit artikel worden de resultaten van het DWO op hoofdlijnen bespreken. Een meer gedetailleerde analyse maakt deel uit van het pdf-boek dat naar alle deelnemers zal worden gestuurd. Het boek dat tijdens de Expert Meeting op 29 juni zal worden uitgereikt zet daarnaast de resultaten van het DWO 2005 af tegen andere, internationale onderzoeken.

Resultaten Technisch

Karakteristieken deelnemende organisaties.

Het onderzoek is voornamelijk ingevuld door grote organisaties. Groot heeft daarbij betrekking op zowel omzet (66 procent heeft een omzet van meer dan € 20 miljoen) als op aantal medewerkers (meer dan 1000 medewerkers – de hoogste categorie).

Naar branche verdeeld gaat het met name om organisaties in telecom, utilities en media (20 procent), banken en verzekeringen (20 procent) en overheid en not-for-profit (18 procent).

Wat opmerkelijk is, is dat maar liefst 57 procent van de organisaties aangeeft tussen een en vijf jaar onderweg te zijn met datawarehousing. Gegeven het feit dat Inmons boek 'Building the Data Warehouse' toch al weer 13 jaar oud is (!) zijn de Nederlandse datawarehouses van recenter datum dan over het algemeen wordt aangenomen. Hierbij behoort wel een opmerking. Respondenten bleken soms moeite te hebben bij het beantwoorden van deze vraag¹. De reden hiervoor is dat niet altijd duidelijk was wat data warehouse genoemd kon worden en wanneer dus daadwerkelijk gestart is.

Tussen de verschillende plateaus doen zich echter crises voor

Over alle deelnemende organisaties heen kan geconstateerd worden dat 80 procent van de geïmplementeerde datawarehouses meerdere afdelingen ondersteunen. Dat betekent dat de meeste organisaties inmiddels de fase achter zich hebben gelaten waarin slechts een afdeling wordt ondersteund.

Ambitie en beheer.

Qua ambitie hebben datawarehouses in tweederde van de gevallen een 'down to earth'-karakter. Het gaat dan om het verbeteren van de efficiëntie van de informatievoorziening en de uitbreiding

van de informatievoorziening. Flexibilisering van de informatievoorziening en ondersteunen van een nieuwe manier van werken en besluitvorming is in eenderde van de gevallen het doel. Ook op financieel gebied blijkt de doelstelling slechts in 26 procent gericht op het vinden van nieuwe mogelijkheden om de winstgevendheid te vergroten. De resterende 74 procent heeft betrekking op het beter benutten van hetgeen potentieel al aanwezig is. Kijkend naar de beheeraspecten blijkt dat in 31 procent van de gevallen er geen vaste procedures bestaan hoe om te gaan met door gebruikers gewenste veranderingen. Het gebruik dat wordt gemaakt van het datawarehouse wordt in de meerderheid van de gevallen (71 procent) ook niet gemeten. Dat scheidt het beeld van een organisatie die druk bezig is om een datawarehouse op te tuigen dat (inmiddels) meerdere afdelingen van informatie voorziet – waarbij men nog niet is toegekomen aan het daadwerkelijk meten van gebruik. Het optuigen van de datawarehouse-omgeving blijkt ook uit de antwoorden op de vragen omtrent eigenaarschap, taken en verantwoordelijkheden. Het eigenaarschap van de datawarehouse-omgeving is in 82 procent van de gevallen toebedeeld. Voor de eigenaarschap van data gaat het om 79 procent van de organisaties.

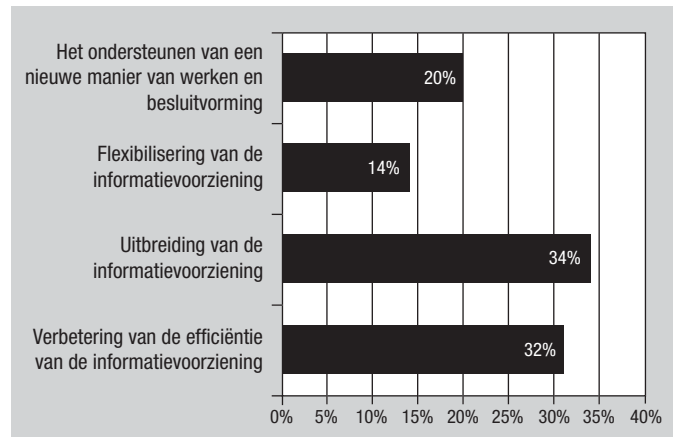
Informatica Powermart/ Powercenter wordt genoemd als meest gebruikte ETL-tool

Discussies over de definities van de gerapporteerde data komen in 73 procent van de gevallen voor. In 28 procent van de gevallen wordt geantwoord dat er geen discussies zijn. Bij 8 procent is geen afstemming nodig; dit zou van toepassing kunnen zijn op de situatie waarin het datawarehouse slechts 1 afdeling ondersteund (ook 8 procent). Echter, 20 procent geeft aan dat definities zijn afgestemd en verder geen issue meer zijn.

Interessant is ook hoe IT inschat hoe de organisatie het datawarehouse ziet. Maar liefst 51 procent antwoordt dat het datawarehouse een van de bronnen voor managementinformatie is. Dat is dus veel minder stellig en ambitieus dan de enige bron voor managementinformatie (wordt maar door 6 procent genoemd). Hoopgevend is dat 20 procent het datawarehouse ziet als de belangrijkste (maar niet enige) bron van managementinformatie.

Kosten en financiering.

De jaarlijkse licentiekosten bedragen voor 20 procent van de bedrijven meer dan € 100.000,- en voor 18 procent bedragen ze echter minder dan € 25.000,-; hier is dus sprake van een tweedeling. Deze tweedeling is ook te zien bij de jaarlijkse ontwikkelkosten: 31 procent meer dan € 250.000,- en 18 procent minder dan € 50.000,- en ook de jaarlijkse beheerkosten laten deze tweedeling



Afbeelding 3: Wat is het doel van de datawarehouse-omgeving op functioneel gebied?

zien: 22 procent meer dan € 250.000,- en 20 procent minder dan € 50.000,- .

Kijkend naar de wijze waarop het datawarehouse wordt gefinancierd en het belang van de kosten voor het datawarehouse, maakt aannemelijk dat de meeste organisaties een datawarehouse zien als een algemene voorziening. Een aantal van 57 procent bekostigt het datawarehouse uit een apart of organisatie-budget. Belangrijk zijn de kosten volgens opgave maar in 26 procent van de gevallen. In de rest van de gevallen is het hoogstens medebepalend, een factor die meespeelt of zelfs een marginale factor.

Ook het antwoord op de vraag naar het belang van de baten bevestigt dit beeld. 'Zichtbare baten' worden meer dan vier keer vaker genoemd als belangrijke factor dan 'Meetbare baten'. Behalve het algemene voorzieningskarakter kan hierbij ook meespelen dat het meten van baten in een datawarehouse-omgeving bijzonder lastig kan zijn.

Omvang en gebruik.

De meeste datawarehouses zijn tussen de 500 MB en 2 TB groot (63 procent), 23 procent is kleiner, 14 procent is groter dan 2 TB. De maandelijkse dataload bedraagt voor 60 procent tussen de 50 MB en de 200 GB.

In maar liefst 54 procent van de gevallen wordt aangegeven dat alle gebruikersgroepen gebruik maken van de datawarehouse-omgeving. Dit onderschrijft de ontwikkeling dat datawarehousing en BI inmiddels niet alleen zijn voorbehouden aan managers en analisten, maar breed in de organisatie worden aangewend. Kijkend naar de toepassingen op het datawarehouse wordt dit verder bevestigd: 23 procent van de organisaties beschikt niet over een OLAP-tool. Anderzijds beschikt 48 procent over twee of meer rapportage-tools. Rapportage blijft de meest toegepaste activiteit op het datawarehouse.

Tools en informatieverspreiding.

Niet verrassend is dat Informatica Powermart/Powercenter wordt genoemd als meest gebruikte ETL-tool (31 procent). Een goede

tweede is Oracle Warehouse Builder wat door 22 procent van de respondenten wordt genoemd. Gezien de koppositie van Oracle bij de databases (63 procent) en het meeleveren van OWB is dit eveneens niet verrassend. Wat wel opvalt is dat Microsoft Data Transformation Services (DTS) maar 8 procent scoort. Dat kan te maken hebben met het feit dat met name grote organisaties de survey hebben ingevuld en dat SQL Server (18 procent) en DTS toch nog steeds niet worden gezien als producten die veel data en gebruikers aankunnen. Wat betreft hoeveelheid data wordt deze conclusie ondersteund, qua aantal gebruikers echter niet. Bij de databases scoren zoals vermeld Oracle en Microsoft het hoogste, gevolgd door IBM DB2 (11 procent). Specifieke datawarehouse databases zoals Sybase IQ en Teradata nemen een verwaarloosbaar percentage in (samen twee procent).

Aan de ontsluitingskant is Business Objects koploper met maar liefst 40 procent, gevolgd door Cognos met 31 procent. In het geval van Cognos hebben we het dan over Impromptu en de opvolger Reporting Services. Ondanks de goede berichten over Reporting Services blijft BO dus nog steeds een stap voor. Derde plaats is weggelegd voor Excel (25 procent) met SAS op een vierde plaats (22 procent). Microsoft Reporting Services blijft in dit veld (nog) achter met acht procent. Voorlopig wordt hier meer over geschreven dan dat het daadwerkelijk wordt gebruikt. Bij de OLAP-tools is de volgorde andersom. Hier leidt Cognos (29 procent), gevolgd door Business Objects (20 procent). De derde plaats is voor Hyperion Essbase, dat net iets beter scoort dan SAS (15 procent versus 14 procent). Opvallend is dat Excel hier maar 5 procent scoort; klaarblijkelijk zijn de mogelijkheden van Pivot Tables nog niet breed ontdekt of toegepast. Wordt al die informatie vervolgens ook elektronisch verspreid binnen de organisatie? In 68 procent van de gevallen blijkt dat zo te zijn. Portals spelen daarbij nog geen grote rol. 43 procent Maakt geen gebruik van een portal. 15 procent Beschikt wel over een portal, maar gebruikt die niet om managementinformatie aan te bieden.

Resultaten Acceptatie

Ook de acceptatievragen zijn voornamelijk door grote bedrijven ingevuld (naar omzet en aantal medewerkers gemeten). Input werd voornamelijk verkregen van overheid en not-for-profit (23 procent), telecom, utilities en media (23 procent) en de zakelijke dienstverlening (18 procent).

Respondenten geven aan dat de BI-applicatie hun in staat stelt om hun werk sneller te doen (86 procent), dat het hun productiviteit en effectiviteit verbetert (respectievelijk 78 procent en 91 procent) en dat de BI-applicatie nuttig is voor hun werk (100 procent is het hiermee eens of geheel mee eens). 87 procent Stelt de BI-applicatie het hun makkelijker maakt om hun werk te doen. Echter, het vergt wel iets om de BI-applicatie te gebruiken. Dat het makkelijk is om de BI-applicatie te leren gebruiken vindt nog 54 procent (mee eens of geheel mee eens). Hiermee is 27 procent het niet eens of geheel mee oneens. Een percentage van 50 vindt het vervolgens makkelijk om de BI-applicatie te gebruiken, maar

slechts 27 procent vindt het makkelijk om de BI-applicatie te laten doen hij wil. Dat de werking van de applicatie duidelijk en begrijpelijk is vindt minder dan de helft (45 procent).

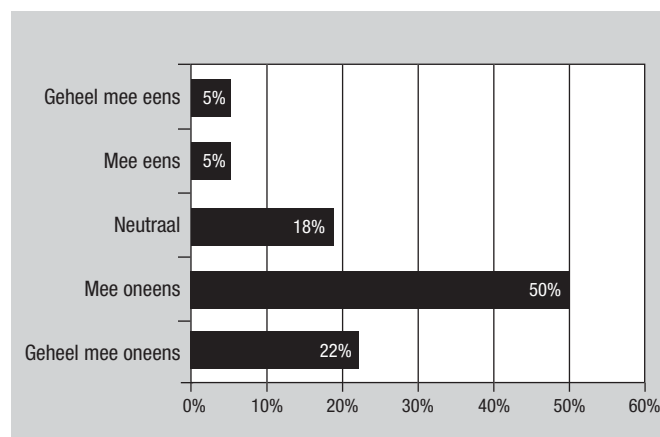
De helft van de respondenten vindt dat er een alternatief is voor de informatie die wordt geleverd door de BI-applicatie. Dat kunnen andere systemen zijn (64 procent) of andere personen (41 procent). De helft van de organisaties geeft aan dat ze geen andere bronnen gebruiken omdat deze betere informatie zou leveren. Hierbij geldt overigens dat maar liefst 41 procent op dit punt neutraal invult en 9 procent aangeeft dit wel te doen.

Het exporteren van gegevens vanuit de BI-applicatie om deze verder te bewerken en zo betere informatie te krijgen wordt door een overtuigende 69 procent gedaan. Inzoomend op de redenen waarom de BI-applicatie wordt gebruikt ontstaat het volgende beeld:

- Is zo gebruikelijk binnen de organisatie (73 procent mee eens/ neutraal);
- Levert nuttige informatie op (91 procent mee eens/geheel mee eens);
- Biedt nieuwe inzichten (68 procent mee eens/geheel mee eens);
- Op een hoger plan brengen (72 procent mee eens/geheel mee eens);
- Legt meer gewicht in de schaal (45 procent neutraal).

Uit de IT-survey blijkt dat het doel van datawarehousing in slechts 26 procent van de gevallen het opdoen van nieuwe opzichten is. De business survey geeft aan dat toch in meer dan tweederde van alle gevallen nieuwe inzichten worden opgedaan.

Vaak wordt een datawarehouse gepositioneerd als single-version-of-the-truth. Rapportages die gebruik maken van het datawarehouse scoren dan hoger dan andere rapportages. Soms worden rapportages waarvan niet kan worden aangetoond dat ze hun informatie van het datawarehouse hebben ontvangen niet eens geaccepteerd. Uit de feedback van de eindgebruikers blijkt dat in de praktijk dit nog niet zo werkt. Slechts 10 procent van de business respondenten ziet de BI-applicatie als enige bron van managementinformatie (bij technische respondenten was dat 6



Afbeelding 4: "Ik zie de BI-applicatie als de enige bron van managementinformatie?"

procent). De BI-applicatie als belangrijkste bron van managementinformatie geldt voor 54 procent van de business en 49 procent van de technische respondenten. 82 procent Gebruikt naast de BI-applicatie nog regelmatig andere systemen voor managementinformatie.

Een groot deel van de business gebruikt de BI-applicatie elke dag (32 procent). Een additionele 18 procent gebruikt de applicatie bijna elke dag. En 55 procent gebruikt de BI-applicatie tussen drie en tien uur per week. Uitgaande van een werkweek van vijf dagen betekent dit dat men minimaal 36 minuten per dag en maximaal twee uur per dag met de BI-applicatie bezig is.

Conclusies

De meerderheid van de deelnemende organisaties (tussen de 69 procent en 72 procent) bevindt zich inmiddels op niveau 2: Data Warehousing Management. De IT-only invalshoek is hiermee voor de meeste organisaties verleden tijd en ingeruild ten faveure van een procesbenadering. Meerdere afdelingen en meerdere toepassingen zijn aangesloten op het datawarehouse. Verbeteren van de efficiëntie van de informatievoorziening en de uitbreiding van de informatievoorziening zijn de doelen die worden nagestreefd. Een belangrijk deel van de governance-aspecten is ingevuld (eigenaarschap van data en datawarehouse, wijzigingsprocedure), maar verdergaande aspecten als meten van het feitelijke gebruik

van het data warehouse en daarop sturen worden in het merendeel van de gevallen nog niet gemeten.

Gebruikers van de BI-applicaties zijn overwegend positief als het gaat om productiviteit, effectiviteit, nut en gemak. Overigens mag de BI-applicatie best wel duidelijker en begrijpelijker zijn en makkelijker om te laten doen wat men wil. Rekening houdend met 12 procent van de organisaties die al op niveau 3 opereren mag geconstateerd worden dat we sinds het eerste Nationaal Data Warehouse Onderzoek goed zijn opgeschoten. Wie verder wil discussiëren over de resultaten van het DWO 2005 zien we graag op de Expert Meeting van 29 juni.

Noot

1. Naast de vragenlijst is een beperkt aantal interviews afgenomen. De hier vermelde waarneming heeft betrekking op de interviews.

Paul van der Linden en Hans Lamboo

Paul van der Linden (Paul.PFH.vanderLinden@AtosOrigin.com) is senior consultant Data Warehousing/BI bij Atos Origin en geeft leiding aan Data Warehousing Cost & Lifecycle Management (CLM).

Hans Lamboo (Hans.Lamboo@Array.nl) is hoofdredacteur van Database Magazine.