

Vermijden van achterstallig onderhoud

APK voor het datawarehouse

Marc Houtkooper

Datawarehouses zijn inmiddels binnen elke grote organisatie een begrip en met de komst van goedkopere hard- en software komen de voordelen van een datawarehouse nu ook binnen het handbereik van middelgrote organisaties. De verwachting is dat met de opkomst van customer relationship management en enterprise resource planning-systemen het aantal datawarehouses de komende jaren alleen maar zal toenemen. Doordat veel datawarehouse-projecten niet het gewenste resultaat hebben opgeleverd, hebben deze binnen veel organisaties een negatieve indruk achtergelaten. Momenteel worden veel standaard-toepassingen met datawarehouses of datamarts geleverd onder de naam analytical applications. De basis blijft echter toch altijd een database waar consistent en geïntegreerd de gegevens tot informatie worden getransformeerd.

Een datawarehouse-omgeving kunnen we het best vergelijken met het productieproces van een fabriek. De grondstoffen zijn de gegevens uit verschillende bronsystemen en worden in de fabriek gemengd en geïntegreerd tot het eindproduct: informatie. Deze informatie is de basis voor iedere organisatie om haar continuïteit te waarborgen en de concurrentiepositie te verstevigen.

De informatiefabriek is dus essentieel voor de bedrijfsvoering en moet daarom dan ook naadloos aansluiten bij de wensen en eisen ten aanzien van de informatiebe-

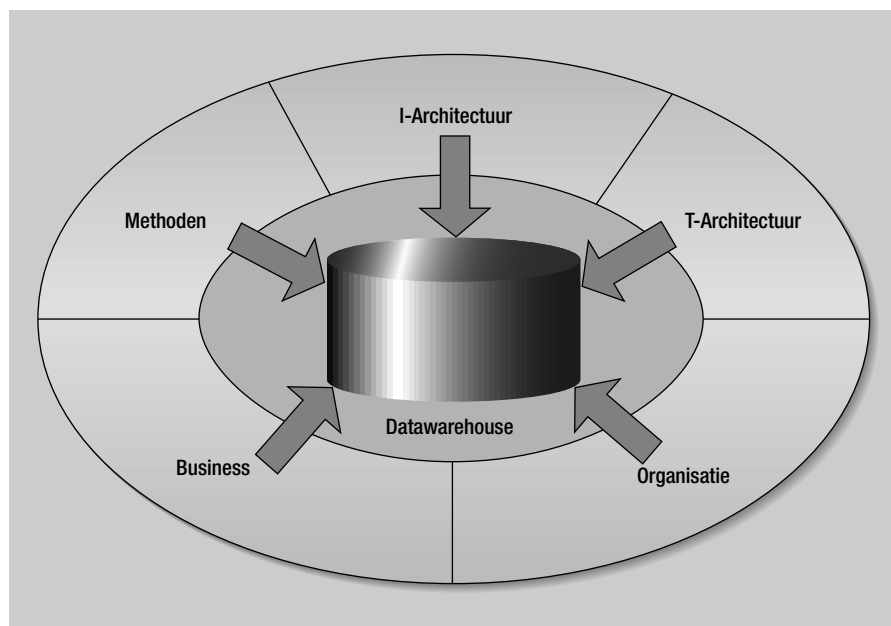
hoefte binnen de organisatie. In de praktijk echter wordt maar al te vaak door achterstallig onderhoud en/of slechte kwaliteit van gegevens het datawarehouse als niet actueel beschouwd.

WANNEER EEN DATAWAREHOUSE ASSESSMENT?

Onterecht worden vaak nieuwe initiatieven gestart om aan de informatiebehoefte tegemoet te komen en worden bestaande datawarehouses uitgefaseerd. Echter door het datawarehouse parallel met de organisatie mee te laten veranderen, kunnen de gemaakte investeringen worden gewaarborgd. Een datawarehouse assessment kan in deze optiek worden gezien als de APK

(Algemene Periodieke Keuring) van de informatiefabriek. Door periodiek de datawarehouse omgeving vanuit verschillende invalshoeken te keuren, wordt deze omgeving optimaal afgestemd op de informatiebehoefte van de organisatie.

Reeds bij de start van het project moeten de informatiebehoeften goed in kaart worden gebracht. Ook onderwerpen als de keuze van de architectuur, de keuze van de software en het oplossen van diverse initiële problemen dragen bij aan een succesvolle ontwikkeling en inrichting van een datawarehouse-omgeving. Uit ervaring blijkt dat veel datawarehouse-projecten problemen ondervinden tijdens de ontwikkeling en implementatiefase. Met name door de onverwachte complexiteit



FIGUUR 1: DE INFORMATIEFABRIEK

van de gegevensintegratie in combinatie met de onervarenheid van de gebruikers en ontwikkelaars blijven de resultaten van een datawarehouse-project achter bij de verwachtingen. Met als gevolg dat de kosten vaak hoger uitvallen dan de oorspronkelijke begroting. Ook hier kan een tussentijdse assessment uitkomst bieden.

Een datawarehouse assessment is met name bedoeld voor de fase na de ontwikkeling van de eerste incrementen. Wanneer bronsystemen, informatiebehoeften, technologieën en andere omgevingsfactoren veranderen, is een tijdige evaluatie of onderzoek naar de status van het datawarehouse een waardevolle investering. Een datawarehouse assessment bewijst met name zijn nut in de fase van het volwassen worden, uitbreiding en completeren van een datawarehouse-omgeving. Enkele voorbeelden wanneer een datawarehouse assessment zinvol is, zijn:

- *Het opnieuw actualiseren van een bestaand datawarehouse.*

Herkenbare feiten zijn dat het gebruik en de toegevoegde waarde van de informatie uit het datawarehouse teleurstellend en beneden verwachting is. Geleidelijk verliest het datawarehouse de aandacht van de organisatie en er worden alternatieve projecten gestart om aan de informatiebehoefte tegemoet te komen.

- *De overgang van ontwikkeling naar beheer.*

Na de fase van ontwikkeling dient er voortdurende gecontroleerd te worden op de kwaliteit van de gegevens, de informatiebehoeften en de volledige datawarehouse-infrastructuur (zowel technisch als functioneel). Vaak wordt deze overgang onderschat en denkt men na de bouw gereed te zijn en worden vervolgens budgetten en begrotingen aangepast. Een datawarehouse-project moet altijd overgaan in een datawarehouse-proces, waarbij de juiste middelen en mensen worden ingezet.

- *Overgang naar een 2e generatie datawarehouse.*

Het gaat hier om de transformatie naar een solide datawarehouse-omgeving. Hierbij zijn de aandachtspunten een geautomatiseerd extractieproces, een volwaardige gegevensintegratie, optimale performance en gebruiksvriendelijke toegangsmogelijkheden. Veel ad-hoc ontwikkelwerk moet worden herkend en omgezet naar structurele processen, die bovendien uitvoerig moeten worden gedocumenteerd. Ook het onderhouden en beheren van de metagegevens is hierbij een essentieel aandachtspunt.

- *Positionering van het datawarehouse als kernproces.*

Deze activiteit vergt met name inspanning aan de gebruikerszijde. Het betreft hier de bewustwording en het vertrouwen in de kwaliteit en beschikbaarheid van de datawarehouse-omgeving als belangrijkste informatieleverancier.

De bovenstaande voorbeelden geven een indicatie van wanneer een datawarehouse assessment zinvol kan zijn. Het is echter aan te bevelen om net als bij de APK van motorrijtuigen eveneens periodiek een datawarehouse-omgeving te keuren middels een assessment. Hierdoor is men gewaarborgd van het feit dat de bedrijfsvoering te allen tijde op de juiste manier van informatie wordt voorzien.

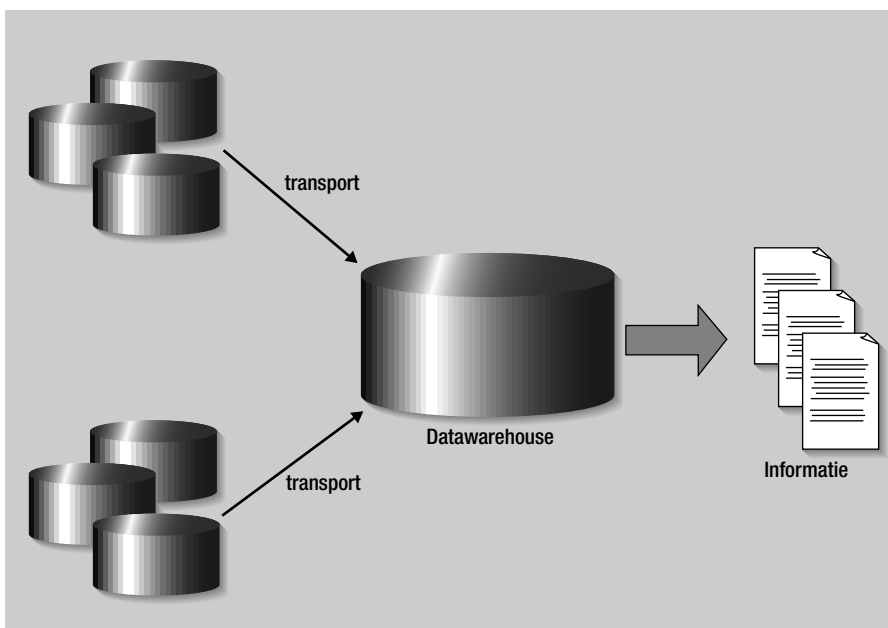
DE INVALSHOEKEN VAN EEN DATAWAREHOUSE ASSESSMENT

Gedurende een datawarehouse assessment analyseert en onderzoekt men vanuit de volgende invalshoeken een datawarehouse omgeving:

1. Business requirements (informatiebehoeften)
2. Informatie-architectuur
3. Technische architectuur
4. Methoden en projectmanagement
5. Organisatie

Tijdens een *business requirements* assessment wordt onderzocht welke informatiebehoeften ten grondslag liggen aan de bedrijfsvoering van de organisatie. Ook wordt onderzocht of de huidige informatievoorziening overeenkomt met deze informatiebehoeften. Aandachtspunten die tijdens deze assessment worden behandeld zijn:

- Actualiteit van de informatiebehoeften en documentatie hierover
- Actualiteit van een business case
- Succesfactoren van de datawarehouse-omgeving
- Betrokkenheid van de sponsor en initiatiefnemer
- Strategische en tactische doelstellingen van de organisatie en het gebruik van informatie hierin.



FIGUUR 2: DE INVALSHOEKEN VAN EEN DATAWAREHOUSE ASSESSMENT.

De *informatie architectuur* assessment sluit aan op de business requirements assessment. Het onderzoek omvat de analyse op de logische gegevensstructuren en of deze volledig zijn ten opzichte van de informatiebehoeften. Ook wordt het ETL (Extractie, Transformatie en Load) proces onder de loep genomen en wordt de organisatie rondom het gebruik van metagegevensbeheer onderzocht.

Aandachtspunten gedurende deze assessment zijn:

- Informatiemodellen en informatiearchitectuur
- Gegevenskwaliteit en gegevensintegratie
- Het ETL (Extractie, Transformatie en Laad) proces
- Toegankelijkheid van gegevens.

Met een *technische architectuur* assessment wordt het gebruik van hardware, software en middleware onderzocht. Met name performance, onderhoud, capaciteitsplanning, backup-procedures en het gebruik van standaardtoepassingen komen aan de orde. Aandachtspunten vanuit deze invalshoek zijn:

- De technische architectuur
- Continuïteit van het ETL-proces
- Performance van het invoer en uitvoerproces
- Gebruikersvriendelijkheid en gebruik van de toepassingen
- Beveiliging van gegevens en informatie.

De assessment op de gebruikte *methodieken* en het *projectmanagement* onderzoekt de datawarehouse ontwikkelcyclus en de bijbehorende methodieken. Een datawarehouse-project is qua aanpak specifiek, omdat voor een snel resultaat meestal wordt gekozen voor een incrementele aanpak. Om die reden vergt het ook een andere kijk op het project management. Ook worden de samenstelling en de vaardigheden van het projectteam onderzocht en wordt de implementatie strategie nader beschouwd, die vervolgens weer samenhangt met de resultaten uit de eerdere assessments.

Aspecten die aan de orde komen tijdens deze assessment zijn:

Om de noodzaak van een datawarehouse assessment aan te tonen hierbij een snelle zelftest, waarin men kan bepalen of een datawarehouse assessment zinvol is binnen de eigen organisatie. Wanneer een vraag niet kan worden beantwoord of wanneer overtuigend nee wordt ingevuld, betekent dit dat er een risico in de datawarehouse omgeving aanwezig is.

Business requirements / informatiebehoeften	Ja	Nee
Is er een business case aanwezig waarin de business requirements zijn gedefinieerd?		
Is beschreven waarom het datawarehouse wordt/is gebouwd?		
Zijn de kritische succesfactoren van het datawarehouse omschreven?		
Zijn de gebruikers en de budgethouders actief betrokken?		
Is gedefinieerd hoe de relatie is met de overige informatieprojecten binnen de organisatie?		
Ondersteund het datawarehouse de strategische doelstellingen van de organisatie?		
Is er een gegevensanalyse op de bronsystemen uitgevoerd?		
Informatie architectuur		
Is de informatiearchitectuur beschreven?		
Zijn er gegevenskwaliteit eisen opgesteld?		
Is het ETL (Extractie, Transformatie en Laad) proces beschreven?		
Zijn de datamodellen dimensioneel en flexibel ontworpen?		
Is er een Staging Area aanwezig?		
Is de informatie correct en bruikbaar voor de eindgebruikers?		
Is de informatie architectuur schaalbaar en eenvoudig te onderhouden?		
Technische architectuur		
Is de technische architectuur beschreven en actueel?		
Werkt het ETL proces automatisch en zonder fouten?		
Is de performance van het laadproces aanvaardbaar?		
Wordt het laadproces binnen de gestelde tijdframe uitgevoerd?		
Is de performance van het uitvoerproces (queries / reporting) aanvaardbaar?		
Is er een capaciteitsplan aanwezig?		
Is er een back-up procedure aanwezig?		
Is er een recovery procedure aanwezig?		
Zijn er voldoende vaardigheden en kennis aanwezig om het datawarehouse te onderhouden?		
Is men tevreden over de inzet van de software?		
Worden alle software toepassingen naar tevredenheid gebruikt?		
Is de beveiliging van de gegevens formeel geregeld?		
Zijn er gescheiden ontwikkel, test en productie omgevingen?		
Is er een change management procedure?		
Zijn er testgegevens aanwezig die gebruikt kunnen worden bij volgende project iteraties?		
Methodieken en Project Management		
Is het project iteratief gedefinieerd?		
Zijn de juiste vaardigheden en kennis aanwezig en beschikbaar?		
Is de planning binnen de gestelde periode haalbaar?		
Zijn de gebruikers en budgethouders ook actief binnen het project?		
Is de scope van het project in het plan van aanpak beschreven?		
Organisatie		
Is de eindverantwoordelijkheid van het datawarehouse omschreven?		
Zijn de functies en rollen voor het beheer bekend en toegewezen?		
Zijn de budgetten voor beheer en ontwikkeling realistisch?		

- De structuur en aanpak van de incrementele projectbenadering
- De inzichtelijkheid van het project
- De vaardigheden en kennis van de projectleden
- De planning en scope van het project
- De samenwerking en communicatie tussen de opdrachtgever, gebruikers en ICT medewerkers.

Met een *organisatie assessment* wordt onderzocht of de organisatie gereed is voor een datawarehouse-omgeving. De projectorganisatie moet op één lijn zitten met de gebruikers en de ICT-afdelingen.

Verantwoordelijkheden, functies en rollen rondom het beheer van het datawarehouse worden in kaart gebracht en geanalyseerd. Vaak wordt de overgang van ontwikkeling naar het beheer ervan onderschat. Na het in productie nemen van een datawarehouse dragen de specialisten de kennis over op een beheerteam en, in het ergste geval, verlaten ze daarna de organisatie. Ook de overdracht van kennis van zogenaamde *key-gebruikers* naar vervangers is een niet te verwaarlozen activiteit. Het succes van een datawarehouse-omgeving wordt hoofdzakelijk bepaald door het draagvlak dat het heeft bij de gebruikers en het management. De aandachtspunten van een organisatie assessment zijn:

- Verantwoordelijkheden bij het beheer en onderhoud (change management)
- De functies en rollen die betrokken zijn bij de datawarehouse-omgeving
- De positie van het datawarehouse in de organisatie.

De vijf hiervoor genoemde invalshoeken geven samen een totaalbeeld van de status van de datawarehouse-omgeving. Een aantal assessments kan als resultaat hebben dat onmiddellijk ingrijpen noodzakelijk wordt geacht, terwijl andere assessments meer de nadruk leggen op de eventuele zwakke punten die via een plan van aanpak kunnen worden verbeterd. In alle gevallen zullen er concrete activiteiten ter verbetering worden aanbevolen.

Door de genoemde assessments en bijbehorende aanbevelingen samen te voegen en hier prioriteiten aan toe te kennen, ontstaat er een geïntegreerd actieplan. Dit

plan moet vervolgens onderdeel gaan uitmaken van de strategische planning binnen de organisatie, zodat binnen een overzichtelijke periode de informatiefabriek weer de juiste informatie gaat leveren om zo de bedrijfsvoering optimaal te ondersteunen.

HET UITVOEREN VAN EEN DATAWAREHOUSE ASSESSMENT

In het algemeen vertegenwoordigen datawarehouse-omgevingen grote en complexe oplossingen voor gegevens-extractie en integratie-problemen. Ook al kunnen deze omgevingen in ontwerp en omvang van elkaar verschillen, toch is de problematiek in regel terug te voeren tot de volgende drie hoofdlijnen:

- De gegevens uit de bronsystemen zijn meestal slecht gedocumenteerd, waardoor juistheid en consistentie moeilijk kunnen worden gegarandeerd.
- De kwaliteit van de gegevens uit bronsystemen laat veelal te wensen over, waardoor de ontwikkeling van extractie- en transformatieprocessen nauwelijks beheersbaar en te plannen zijn.
- Doordat veel datawarehouses door ICT-afdelingen zijn geïnitieerd, zijn de business requirements vaak slecht gespecificeerd en komt het datawarehouse niet tegemoet aan de gewenste informatie-behoefte.

Een datawarehouse assessment zorgt ervoor dat oplossingen en verbetervoorstellen van een datawarehouse-omgeving binnen een doorlooptijd van zes tot acht weken worden gerealiseerd. Het is natuurlijk niet altijd noodzakelijk om alle vijf deelasessments uit te laten voeren. Vaak kunnen één of twee deelasessment(s) voldoende zijn om de belangrijkste knelpunten inzichtelijk te maken.

Het advies is om een datawarehouse assessment uit te laten voeren door een objectieve partij. Door deskundige buitenstaanders het datawarehouse te laten onderzoeken, worden deze niet geconfronteerd met de ontwikkelhistorie en kunnen

er sneller nieuwe inzichten ontstaan. Binnen omvangrijke organisaties is het veelal mogelijk om uit andere datawarehouse-omgevingen deskundigen in te zetten. Een datawarehouse assessment wordt altijd afgesloten met een eindverslag en een presentatie, waarin de huidige situatie en de eventuele knelpunten worden benoemd. Indien mogelijk wordt ook een kosten/baten-analyse toegevoegd. Hiermee kunnen eventuele noodzakelijke investeringen worden verantwoord.

DYNAMIEK

Datawarehouse-omgevingen behoren net zo dynamisch te zijn als de organisaties waar zij voor werken. Maar al te vaak worden organisatorische veranderingen niet doorgevoerd binnen de datawarehouse-omgeving. Dit is het begin van ontevredenheid bij de gebruikers en de roep om nieuwe informatiesystemen. Door het datawarehouse in lijn te houden met de organisatie kunnen aanzienlijke desinvesteringen worden voorkomen. Een datawarehouse-omgeving vertegenwoordigt nooit een tijdelijke of eenmalige oplossing. Door het dynamische karakter van de organisatie en de technologie is het datawarehouse een continue proces van evalueren, evolueren en aanpassen. Daarom levert het periodiek uitvoeren van een datawarehouse assessment een belangrijke bijdrage aan een waardevolle en kwalitatief goede informatievoorziening, waarbij de gedane investeringen blijven gewaarborgd. ●

Referenties:

- R. Kimball, L. Reeves, M. Ross, W. Thornthwaite: *The Data Warehouse Lifecycle Toolkit*. J. Wiley & Sons, 1998.
- Arthur Moore and David L. Wells: *How to do a Data Warehouse Assessment*, 1999.

Marc Houtkooper (marc.houtkooper@newcom.net) is consultant bij Newcom Information Systems.