

Op zoek naar verbetering van de business performance

# Enterprise Information Management (1)

Mike Ferguson en Jan Henderyckx

**In dit eerste deel van een serie artikelen introduceren we Enterprise Information Management (EIM) en onderzoeken we wat er gedaan moet worden om EIM te implementeren teneinde controle over de data in uw onderneming te krijgen.**

Er bestaat geen twijfel over dat data een waardevolle business asset vormen en van vitaal belang zijn voor de business operations en het voortbestaan van een bedrijf. Data zijn voor een onderneming wat water is voor de mens. Een mens kan zowat drie dagen zonder water; daarna bezwijkt het lichaam en houdt het belangrijkste, de geest, ermee op. Hoe lang kan uw proces of onderneming overleven zonder data? Data, en in het verlengde ervan alle informatie, zijn een belangrijke resource die het verschil kunnen uitmaken tussen gewoon zaken blijven doen of uitblinken. Daarom is het zo belangrijk dat we data beheren, schoonhouden en op een consistente, algemeen begrepen manier beschikbaar stellen aan alle bedrijfsonderdelen die het nodig hebben. Om dit te bewerkstelligen hebben we Enterprise Information Management nodig.

## Wat is EIM?

Enterprise Information Management is de verzameling processen die ervoor zorgen dat gestructureerde en ongestructureerde data in de onderneming zo worden beheerd, dat ze in de business kunnen worden gebruikt. Het omvat het management van:

- databenoeming en datadefinities;
- enterprise metadata management;
- datamodellering;
- enterprise datakwaliteit (profiling en cleansing);
- enterprise data-integratie (inclusief dataconsolidatie, datafederatie en datasynchronisatie);
- masterdata management;
- datamonitoring;
- enterprise content management.

Jan Henderyckx is spreker op ons congres Trends in Informatie-management op 18 maart 2008.

Er zijn drie belangrijke redenen waarom we EIM nodig hebben. Ongecontroleerde en niet behoorlijk beheerde data hebben een aanzienlijke impact eerst en vooral op de business operations, vervolgens op besluitvorming en performance management, en tot slot op de compliance.

Als in een onderneming de data niet worden beheerd, heeft dit een rechtstreekse invloed op de business operations.

Medewerkers, klanten, partners en toeleveranciers worstelen allen met het vinden van informatie. Bovendien kunnen incomplete en inaccurate data fouten in de processen veroorzaken. Doordat de gewenste data niet tijdig voorhanden zijn of onbetrouwbaar blijken, reageert de business traag. Ontevreden klanten zijn vaak het gevolg.

## Het niet managen van data heeft een impact op de compliance

Besluitvorming en performance management worden beïnvloed doordat het niet accuraat beheren van data kan leiden tot slechte of foutieve beslissingen. Het kan zelfs mensen ervan weerhouden hoe dan ook een besluit te nemen, omdat ze niet beschikken over de vereiste betrouwbare informatie. Ook kunnen zich reconciliatieproblemen voordoen in het performance management als de metrische berekeningen niet worden begrepen en inconsistent zijn. Tot slot heeft het niet managen van data een impact op de compliance, doordat inaccurate rapportering kan leiden tot schending van wet- en regelgeving.

## Uitdagingen en risico's

Het oplossen van het EIM-probleem is voor de meeste bedrijven een grote uitdaging en is niet meteen iets dat 'in een hand-omdraai' kan worden geregeld. Hoe groter de onderneming en hoe meer gedistribueerde systemen, hoe groter de uitdaging. Daarbij geldt vooral dat de uitdaging zich op velerlei vlak situeert:

- Verschillende versies van de waarheid. Het feit dat data meer dan eens worden opgeslagen, vormt op zich geen probleem

zolang dit een business doel dient, zoals een operational datastore. In alle andere gevallen worden de data opgeslagen in verschillende systemen, wat kan leiden tot grote problemen bij het beheer van klanten en producten, bij interacties met partners en tot problemen inzake compliance;

- Gebrek aan betrouwbare informatie. Data die incompleet, verouderd of inaccuraat zijn of verkeerd geïnterpreteerd, zijn moeilijk te begrijpen. Bovendien is het moeilijk te controleren hoe de data naderhand worden gebruikt. Slecht beheerde datadefinities voor dezelfde data in verschillende systemen leiden tot een gebrek aan begrip en bijgevolg tot een gebrek aan vertrouwen;
- Niet beheerde business data. Sommige gebruikers hebben data gecreëerd in eigen tabellen, spreadsheets, Access databases etcetera, die wel worden gebruikt in de dagelijkse business operaties, maar die niet formeel worden beheerd. Bovendien werken deze zogenaamde 'cottage industry datastores' meestal niet als ze gebaseerd worden op data-subsets afkomstig uit datastructuren in centrale operationele systemen, die hetzij wijzigen, hetzij ophouden te bestaan. Dit is één van de fundamentele uitdagingen waar elk informatie re-engineering project mee geconfronteerd wordt. Het is immers onmogelijk om een impact-analyse te doen op processen die niet gedocumenteerd zijn. Daarbij komt nog dat er meestal geen record management heeft plaatsgevonden van quotes, facturen, verklaringen, etcetera;
- Slecht beheerde synchronisatie van data in verschillende systemen. Subsets van data, gelezen en gekopieerd uit centrale operationele systemen, worden niet beheerd, wat resulteert in data die niet synchroon zijn met mutaties in de operationele data;
- Conflicten bij het vastleggen van data. Incorrect vastgelegde opslagperiodes in applicaties leiden tot hogere kosten, inconsistentie en moeilijkheden bij het garanderen van compliance;
- Teveel informatie en niet weten wat belangrijk is; onder meer het niet gebruiken van waardevolle maar ongestructureerde informatie, het niet gebruiken van vraag-signalen om de supply chain te voeden en het niet gebruiken van klant-analyses bij het onderling afstemmen van marketing en sales.

Uitdagingen gaan hand in hand met risico's. Misschien wel het grootste risico van allemaal is weigeren te erkennen dat het *een probleem is van de onderneming als een geheel*. Het oplossen van de problemen op business niveau kan tot enige verbetering leiden, maar lost het EIM-probleem niet op. Om dit op de juiste manier te doen, is het noodzakelijk te erkennen dat het een ondernemingsbreed probleem is, dat invloed heeft op de mogelijkheid tot handelen, besluiten te nemen en compliant te blijven, en dat bijgevolg ook op ondernemingsniveau moet worden aangepakt. Eén van de meest effectieve manieren om zo'n sponsorship te verkrijgen is te focussen op iets waar de bestuurders oor naar hebben. Bestuurders kunnen niet

om het feit heen dat het niet managen van data de compliance beïnvloedt en het negatieve effect ervan op de algehele business performance. Een bedrijf dat niet voldoet aan de vigerende regel- en wetgeving loopt niet alleen het risico van straffen en boetes, maar riskeert het vertrouwen van de klanten te beschamen. Dit kan leiden tot teleurstellende verkoopcijfers en dalende aandeelkoersen. Aangezien de meeste bestuurders een bonusafpraak hebben gebaseerd op de waarde van de aandelen, is het in hun eigen belang om compliant te zijn en alles te doen wat ze kunnen om de business performance te verbeteren.

## Hoe weet u of uw bedrijf EIM nodig heeft?

Of u EIM nodig heeft blijkt meestal uit klassieke symptomen als:

- het ontbreken van metadata management. Er zijn geen datamodellen of er worden achterhaalde modellen gebruikt; het bedrijf heeft geen data-architect; geen algemeen begrip van de betekenis van de data, er is geen algemeen geldende business begrippenlijst; in verschillende applicaties en BI-systemen worden verschillende definities gebruikt voor dezelfde data;
- slechte datakwaliteit, er wordt geen software gebruikt voor data profiling en cleansing;
- een enorme hoeveelheid redundante en versnipperde inconsistente data in verschillende systemen;
- het gebruiken van veel stand-alone BI-systemen met verschillende ETL-tools en overlapping in het ETL-proces. Dit veroorzaakt congruentieproblemen, 'Excel-mania' en ongecontroleerd knippen & plakken;
- het slecht beheren van masterdata. Versnipperde masterdata met diverse variaties en subsets van dezelfde data in verschillende systemen; de masterdata attributen worden door verschillende systemen onderhouden zonder enige controle over conflicterende data; tussen systemen onderling vindt een tweewegs dataflow plaats om data te synchroniseren (bijvoorbeeld via FTP van files); men lijdt aan 'FTP-mania', om data door te geven worden Excel-attachments onbeheerd

## Enterprise Information Management (I)

In een serie artikelen introduceren Mike Ferguson en Jan Henderyckx Enterprise Information Management (EIM) en onderzoeken zij wat er gedaan moet worden om EIM te implementeren, teneinde controle over de data in de onderneming te krijgen. Het is hun bedoeling om in de komende maanden diepgaand te onderzoeken welke technologieën, methodieken en best practices er omtrent EIM voorhanden zijn. Op deze manier willen ze tot een mechanisme komen dat zowel gestructureerde als ongestructureerde data aankan en zo een bijdrage levert aan de verbetering van de algehele business performance.

- doorgezonden en data worden opnieuw ingebracht;
- het outsourcen van business processen zonder proces-integratie, wat de data alleen maar verder fragmenteert;
- versnipperde, onbeheerde en ongestructureerde content zonder Enterprise Content Management Systeem. Er is geen taxonomie, geen management van records;
- gebrek aan wendbaarheid doordat geen gebruik kan worden gemaakt van de innovatiemogelijkheden. Dit gaat gepaard met snel stijgende kosten door inflexibele systemen en de noodzaak tot veranderen.

Als u na grondige analyse van bovenstaande lijst tot de conclusie bent gekomen dat uw organisatie dringend EIM moet implementeren, rijst de volgende vraag: wat zijn de belangrijkste vereisten voor EIM? Hoe moet een stevig fundament worden gelegd? Het gaat in essentie om drie belangrijke zaken:

- een organisatorische structuur toegewijd aan EIM als deel van een Integration Competency Centre (ICC);
- een bedrijfsbrede standaardsuite van end-to-end technologie voor EIM in één platform;
- een strategie voor enterprise data governance, inclusief een continue verbeteringsmethodiek.

Elk van deze punten behoeft uiteraard wat meer toelichting.

## Organisatorische voorbereiding

Vanuit organisatorisch perspectief is het belangrijk te erkennen dat er een EIM-team moet worden opgericht dat onderdeel is van een IT Integration Competency Centre (ICC). Dit toegewijde team moet op de hoogte zijn van het volledige spectrum van ontwerp-alternatieven voor datamanagement en data-management-patternen definiëren, waarvan de scope breed genoeg is om de meeste datamanagement-problemen aan te kunnen.

Meer bepaald moet het EIM-team verantwoordelijk zijn voor:

- ondernemingsbrede databenoeming en definitiestandaards (bekend als 'shared business vocabulary');
- het enterprise datamodel;
- de business rules voor databeveiliging en datakwaliteit;

- het ontwikkelen van data-integratie;
- het masterdata management;
- het enterprise content management;
- het taxonomie-ontwerp;
- de integratie met andere technologieën.

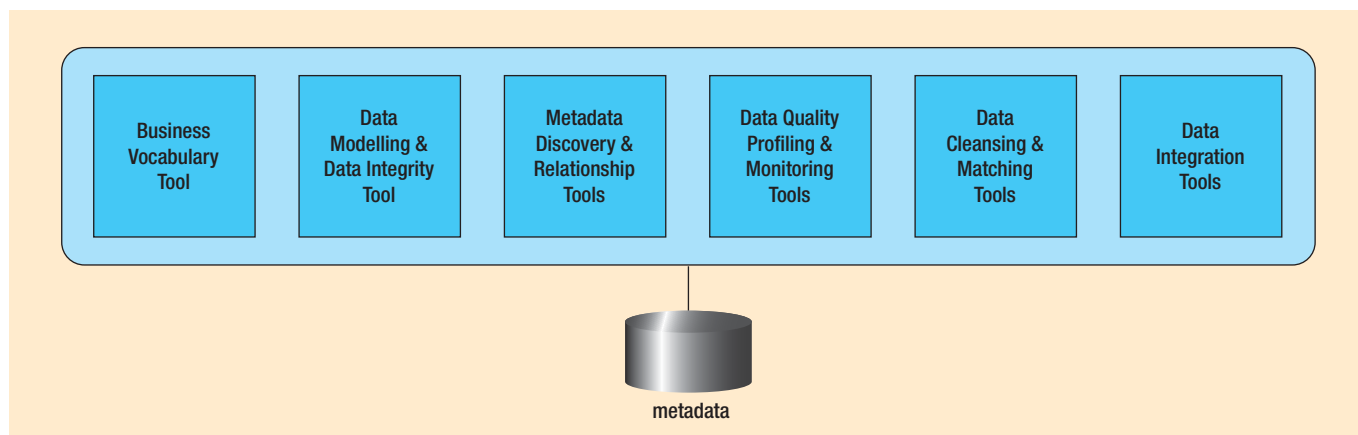
Om zo doelmatig mogelijk te zijn, is het aan te bevelen dat ondernemingen eerst overwegen om de IT-professionals verantwoordelijk voor het beheer van operationele data, BI-systeem data-integratie en enterprise content management samen te voegen. Deze data-architecten zullen moeten samenwerken om het EIM-probleem op te lossen.

In aanvulling op het EIM-team dient er een soort Raad voor Data Governance Change Control te bestaan, die verantwoordelijk is voor de bedrijfsbrede 'shared business vocabulary' (SBV) en voor de goedkeuring van wijzigingsverzoeken voor in onbruik geraakte data items, het wijzigen van bestaande data items en het toevoegen van nieuwe data items die gebruikt worden in de business. Wijzigingsverzoeken moeten eerst deze Raad als een soort 'poortwachter' ter goedkeuring passeren, vooraleer de wijzigingen kunnen worden doorgegeven aan het EIM-team bij IT. Het is aanbevolen dat er een data-architect in deze Raad zit.

Het opruimen van de informatiewarboel vereist tevens een volwassen organisatie aan de business kant, waar fundamentele principes, zoals het data-eigendom, kunnen worden vastgesteld. Niet goed gedefinieerd ownership uit zich vaak in informatie-anarchie. In elke tak van de onderneming moet daarom een data-beheerder worden aangesteld die verantwoordelijk is voor de goedkeuring van het schrappen, wijzigen en toevoegen aan de 'shared business vocabulary', om te kunnen voldoen aan nieuwe vragen vanuit de business en om de integriteitsregels vast te stellen. Deze wijzigingen gaan dan naar de bovenvermelde Raad ter goedkeuring.

## EIM-technologieplatform

De tweede belangrijke eis waaraan moet worden voldaan is de installatie van een EIM-technologieplatform. Dat is niet zo maar



Afbeelding 1: Enterprise Information Management Suite.

---

een tool; het is een suite van geïntegreerde tools voor het beheren van elk type datamanagement-probleem. Het is niet alleen een tool voor ETL of datakwaliteit of enterprise information integration, hoewel al deze gebieden wel moeten worden bediend. Er wordt wel eens gezegd dat voor wie alleen een hamer bezit, alle problemen spijkers zijn. Welnu, het is van kapitaal belang in te zien dat bij EIM zeker niet alle problemen spijkers zijn. Vandaar dat verschillende, geïntegreerde tools moeten worden ingezet die dezelfde metadata gebruiken. Afbeelding 1 illustreert dit proces. Het verdient ook voorkeur dit technologieplatform aan te schaffen en liever niet zelf te bouwen.

Het platform moet de volgende zaken zeker bevatten:

- data en metadata connectors om toegang te krijgen tot verschillende data- en metadata-bronnen en doelen;
- een vocabulary-tool voor eindgebruikers om formeel nieuwe business data te definiëren en door te kunnen bladeren op zoek naar business datadefinities die al bestaan binnen de onderneming;
- een datamodelleer-tool;
- discovery van metadata en metadata-relaties, en identificatie-tools;
- enterprise datakwaliteits-tools voor data profiling en rules-gedreven cleansing;
- Data-integratie software voor Enterprise Information Integration;
- Data-integratie software voor event-driven real-time en batch ETL;
- De mogelijkheid tot data-synchronisatie om wijzigingen door te voeren naar andere systemen, bijvoorbeeld met een enterprise service bus (ESB);
- Software voor enterprise content & record management voor management van de life cycle van documenten, webcontent en digitale media;
- Standaard API's (bijvoorbeeld webservices, JSR 170 & JSR 283, JMS);
- Data-lineage voor data-audits en traceability;
- Databeveiliging en administratie;
- Een metadata repository om de verschillende tools in de suite dezelfde metadata te laten gebruiken.

Via deze tools moet het mogelijk zijn elk type datamanagement-probleem het hoofd te bieden, zoals: datareplicatie; datawarehousing; datafederatie; dataconsolidatie; datamigratie; masterdata management; datasynchronisatie.

Voor wie een dergelijke tool-suite wil aanschaffen: op dit moment zijn (in alfabetische volgorde) DataFlux, IBM, Informatica en SAP (door de overname van BO) toonaangevend in de markt.

Om on-demand toegang tot informatie te kunnen geven, moet het EIM-platform:

- voorzien in een abstractielaag die de locatie van de informatie verborgt. De informatie-afnemer logt alleen maar aan op

het EIM-platform ongeacht de actuele locatie van de bron-informatie;

- voorzien in diverse interfaces: een SQL en XQuery interface om informatie te onttrekken, de mogelijkheid om XML-documenten te produceren, plug-in portlet technologie, message queuing voor event-driven data cleansing en integratie;
- kunnen omgaan met: performance van de bevragingen, inclusief datacaching om de performance van on-demand verzoeken te verbeteren (hierbij kan de cache gedenormaliseerd zijn, geaggregeerd worden, pre-joins bevatten en dergelijke), de complexiteit van sommige bevragingen die noopt tot de constructie van tijdelijke objecten, tijdelijke vereisten in termen van object- en relatie-consistentie, de levenscyclus van de entiteit, Information Transformatie.

## Enterprise Data Governance Strategy

Het invoeren van EIM is veel meer dan het definiëren van een verzameling standaarden en het installeren van wat software. De eerste twee stappen maken het alleen maar mogelijk om de fundamentele verandering in te zetten naar een continue verbetermethodiek waardoor de kwaliteit en het beheer van informatie bedrijfsbreed enorm wordt verbeterd.

## Misschien is het grootste risico weigeren te erkennen dat het een probleem is van de onderneming als een geheel

De aanpak bestaat uit vijf stappen, te beginnen bij de definitie van de 'shared business vocabulary' en modellen, waarna de verbinding wordt gemaakt met de bestaande artefacten gedurende de identificatiefase. Als deze initiële fasen zijn uitgevoerd, kan men, na de daaropvolgende inschattings- en integratiefasen, beginnen te denken aan het ter beschikking stellen van uitvoerbare componenten. De laatste fase is het daadwerkelijke service delivery management.

In deel 2 van deze serie wordt dieper ingegaan op deze Continuous Improvement Methodology en de wijze waarop het in uw bedrijf kan worden ingevoerd.

### Mike Ferguson en Jan Henderyckx

Mike Ferguson is Managing Director van Intelligent Business Strategies Limited. Als analist en consultant is hij gespecialiseerd in Business Intelligence en enterprise business integration.

Jan Henderyckx is onafhankelijk consultant, spreker en auteur op het gebied van informatie-architectuur en databases.

Het copyright van de in deze serie beschreven methode voor Enterprise Information Management ligt bij de auteurs.