

Gegevensintegratie essentiële voorwaarde voor e-business

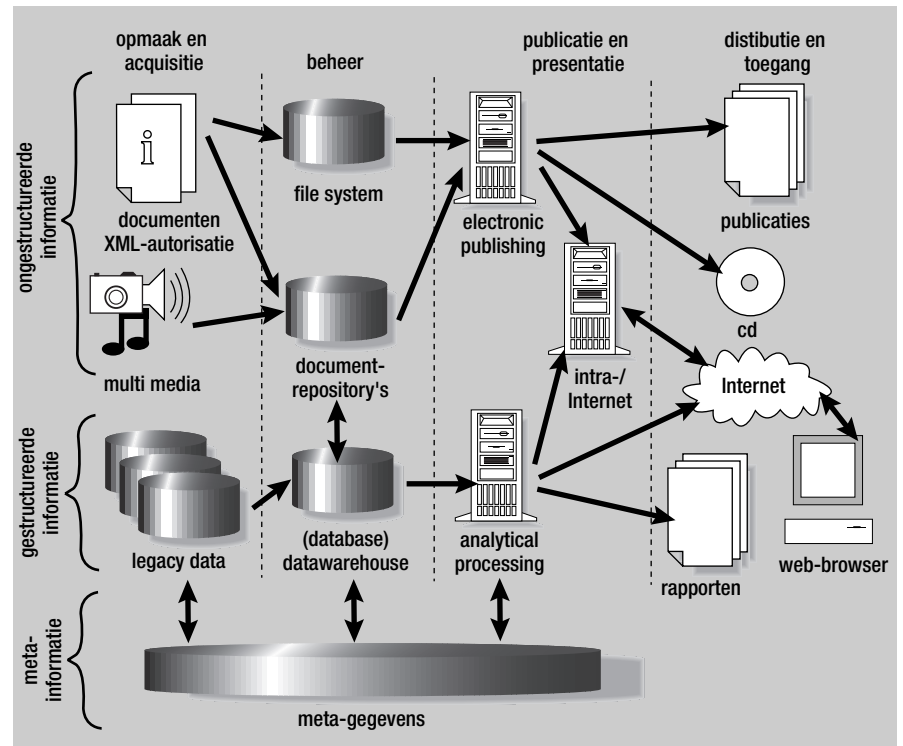
De bodemschatten van de nieuwe economie

Freek Kamst

D Het millenniumspook is verjaagd en iedereen staat weer met beide benen op de grond. De draad van eerder in gang gezette ontwikkelingen wordt opgepakt. Alle strategieën richten zich opnieuw op verbetering van het proces van informatie-uitwisseling. Het accent ligt daarbij op het ondersteunen van het strategisch beslissingsproces, en -in combinatie daarmee- vooral ook het ontwikkelen van een e-business-visie. Maar wat is 'e-business' nu eigenlijk? En hoe belangrijk daarbij is gegevensintegratie?

Begin jaren negentig kwam een trend op om vanuit transactionele gegevens beslissingsondersteunende informatie te winnen. In de laatste jaren van de twintigste eeuw is hieraan beduidend minder aandacht besteed, omdat de automatiseringsbudgetten voornamelijk waren gericht op het zoeken naar mogelijke *millennium bugs*. Deze activiteiten hadden hoofdzakelijk betrekking op de transactionele systemen en kregen als vanzelfsprekend een hoge prioriteit. Toch zijn ook talrijke organisaties bezig gebleven informatie en kennis te putten uit bestaande aanwezige gegevens; de opkomst van datawarehouses en datamarts was hiervan het gevolg.

Ook de opkomst van Internet kunnen we in de geschiedenisboekjes bijschrijven als een belangrijk kenmerk van de jaren negentig. De globalisering van informatie is met dit medium eenvoudig te realiseren en het geeft organisaties totaal andere strategische doelstellingen met betrekking tot hun businessprocessen. Nieuwe doelen, die men de laatste tijd ook wel samenvat onder de noemer e-



FIGUUR 1: VOORBEELD VAN EEN HUIDIGE INFORMATIE-ARCHITECTUUR, ONDERVERDEELD IN VIER FUNCTIONELE HOOFDGBIEDEN.

business. De Gartner-groep beschrijft e-business als de som van e-commerce, customer relationship management (CRM), collaboration-technologieën (integratie van software en middleware), supply chain management (ERP-systemen), business intelligence en kennismangement.

E-BUSINESS = EC + CRM + CT + SCM + BI + KM

De argumenten van de som zullen geheel of gedeeltelijk moeten worden ingevuld,

voordat een organisatie een e-business-toepassing succesvol kan implementeren. Omdat de meeste ondernemingen nog hard bezig zijn met de implementatie van bijvoorbeeld CRM-systemen, zal het, zo verwacht Gartner, nog wel tot 2004 duren voordat de eerste vruchten van het e-business-denken kunnen worden geplukt. Hierbij moeten we opmerken dat de invulling van een e-business-strategie niet alleen de implementatie van een aantal applicaties omvat. Het herinrichten van bestaande en het creëren van nieuwe bedrijfsprocessen vormen een complex en

sociaal gevoelig traject, dat vooraf en tijdens de implementatie niet valt te onderschatten. Ook terugkoppeling en verificatie naar de bestaande bedrijfsprocessen zijn een essentieel onderdeel van de implementatie. Integratie van gegevens tussen de bedrijfsprocessen is hierbij een van de belangrijkste aandachtspunten. Immers, de betekenis van een entiteit als bijvoorbeeld 'relatie' zal binnen alle e-business-processen eenduidig moeten zijn.

FUNCTIONELE ARCHITECTUUR

Voordat we overgaan tot het beschrijven van mogelijke architecturen voor een geïntegreerde informatie-omgeving, geven we eerst een overzicht van de traditionele functionele informatie-architectuur. Hierin onderkennen we vier functionele hoofdgebieden:

- opmaak en acquisitie (capture);
- beheer (management);
- publicatie en presentatie (publishing);
- distributie en toegang (distribution and access).

De hoeveelheid gegevens binnen deze functionele gebieden is de laatste jaren flink toegenomen door de implementatie van nieuwe systemen en processen.

Voorbeelden daarvan zijn CRM-systemen, datawarehouses, BI enzovoort.

Kenmerkend is dat al deze systemen integratie en standaardisatie van gegevens vereisen en een belangrijke invloed uitoefenen op de bestaande processen binnen organisaties. Door de vier functionele gebieden stromen drie informatietypen:

- *Gestructureerde informatie.* De gegevens bevinden zich voornamelijk in de relationele databases van OLTP-systemen (online transactional processing) en liggen voornamelijk opgeslagen in genormaliseerde datamodellen. Hierdoor zijn deze gegevens meestal niet geschikt voor complexe rapportages. Ook de data in legacy-systemen zijn zeer waardevol, maar lastig te ontsluiten. Door deze gegevens te denormaliseren en in datawarehouses of datamarts te laden, worden zij toegankelijker voor de organisatie.
- *Ongestructureerde informatie.* Deze

informatie bevindt zich voornamelijk in documenten, e-mail, foto's, videos en tekeningen. De gegevens worden geproduceerd door tekstverwerkers, groupware, DTP-systemen (desktop publishing) en bijvoorbeeld engineering-systemen en zijn meestal opgeslagen in filesystemen. Maar deze categorie data is een van de belangrijkste houders van bedrijfsinformatie. Vaak is de kennis in deze informatie verborgen en is die slechts zichtbaar te maken door verbanden te leggen met andere documenten en/of gestructureerde gegevens (databases).

- *Meta-informatie.* Metagegevens bevatten informatie over de gestructureerde en ongestructureerde informatie binnen een organisatie. De komst van data-warehousing en BI heeft het inzicht met zich meegebracht dat metadata onmisbaar zijn, gezien de toenemende complexiteit van informatiesystemen. Ze bevinden zich in documenten, repository's, datawarehouses en diverse andere ongestructureerde en gestructureerde bronnen. Een eenduidige vastlegging van metadata is binnen de meeste organisaties nog niet gerealiseerd.

GEÏNTEGREERD DATAMANAGEMENT

Volgens de diverse onderzoeksbureaus staan we aan het begin van een technologische golf, en als ondernemingen en

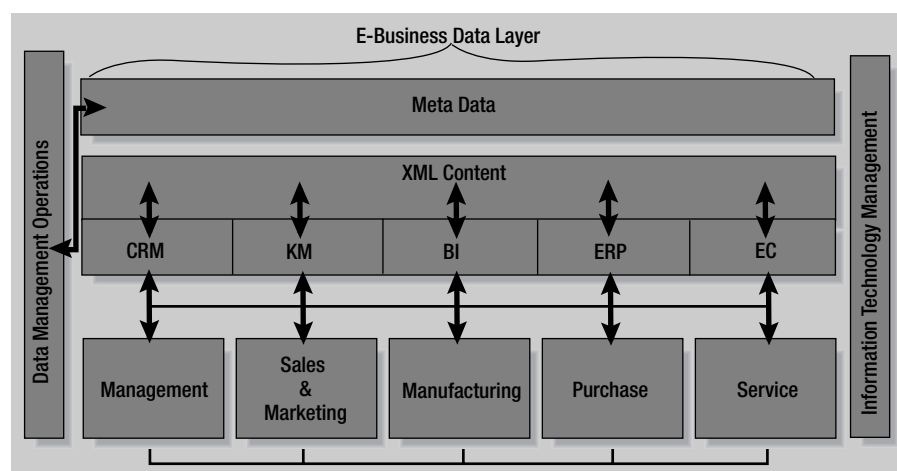
organisaties willen meesurfen op de 'nieuwe economie' die daarop komt aanrollen, zullen zij prioriteit moeten geven aan twee essentiële zaken:

- het vooropstellen van het belang van kennis en informatie;
- het centraal stellen van de klant (*customer asset management*).

Deze stellingen lijken eenvoudig, maar houden concreet in dat organisaties de muren tussen hun afdelingen moeten neerhalen en een nieuw proces tussen stafafdelingen en management creëren. Dit proces (*data management operations*) behelst de integratie van elektronische processen, metadata-beheer en het sturen van de informatiestromen binnen een organisatie. De traditionele IT-afdeling zal zich volledig gaan richten op het technisch beheer van de IT-processen (beheer, implementatie, beveiliging enzovoort).

Tools

Het integreren van informatieprocessen is geen nieuw verschijnsel. Datawarehouse- en BI-projecten zijn voorbeelden van geïntegreerde informatieverwerking, en de valkuilen en ervaringen met dergelijke projecten kunnen van pas komen bij de inrichting van een e-business-omgeving. Inmiddels zijn er professionele tools voor data-integratie op de markt die zijn toegesneden op de behoeften van een e-business-onderneming. Met deze tools kan gegevensuitwisseling tussen diverse processen tot stand worden gebracht zonder dat daarvoor veel inspanning op het



FIGUUR 2: VOORBEELD VAN EEN E-BUSINESS-GEGEVENSSTRUCTUUR EN DE BIJBEHOORENDE INTERACTIE MET DE BEDRIJFSPROCESSEN.

gebied van software-ontwikkeling is vereist. Het gaat om producten van dezelfde leveranciers als die van de tools voor extraction, transformation en loading (ETL) die men bij de exploitatie van data-warehouses en datamarts gebruikt. Veel van de leveranciers daarvan zien nieuwe kansen in de sterk opkomende e-businessmarkt en positioneren hun producten dan ook in deze richting.

Nieuw aspect is de integratie met ongestructureerde gegevens. Deze integratie wordt steeds meer werkelijkheid door de gegevens conform de XML-standaard vast te leggen. De verwachting is dat binnen enkele jaren fileservers zullen hebben plaatsgemaakt voor XML-repository's,

De categorie 'ongestructureerde data' is een van de belangrijkste houders van bedrijfsinformatie

waarin gegevens als tekstdocumenten, video, geluid, foto's en tekeningen zijn vast te leggen. Koppeling van deze gegevens aan transactionele en business-informatie in databases en datawarehouses zal nieuwe mogelijkheden bieden in het verwerken van informatie.

Ook klinkt steeds meer de roep metadata in XML-repository's vast te leggen. Ovum geeft in een van haar laatste rapporten¹ aan dat XML een prima standaard is voor het vastleggen van deze gegevens, maar voegt daar in één adem de waarschuwing aan toe dat XML niet als dé oplossing mag worden beschouwd. XML biedt de mogelijkheid informatie in een uniform formaat met een gemeenschappelijke syntax te beschrijven. Gemeenschappelijke modellen zorgen er vervolgens voor dat de inhoud bruikbaar is voor applicaties, zonder dat elke applicatie de interne structuur van een andere applicatie hoeft te kennen. Door XML als standaardtaal voor data-uitwisseling in te voeren, is de integratie van gestructureerde, ongestructureerde en meta-informatie technisch haalbaar geworden. Alles wijst erop dat steeds meer leveranciers XML als standaard in hun applicaties gaan opnemen, waardoor er op ter-

mijn beheertools en applicaties op de markt zullen komen die beheer en exploitatie van een geïntegreerde gegevensomgeving zullen vereenvoudigen.

KENNISMANAGEMENT

Integratie van gestructureerde en ongestructureerde gegevens draagt onder meer bij tot een optimale exploitatie van de intellectuele 'bodemschatten' van een organisatie. Voegt men daaraan persoonsgebonden kennis toe, dan is er sprake van kennismanagement (*knowledge management*). Dit gebied is een relatief nieuw proces. Volgens de trendwatchers zal e-business in combinatie met kennismanagement binnen luttele jaren nieuwe concurrentieverhoudingen kunnen scheppen. Daarom gaan we hier wat dieper op dit proces in. Waarom moet een e-business-organisatie een proces van kennismanagement exploiteren? Omdat alle organisaties steeds meer worstelen met een omgeving die voortdurend sneller verandert. Door e-business accelereert dit fenomeen alleen maar meer, waardoor beslissingen steeds vaker vanaf de werkvloer door niet-leidinggevenden zullen moeten worden genomen. Informatie en kennis worden dus essentieel op de werkvloer en zullen overal en altijd beschikbaar moeten zijn.

Zeker is dat we kennismanagement moeten positioneren als een bedrijfsproces dat tot taak heeft het beheer van de intellectuele eigendommen van een organisatie te waarborgen. Die eigendommen zijn

onder meer de inhoud van databases, documenten, intranet en andere concrete gegevensbronnen (*expliciete kennis*), maar ook bij de medewerkers aanwezige kennis, ervaring en talenten (*impliciete kennis*) vallen onder deze intellectuele eigendommen. Bij het initiëren van kennismanagement binnen een e-business-omgeving is het aan te bevelen eerst de expliciete kennis te organiseren. Dit omdat concrete gegevensbronnen direct zijn te identificeren en daardoor het resultaat binnen de organisatie relatief snel zichtbaar wordt.

Kennismanagement is een proces dat de gehele organisatie aangaat en valt te vergelijken met de invoering van een CRM-systeem of, bijvoorbeeld, een kwaliteitssysteem. Gemeenschappelijk kenmerk van dergelijke projecten is dat zij niet door een centrale projectgroep zijn uit te voeren; ze moeten in deelprojecten en bij decentrale werkgroepen worden ondergebracht. Een centrale stuurgroep, waarin ook het management is vertegenwoordigd, zorgt voor sturing en beleid, en als alle gebieden zijn onderzocht en geïmplementeerd, kunnen de kennismanagement-taken op diverse plaatsen binnen de organisatie worden ingebed. Men kan het proces in een aantal functionele gebieden onderverdelen. Aan de hand van de omvang van de organisatie kunnen de werkgroepen organiseren zoals is omschreven in het kader *Implementatie kennismanagement*.

Niet alle genoemde functies zijn relevant voor een kennismanagement-proces binnen een organisatie. Waar nodig dient zij een werkgroep te forme-

XML

Extensible Markup Language (XML) is een metataal, dat wil zeggen een taal waarin andere, nieuwe coderingstalen kunnen worden beschreven (bijvoorbeeld HTML). In XML kunnen bijvoorbeeld de coderingsregels voor een adres, product of iedere nieuwe tekststructuur worden vastgelegd. Er wordt dus een documenttype gedefinieerd (een brief, een handboek, prijslijst enzovoort). Hierdoor ontstaat de mogelijkheid de via Internet aangeboden informatie zo te typeren, dat deze voor specifieke informatiebehoefte kan worden ontsloten. XML is bovendien ook zeer geschikt om in de achtergrond als taal voor gegevensuitwisseling dienst te doen. Microsoft heeft bijvoorbeeld haar "active channels" intern gecodeerd in XML. Binnen het gebied van metadata-beheer wordt ook XML als standaard uitwisselingsformaat aanbevolen.

ren. Ook is het niet altijd noodzakelijk automatiseringsmedewerkers in een werkgroep te plaatsen. De 'werkgroep kennisacquisitie' kan bijvoorbeeld bestaan uit medewerkers van de afdeling personeelszaken. 'Data management operations' echter moet telkens een sturend orgaan zijn in het project, om zo de integratie van gegevens en

het beheer van de metadata integraal te waarborgen.

START SMALL, THINK BIG

Snel doorgroeien naar een e-business-omgeving lijkt een schier onmogelijke opgave. Alles in één keer uitvoeren is dan

ook niet mogelijk. Belangrijk is dat het management een visie en strategie creëert, waarin het aangeeft hoe het de eerstvolgende jaren de noodzakelijke e-business-componenten wil gaan invullen. Daarnaast is standaardisering en integratie van gegevens een van de belangrijkste uitgangspunten.

Voor alle besproken gebieden geldt eigenlijk "start small, think big". Dit motto is namelijk in de datawarehouse-ontwikkeling een succesvol ijkpunt gebleken. Alle gebieden die met de invoering van e-business-processen te maken hebben, zijn gebieden die de gehele organisatie raken. En daarmee zijn het projecten met een hoog risico qua omvang en acceptatie. Wanneer de e-business-strategie kan worden vertaald naar concrete plannen, zullen stuur-, project- en werkgroepen worden geformeerd, waarbij de integratie van gegevens als een rode draad door de diverse projectgroepen zal lopen. Om dat te bereiken moet bij elk project 'data management operations' als stafgroep in de projectorganisatie zijn vertegenwoordigd. Zo is sturing van de informatiestromen en uniforme uitwisseling van gegevens binnen de organisatie gewaarborgd.

Kortom: een van de belangrijkste uitdagingen bij de implementatie van een e-business-omgeving is de voor een klantgerichtere werkwijze zo noodzakelijke integratie van gestructureerde en ongestructureerde gegevens. Het management zal duidelijk het belang hiervan constant moeten uitdragen, opdat het draagvlak binnen de gehele organisatie aanwezig blijft. De kreet "There's no business without e-business" zal de komende jaren de lijfspreuk van iedere manager worden.

Referenties:

1. M. Budd, C. Harris-Jones, D. Wells: *Repositories and XML - Technology Choices for Metadata management*. Ovum Ltd, 1999.
2. C. Rozwell, T. Berg: *Strategic Analysis Reports*. Gartner Group 1999.
3. D. Coppens, I. Van de Zand: *Newcom Consulting Group*. Newcom Information Systems 2000. ●

Freek Kamst (Freek.Kamst@newcom.nl) is senior consultant bij Newcom Information Systems.

Implementatie kennismanagement

- *Kennisacquisitie*. Dit gebied heeft tot taak de impliciete (persoonsgebonden) kennis vast te leggen. Vaardigheden, talenten, ervaring en kennis worden centraal vastgelegd. Ook documenten, presentaties en video- en audiomateriaal op dit terrein wordt in dit deelgebied beheerd.
- *Zoeken en bevragen*. Is beperkt tot de expliciete kennis binnen een organisatie. Men houdt zich bezig met het zoeken en opvragen van informatie zonder dat men kennis heeft van de inhoud of locatie van de bron. Transparante toegang tot verschillende gegevensbronnen, zoals datawarehouses, datamarts en/of document-repository's, is een aandachtspunt van dit deelgebied. Momenteel valt deze functionaliteit meestal onder het beheer van de IT-afdeling.
- *Distributie*. Hieronder valt de functionaliteit van het distribueren van informatie naar individuen en/of groepen die daarom verzocht hebben. Distributie kan bijvoorbeeld via e-mail, workflow management-systemen of *push*-technologie via Internet.
- *Samenwerking*. Een van de belangrijkste deelgebieden is samenwerking en integratie in het onderzoeken van procedures en technologieën die de samenwerking en uitwisseling van onderlinge kennis tussen medewerkers stimuleren. Naast groupware zijn natuurlijk oplossingen als video-conferencing een aandachtspunt. Ook het bevorderen van bijvoorbeeld intercollegiale toetsing valt in de focus van dit gebied.
- *Synthese*. Richt zich op het onderzoek naar nieuwe inzichten in de bestaande informatie. Door toepassing van datamining-technieken is uit de bestaande gegevens nieuwe informatie af te leiden die bijdraagt tot kennis. Ook extractie van gegevens uit in- en externe bronsystemen, om daar vervolgens analytische technieken op los te laten (OLAP) is een van de onderwerpen.
- *Profileren*. Onderzoekt of mensen in groepen zijn te verdelen naar gemeenschappelijke interesses, mede om ze te voorzien van relevante informatie. Kan zowel intern als extern gericht zijn.
- *Oplossen en aanbevelen*. Wordt informatie gebruikt voor het nemen van beslissingen (bijvoorbeeld kredietwaardigheid, diagnose van apparatuur etcetera), dan kan men een oplossing of aanbeveling automatisch in workflow management-systemen laten aandragen. In een kennismanagement-omgeving is het mogelijk dergelijke oplossingen en aanbevelingen af te leveren aan groepen of individuen aan de hand van hun profielen. Dit gebied onderzoekt de mogelijkheden en beschikbare technologieën.
- *Onderhoud*. Zorgt voor de procedures en mogelijkheden om kennismanagement te onderhouden. Gegevens zullen voortdurend aan een stroom van wijzigingen onderhevig zijn, en in dit gebied is moet vastleggen hoe het versiebeheer en wijzigingsbeheer is te implementeren. Ook rapportage en signalering van wijzigingen en obsoleete gegevens aan de eigenaars daarvan en het management maken deel uit van dit gebied.