

6.1 De geautomatiseerde integratie van informatie

Drs. R. Hamers

Samenvatting

In dit artikel wordt ingegaan op de integratie van bestaande informatiesystemen. Het artikel begint met integratie binnen de eigen organisatie. Hierin komt naar voren dat bedrijfsprocessen vaak ieder over een eigen systeem beschikken. Geleidelijk aan ziet men pakketten op de markt komen die een groot deel van alle bedrijfsprocessen in één systeem kunnen integreren. Als we doorgaan op deze weg dan komen we op het punt dat we informatie van buiten de eigen organisatie willen integreren met interne informatie. Dat kan gebeuren middels een uitwisseling van een enkel bericht (met behulp van bijvoorbeeld EDI) of door de levering van complete bestanden door externe organisaties. Het integreren van gestructureerde met ongestructureerde data, het integreren aan de hand van aggregeren en de noodzaak van een datadictionary vormen aandachtspunten in dit betoog. Daarna komt de integratie van systemen, netwerken en hardware aan de orde, met bijzondere aandacht voor marketing- en managementinformatiesystemen, omdat in deze systemen diverse integratie-aspecten gelijktijdig optreden. We sluiten af met een verhandeling over privacy-gevoeligheid en specifieke methoden en technieken die betrekking hebben op de integratie van informatie.

geïntegreerd systeem

geïntegreerde data

management informatie-systeem

1 Inleiding

De geautomatiseerde integratie van informatie duidt op een nieuwe richting, die een (geautomatiseerde) informatieverwerking nu en in de toekomst zal gaan volgen. Waar in het verleden geïsoleerde informatiesystemen zijn ontwikkeld, ontstaat thans een sterke behoefte om de informatie-eilanden in eerste instantie met elkaar te verbinden en later uiteindelijk te integreren. Met het integreren bedoelt men vaak 'een opname in een geheel, samen-

PraktijkGids Bedrijfsinformatiesystemen
DECEMBER 1995

voeging' of ook wel 'de harmonische samenwerking van groeperingen, onderdelen'. Geïntegreerde informatiesystemen kunnen derhalve gedefinieerd worden als een groep harmonieus samenwerkende informatiesystemen, waarbij de informatietechnologie een middel is om die integratie efficiënt en effectief tot stand te brengen.

Op dit moment zijn in veel organisaties geïsoleerde informatiesystemen aanwezig, die door allerlei handmatige procedures en ingewikkelde interface-programmatuur met elkaar verbonden zijn. Een overdracht van de eerder ingebrachte gegevens vindt meestal niet plaats. Om een volgende applicatie op te starten moet men vaak opnieuw de inhoud van het complete gegeven intoetsen. Het ontstaan van geïsoleerde informatiesystemen kent een aantal organisatorische oorzaken, zoals bijvoorbeeld:

- men wilde niet betalen voor de door anderen te leveren gegevens of die anderen wilde deze gegevens gewoonweg niet leveren (vanuit de vrees om daarop afgerekend te worden), zodat bepaalde gegevens weer opnieuw geregistreerd en opgeslagen werden;
- het ontbreken van richtlijnen voor uniforme hardware, software en netwerken, zodat verschillende typen computers van diverse leveranciers met afwijkende versies van programmatuur en uiteenlopende (ontwikkel)tools en softwarepakketten met allerlei netwerkprotocollen, in de organisatie hun intrede konden doen.

En daarnaast tevens technische oorzaken:

- met de steeds snellere vernieuwing binnen de informatietechnologie is men in staat om telkens weer modernere systemen op andere platforms te ontwikkelen, die slechts met grote moeite of kosten met verouderde systemen en database-structuren te koppelen zijn;
- de grotere complexiteit en onderlinge afhankelijkheid van met elkaar samenwerkende systemen. Het is makkelijker en sneller om een nieuw systeem te bouwen of een bestaand systeem te modificeren zonder rekening te hoeven houden met de eisen en wensen vanuit andere systemen en gebruikers(groepen);
- geografisch verspreide gebruikers beschikten niet over de adequate netwerkfaciliteiten om gezamenlijk en tegelijkertijd van hetzelfde systeem c.q. dezelfde database gebruik te maken.

In grote lijnen zullen er binnen de eigen organisatie wellicht nog wel meer oorzaken te vinden zijn. Maar in essentie zijn veel organisaties nu in een situatie beland, waarbij de meeste operationele taken geautomatiseerd zijn. Men komt dus nu in een fase terecht waarbij de huidige systemen geïnventariseerd worden en men dan ineens goed beseft hoeveel (geïsoleerde) systemen en separate

gegevensverzamelingen zijn ontstaan, die ieder apart een (overlappend) deel van de informatieverwerking voor hun rekening nemen. Een aantal systeemfuncties kunnen soms meerdere keren voorkomen, omdat men toentertijd dacht dat deze functies weliswaar bijna identiek waren maar voor de eigen situatie niet goed in die vorm waren te hanteren. Dat leidt hopelijk tot het inzicht dat men op geheel andere wijze het bedrijfsproces moet inrichten (Business Process Redesign) en – daaropvolgend – de herziening van de systeem- en gegevensarchitectuur.

2 Binnen de organisatie

2.1 Bedrijfsprocessen in meerdere systemen

Bedrijfsprocessen zijn op een zodanige wijze georganiseerd en samengesteld dat de taken en doelstellingen van de organisatie daarmee gerealiseerd kunnen worden. Deze processen transformeren de binnenkomende input (opdrachten, orders van klanten) naar de gevraagde output (diensten, producten) en zijn weer verder te ontleden in kleinere deelprocessen. Die deelprocessen plaatst men in een organisatiestructuur, zodat divisies en afdelingen weten welke (deel)processen daar uitgevoerd moeten worden. Op een hoog abstractieniveau is er een volgorde te onderkennen in deze deelprocessen, die ieder een fase van het transformatieproces voor hun rekening nemen. Het benaderen van een (nieuwe) klant is veelal de eerste fase, met daarna de acceptatie van de opdracht c.q. order, de omzetting van een order naar een leveringsovereenkomst, het fabricage-, het dienstverlenings- en het uitleveringsproces, en de financiële afwikkeling van dit alles. Afhankelijk van het bedrijfstype zijn hier weer extra fases of andere samenstellingen te onderkennen.

Vanuit de automatiseringshistorie is geleidelijk aan bij elk (deel)proces een daarvoor geëigend informatiesysteem ontwikkeld of aangekocht. Ten behoeve van het bedrijfskritische fabricageproces zijn gedurende de jaren '70 onder eigen beheer (aparte) systemen op het mainframe ontwikkeld voor de uitgifte van werkorders, het beheersen van de voorraad, het bewaken van de voortgang enzovoorts. Soms verschillen deze systemen per producttype, omdat bepaalde producten of diensten een totaal andere verwerking kennen en dus andere informatiebehoeften tot gevolg hebben. Voor de financiële afwikkeling is wellicht kort daarna in de jaren '80 een pakket voor de afdelingscomputer aangekocht met daarin een kostprijscalculatie, de facturering, een debiteurenadministratie en een grootboek. Later kwamen vanaf ongeveer 1990 de kant en klare client/server- en PC-toepassingen de organisatie binnengerold voor de projectenadministratie, de

produktontwikkeling (CAD/CAM), de agendering van de klantbezoeken, de tijdsverantwoording, een tekstverwerker en dergelijke.

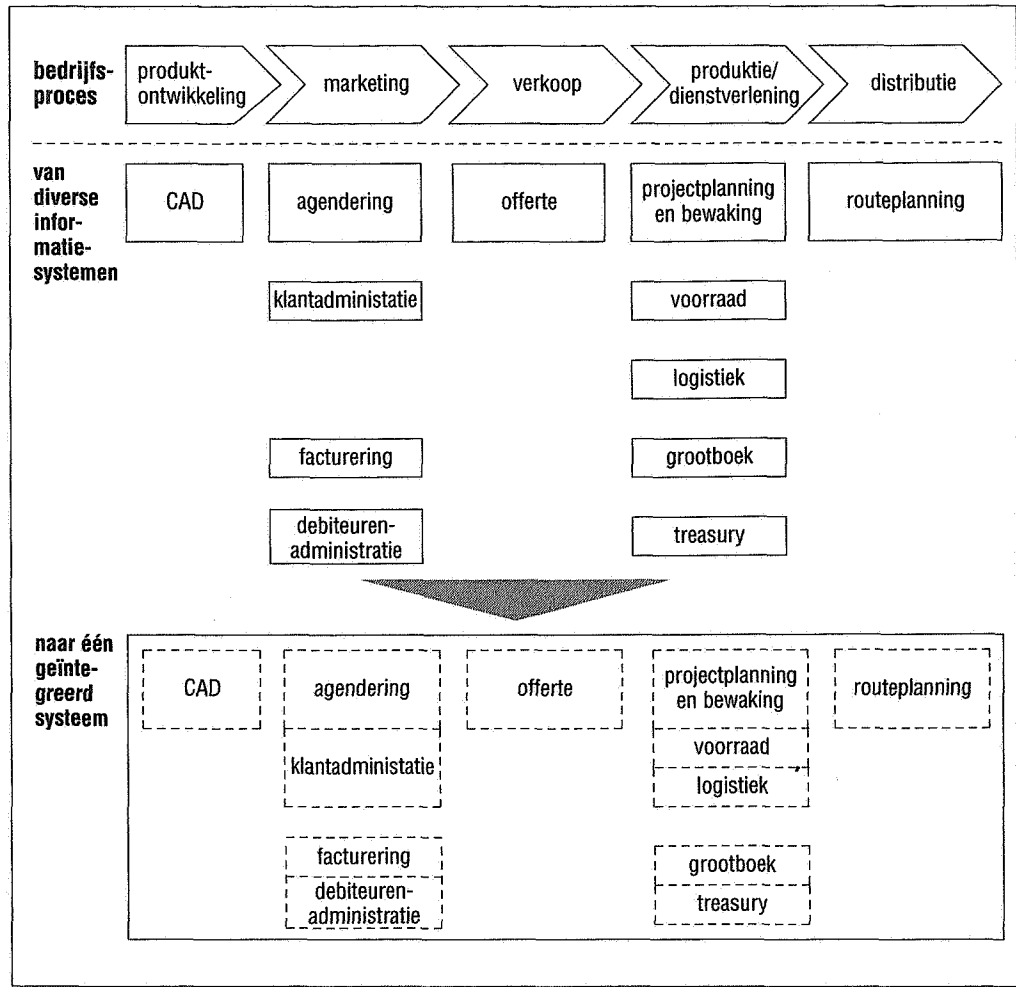
2.2 Koppeling van processen in één systeem

Zo is er op deze manier een verscheidenheid ontstaan aan informatiesystemen op diverse platforms, die elk een deel van het totale bedrijfsproces voor hun rekening nemen. Daarbij is het niet zo dat alle systemen naadloos op elkaar aansluiten of van dezelfde gegevensverzameling gebruik maken.

Een voorbeeld: klantenbestanden kunnen meerdere keren voorkomen. Het oorspronkelijke klantenbestand staat op de afdelingscomputer (voor de facturering), terwijl een historische kopie op een stand-alone PC is geplaatst ten behoeve van de agendering van klantbezoeken. Bovendien zijn aan deze kopie de namen en adressen van potentiële klanten toegevoegd, omdat dat in het originele bestand niet was toegestaan. De toevoeging was mogelijk gemaakt door een slimme medewerker die met een database-pakket op de PC het bestand kon modifieren. Met als gevolg dat hier een sterke redundantie aan gegevens is ontstaan en bovendien het gevaar dat de gegevens van bestaande klanten niet meer actueel worden gehouden in het gekopieerde bestand.

In een niet geïntegreerde systeemomgeving vergt elke processtap het activeren van een ander systeem, waarbij de oorspronkelijke gegevens plus de tussenresultaten telkens opnieuw ingebracht moeten worden. In een geïntegreerde systeemomgeving verloopt dat geheel anders (zie figuur 1). Alle processtappen worden op volgorde binnen één systeem afgehandeld, waarbij de subsystemen op één platform gebruik maken van de reeds beschikbare en zojuist vastgelegde gegevens en de eventuele extra tussenresultaten op elektronische wijze met elkaar uitwisselen.

In deel 6 van deze Praktijkids zal ruimschoots aandacht worden geschonken aan interne integratie van informatiesystemen binnen één organisatie. Dat gebeurt aan de hand van de beschrijving van diverse praktijksituaties, die op succesvolle wijze de interne integratie hebben weten te bereiken. Praktijksituaties die voortkomen uit een herontwerp van het bedrijfsproces (Business Process Redesign (BPR) en Workflow Management (WFM)) en uit situaties waarin men de bestaande systemen en gegevensverzamelingen heeft gefragmenteerd naar een module-achtige constructie, waarin zo min mogelijk overlap en redundantie voorkomt. Ook worden situaties besproken waarin een reeks van systemen of een gecombineerde reeks van systeemfuncties zijn vervangen door een geïntegreerd software-pakket met een database-pakket. Ook de inte-



gratie van analoge informatie (uit fabricageprocessen) en digitale informatie (in de administratieve processen) komt in dit kader aan bod.

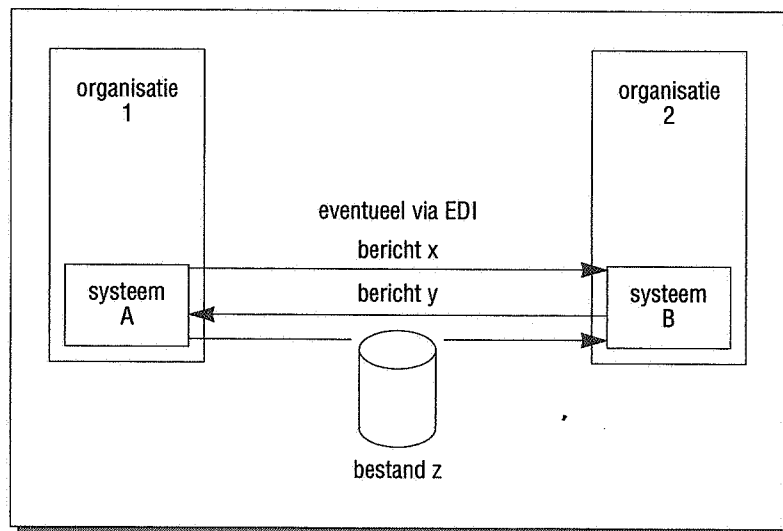
*Figuur 1
Van geïsoleerde naar
geïntegreerde systemen*

3 Over de grenzen van de organisatie

3.1 Berichtenverkeer

De geïntegreerde verwerking van informatie blijft natuurlijk niet beperkt tot de grenzen van de eigen organisatie. Intussen bereiken ook andere organisaties in het binnen- en buitenland een gelijkwaardig niveau van automatisering. Zij hebben databanken opgebouwd die ook voor andere organisaties of personen interessante zaken bevatten. Dergelijke databanken kunnen aanvullende informatie geven over de omgeving waarin zich een organisatie

bevindt, bijvoorbeeld over concurrenten en hun omzetten in bepaalde gebieden of marktsegmenten; misschien zelfs over een dreigend faillissement van een grote debiteur. Het kan ook gaan om informatie die in de eigen databanken reeds was opgeslagen, maar door de toevoeging van externe informatie een completer en duidelijker beeld geeft van bepaalde feiten, gebeurtenissen, organisaties of personen. Dat maakt het zeer aantrekkelijk om bij externe organisaties selectief een beperkt aantal gegevens of gewoonweg een volledig bestand aan te vragen die in de eigen databank geïntegreerd moet worden. Zie hiervoor figuur 2.



*Figuur 2
Integratie over de
grenzen van de
organisatie*

In de praktijk zijn hiervan meerdere voorbeelden te vinden. Een bank controleert online vanuit de eigen applicatie de kredietwaardigheid van een klant bij de BKR (Bureau Krediet Registratie in Tiel) met behulp van een elektronische gegevensuitwisseling. Een leverancier 'kijkt' in het voorraadsysteem van zijn afnemer (een autofabrikant), constateert een tekort aan een bepaald artikel en besluit tot uitlevering van een vastgestelde hoeveelheid (we noemen dit het Just In Time-principe). Het belastingformulier kan door een burger op de PC thuis worden ingevuld en met een diskette naar de Belastingdienst worden verstuurd. Een rekeninghouder vraagt 's avonds vanaf de PC het saldo bij zijn bank op en verzendt op basis daarvan enkele betalingsopdrachten. Misschien zijn we nog niet zo ver verwijderd van het moment dat mutaties op de bankrekening elektronisch overgebracht worden naar een software-pakket, die voor de thuissituatie de financiële administratie verzorgt en waarmee men na afloop van het kalenderjaar (een deel) van het belastingformulier produceert en meteen via een lijnverbinding doorstuurt naar de overheid. De integratie van informatie hoeft dus niet beperkt te zijn tot bedrijven maar kan ook plaatsvinden in de privé-omgeving.

3.2 EDI

Ook met behulp van gestandaardiseerde berichten (bijvoorbeeld met EDI = Electronic Data Interchange) is een integratie te bereiken. In de transportwereld wordt EDI veelvuldig gebruikt om pakbonnen, facturen en andere berichten elektronisch te verzenden. Voor dergelijke berichten zijn in nationaal en internationaal verband afspraken gemaakt over de structuur, de volgorde van de berichtonderdelen en over de gecodeerde inhoud van de gegevens-elementen. Dat vereist vaak langdurige onderhandelingen en procedures om de meeste wensen van de betrokkenen in één veelomvattende norm onder te brengen. Op dit moment hebben verschillende bedrijfsbranches al een aantal eigen normen ontwikkeld. Daarmee zijn de desbetreffende bedrijven in staat om middels speciale interface-programmatuur berichten in te lezen en te versturen naar bedrijven binnen de eigen branche en tevens naar andere partijen, zoals de douane. Deze elektronische berichten worden via een netwerk verzonden en versnellen op die wijze de administratieve procedures en verkleinen de hoeveelheden aan rond te sturen transportpapieren. Daarom hebben ook andere bedrijfsbranches interesse getoond om op grote schaal een gestandaardiseerd berichtenverkeer via een netwerk tot stand te brengen en dat tot norm te verheffen. Vaak beginnen enkele pioniers binnen een bedrijfsbranche voorzichtig met een eigen opzet en hebben dan een zo'n onverwacht groot succes, dat de officiële branche-organisatie niet achter kan en wil blijven.

Een huisarts verzendt bijvoorbeeld zijn uitgeschreven recept via een elektronisch bericht naar een verderop gelegen apotheek, zodat de patiënt bij aankomst in de apotheek meteen zijn geneesmiddelen in ontvangst kan nemen. De desbetreffende branche-organisatie heeft hiervoor al belangstelling getoond en wil dit initiatief in algemene vorm bruikbaar maken.

3.3 Levering van bestanden

Bij organisaties die over faciliteiten voor telebankieren beschikken, bestaat de behoefte om in bestandsvorm de saldi en mutaties op de bankrekening rechtstreeks op elektronische wijze door te sluizen naar bijvoorbeeld het grootboek, naar de treasury en naar andere financiële subadministraties. Binnenkomende betalingen van debiteuren worden geconfronteerd met de openstaande facturen. Dat gebeurt via het factuurnummer en eventueel de factuurdatum in de omschrijvingsrubriek van de ontvangen bijschrijving. Een computerprogramma scant de omschrijvingsrubriek en tracht met het herkende factuurnummer en het rekeningnummer van de betaler de bijbehorende factuur in de debiteurenadministratie op te sporen. Dit wordt ook wel de reconciliatie van betalingen genoemd.

Daarnaast komen ook andere voorbeelden voor. Een bedrijf wil zijn klantenbestand doorlichten of aanvullen met gegevens uit het bestand van de Kamer van Koophandel. De Rijksdienst voor het Wegverkeer moet nagaan of de auto's met een geregistreerd kenteken met minimaal een WA-verzekering zijn voorzien.

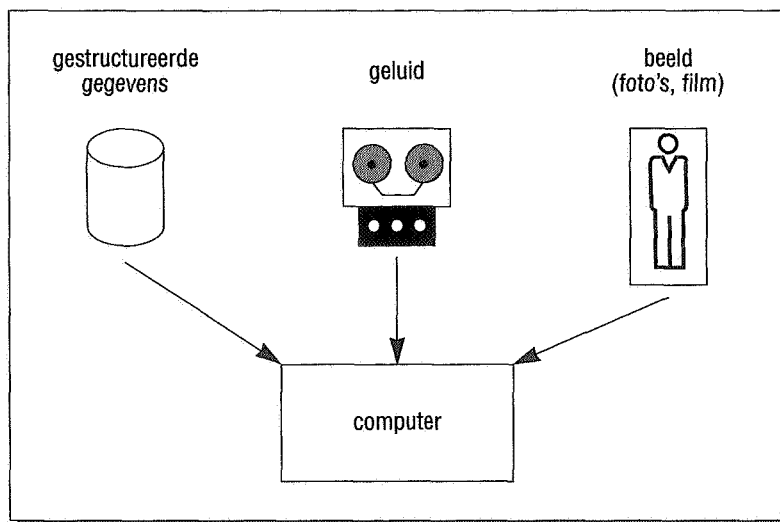
3.4 Bijzondere problematiek

Het integreren van informatie binnen en over de grenzen van de eigen organisatie heen kent voor beide gevallen een bijzondere maar bijna identieke problematiek, waarbij men in sterke mate afhankelijk is van de toeleverende partij of het toeleverende systeem. Telkens komt het probleem naar voren hoe de data uit het ene systeem afgestemd kan worden op de data van het andere systeem; ongeacht of dat andere systeem nu binnen of buiten de grenzen van de organisatie ligt. Bijna altijd wordt men geconfronteerd met ongelijke coderingsstelsels in de gegevenselementen, uiteenlopende identificatiekenmerken, de actualiteit, een onbekende mate van bestandsvervuiling, enzovoort.

4 Data

4.1 (On)gestructureerd

Tot voor kort waren de meeste informatiesystemen ingericht voor de registratie van transacties uit het operationele bedrijfsproces. In grote databases werden de gegevens keurig in aparte bestanden, records en gegevenselementen geordend, gestructureerd en voor wat betreft het aantal geheugenposities begrensd. Elk gegevenselement bevatte numerieke of alfanumerieke waarden of coderingen, die door de programmatuur op een bepaalde discrete inhoud kan worden getest en gebruikt voor het uitvoeren van allerlei berekeningen. Veel database-pakketten kunnen in omvangrijke databases zeer snel een specifiek aangewezen gegeven ophalen, voor andere gebruikers afschermen en de gewenste verandering of verwijdering aanbrengen. Bij integreren van data en systemen, zoals dat in het voorafgaande is beschreven, komt met name de integratie van gestructureerde data naar voren. Maar de geautomatiseerde integratie van informatie kent ook het verwerken van ongestructureerde data (zie figuur 3). Data in de vorm van vrije teksten, tekeningen, afbeeldingen, foto's, film en geluid doen hun intrede in de geautomatiseerde informatieverwerking, waarbij meerdere vormen samen en tegelijkertijd op het scherm of op papier gepresenteerd kunnen worden. Waar in het recente verleden alleen gestructureerde data in databases werd opgeslagen, is het nu nodig om simultaan ook ongestructureerde data daarin op te slaan (inclusief de koppeling tussen de gestruc-



*Figuur 3
Integratie van
gestructureerde en
ongestructureerde
gegevens*

tureerde en ongestructureerde data) of om andersoortige databases te creëren, die daarvoor speciale voorzieningen hebben getroffen. Met verschillende PC-pakketten is men in staat om afbeeldingen en gescande foto's op te nemen in een tekst, eventueel voorafgegaan door een schematische presentatie van gestructureerde gegevens die uit een database zijn geëxtraheerd en met een spreadsheet verder zijn nabewerkt.

4.2 Integreren door aggregeren

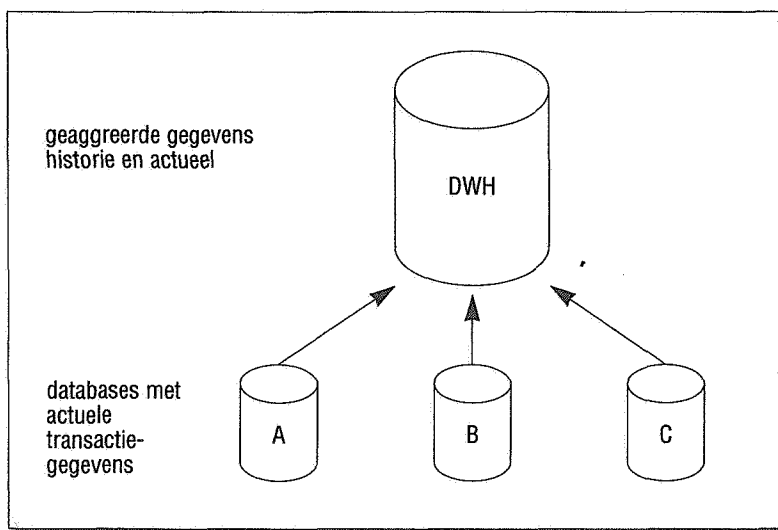
Aggregeren is een bijzondere vorm van integreren. Zoals zojuist is weergegeven, bevatten de meeste databases gestructureerde data van individuele transacties uit het operationele bedrijfsproces. Het betreft dan veelal de meest actuele weerspiegeling van een toestand, feit of gebeurtenis uit de werkelijkheid. Met andere woorden: een database bewaart meestal de laatst gewijzigde inhoud van een gegeven. Vroegere inhoud wordt door de nieuwe inhoud overschreven en raken dus verloren.

Toch kan het zinvol zijn om op vaste momenten (bijvoorbeeld op het einde van iedere maand) gegevens naar bepaalde inzichten te inventariseren, te groeperen, per groep te totaliseren en die totalen vervolgens separaat op te slaan. Bijvoorbeeld: men wenst de totale omzet per klant of klantencategorie, regio en artikel(groep) in te zien. Bij een dergelijke simpele totalisering kunnen niettemin meerdere bestanden betrokken zijn. De omzet is afkomstig uit de getotaliseerde facturen van die maand, de regio en de klantencategorie uit het klantenbestand en de artikelgroep uit het bestand met artikelen. Deze zaak wordt nog wat gecompliceerder als er verschillende klantenbestanden geraadpleegd moeten worden, omdat elke artikelgroep door een andere divisie binnen het con-

cern geproduceerd wordt. Elke divisie heeft naar verwachting op zijn eigen wijze en met afwijkende software zijn klantenbestand opgebouwd en onderhouden. Op deze manier komt door het (geautomatiseerd) aggregeren het integratie-aspect van de intern beschikbare informatie in de volle omvang te voorschijn.

Maar dat integratie-aspect komt ook op een paar andere manieren te voorschijn. Tot welk niveau en met welke doorsnijdingen is namelijk een aggregatie gewenst c.q. echt nodig? Als bij het einde van elke maand die totale omzet per gewenste doorsnijding berekend wordt, dan heeft dat een explosie van geaggregeerde gegevens tot gevolg. Veel organisaties hebben daarom een Data Warehouse opgebouwd om historische standen van geaggregeerde gegevens apart te bewaren naast de databases met de meest actuele transactiegegevens (zie figuur 4).

Figuur 4
Integreren door het
aggregeren naar een
Data Warehouse
(DWH)



Een Data Warehouse heeft op zijn beurt weer typische problemen ten aanzien van de integratie in de tijd. Stel dat een klant met een grote afname verhuist van regio A naar regio B dan kan dat betekenen dat in regio A de omzet van maart een florissante hoogte te zien geeft, maar dat tegelijkertijd de omzet van april dramatisch is gekelderde. In regio B is het tegenovergestelde gebeurd. Moet men dan toch de historische omzetgegevens in het Data Warehouse daarop corrigeren?

Het aggregeren als een bijzondere vorm van automatisch integreren vanuit de interne en in latere instantie ook vanuit externe bestanden, kent dus een aparte problematiek, die ook in deel 6 van deze Praktijkids een uitvoerige behandeling krijgt.

4.3 Datadictionary

De integratie van gegevens uit verschillende bestanden, hetzij intern of extern, legt meteen een probleem bloot dat door velen telkens naar de toekomst is verschoven. Het betreft het structureren en definiëren van (gezamenlijk gebruikte) gegevens, het afbakenen van de gegevensinhoud en het registreren of documenteren van deze zaken in een datadictionary. Vaak is wel de naam van een gegevenselement bekend, maar de precieze betekenis niet. Nog afgezien van het feit dat het domein van de (gecodeerde) gegevensinhoud niet goed is bepaald en de gebruikte codes een onduidelijke interpretatie kunnen hebben. Dat alles kan bij één bestand al een fors probleem zijn, laat staan voor meerdere bestanden die ogenschijnlijk sterk op elkaar lijken en waarbij integratie moet plaatsvinden. Is het klantnummer in het ene bestand hetzelfde als het cliëntvolgnummer uit het andere bestand, ondanks dat het aantal numerieke posities van elkaar verschilt? Wijzen de nummers in die twee gegevenselementen nu dezelfde klant aan?

Daarnaast ontstaan door het aggregeren nieuwe gegevens (in een Data Warehouse) en dus weer nieuwe definities aan de hand van gecumuleerde transactiegegevens. Dat wordt veroorzaakt door het feit, dat volgens de theorie een genormaliseerde gegevensstructuur geen afgeleide gegevens kan/mag bevatten, omdat deze afgeleide gegevens tijdens een verwerkingsproces toch opnieuw berekend of bepaald kunnen worden. Een extra opslag in de gegevensstructuur is dus voor het normaliseren niet toegestaan. In de praktijk is dat natuurlijk anders en bestaan er wel degelijk gegevenselementen met getotaliseerde bedragen en aantallen, maar dat betreft meestal een totalisering naar één dimensie. Bij het aggregeren worden totalen gecreëerd, die naar meer dan één dimensie zijn opgebouwd, zoals bijvoorbeeld een totaal per klanten- en artikelgroep van de omzet vanaf april tot en met juni. Deze geaggregeerde gegevens moeten een eigen naam en definitie krijgen.

Ondanks het feit dat inhouden van databases en bestanden wellicht goed gedocumenteerd zijn (vanuit een datadictionary) is het toch raadzaam om vooraf aan de integratie van de data de desbetreffende bestanden te controleren op juistheid. Het is verrassend en tegelijkertijd frustrerend om te constateren hoe groot de omvang van de bestandsvervuiling kan zijn. Een klantenbestand kan veel te veel klantnummers bevatten omdat één klant meerdere keren voorkomt met een ander nummer. Een gegevenselement dat een bedrag zou bevatten blijkt bij nader inzien een geboortedatum te hebben. Dat was makkelijker en sneller te realiseren dan een systeemwijziging en reorganisatie van de database.

De codering van het bedrijfstype van de klant heeft een andere samenstelling en interpretatie gekregen, zonder dat de programmatuur aangepast hoefde te worden. Uit ervaring bleek dat de nieuwe codering toch door de oude screeningsregels geaccepteerd werd. Niettemin is de inhoud van dit gegevenselement nu gebaseerd op twee verschillende coderingsstelsels.

Dat maakt soms extra conversieprogrammatuur noodzakelijk om de integratie correct te laten verlopen. In het uiterste geval zal eerst het bestand drastisch opgeschoond moeten worden. Het aanscherpen van de screeningsregels is dan een maatregel om de integratieproblemen voor de toekomst zoveel mogelijk te beperken. Het inhoudelijk beheren van gegevensverzamelingen is een aspect dat binnen veel organisaties nog onvoldoende aandacht krijgt. Het automatisch integreren van informatie is niet zomaar het snel aan elkaar koppelen van een paar systemen of het coderen van een interface-programma, maar wel het op orde brengen van de data. En dat kost vaak de meeste (doorloop)tijd en geld (80 à 90% van de totale kosten!). Dus integreren is eigenlijk vooral het opsporen en het in orde maken van de wederzijds benodigde data. Het vertalen van een bepaalde informatiebehoefte naar de bijbehorende werkelijk beschikbare en correcte data, leidt wellicht naar een nieuw specialisme binnen de informatica.

5 Systemen

5.1 Hardware en netwerken

Het integreren van systemen is reeds in paragraaf 2 van dit artikel aan de orde geweest. Deze systemen zijn vaak op verschillende platforms binnen de organisatie aanwezig. Systemen voor het beheren van grote transactiegeoriënteerde databases bevinden zich op het mainframe. Systemen die binnen één of meer afdelingen in gebruik zijn, zijn op een server geplaatst. Enduser tools hebben hun plek gekregen op de PC (de client). Via een netwerk (LAN) worden al deze systemen en platforms met elkaar verbonden middels bepaalde netwerkprotocollen. Het LAN maakt een geautomatiseerde integratie mogelijk van allerlei systemen, die op diverse hardware van verschillende leveranciers draaien. Dat blijft niet beperkt tot een netwerk binnen de eigen organisatie. Via grote (inter)nationale netwerken (zoals Internet) zijn organisaties en particulieren in staat om met elkaar te communiceren en gegevens uit te wisselen. Het sterk toenemende gebruik van Internet spreekt voor zich. Zelfs commerciële bedrijven presenteren zich via dit wereldwijde netwerk en stellen inhoud van databases beschikbaar ter inzage. Op zich is dit natuurlijk niet zo'n verwonderlijke ontwikkeling.

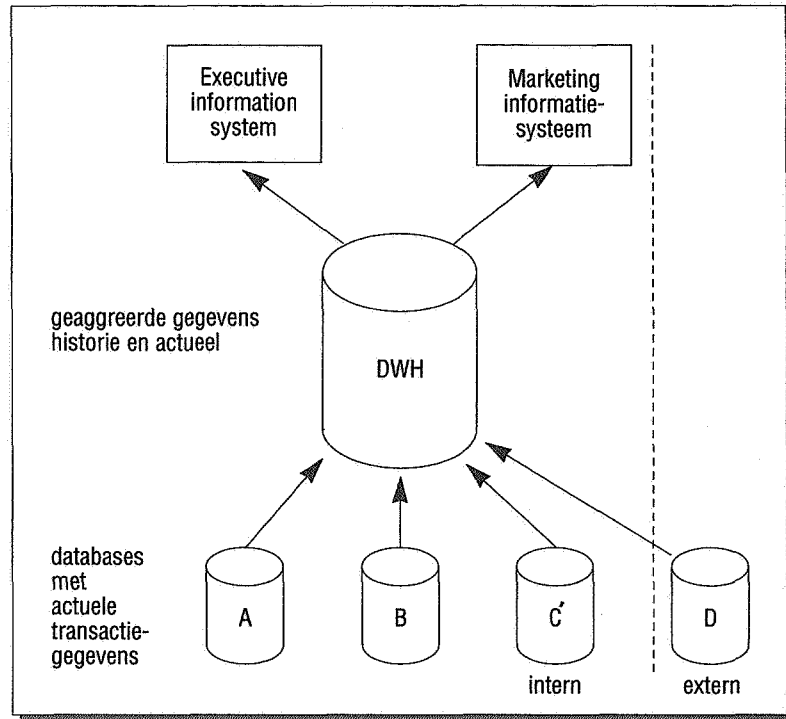
Daarnaast ondergaat de hardware tevens een integratietendens. Steeds meer PC's kan men gebruiken als faxapparaat, om een telefonische verbinding tot stand te brengen, om met elkaar te vergaderen (video-conferencing), voor het produceren van geluid en voor spelletjes. De PC neemt steeds meer functies van andere apparaten over en dat is ook als vorm van een geautomatiseerde integratie te beschouwen. Dat geldt ook in zekere mate voor de (gedeeltelijke) migratie van applicaties vanaf het mainframe naar een client/server-configuratie, waarbij de zwaardere systemen op het mainframe 'achterblijven'. De lichtere systemen worden overgebracht naar de client/server-omgeving, zodat er een betere integratie is te bereiken met software-pakketten en met andere systemen, die dichtbij de eindgebruikers reeds voorhanden zijn.

Op diezelfde PC is nu het besturingssysteem Windows sterk in opkomst. Daar ziet men vele toepassingen en PC-tools onder één besturingssysteem met een heel gebruikersvriendelijke verschijningsvorm op het scherm komen. Die toepassingen kunnen via bepaalde technische voorzieningen rechtstreeks met elkaar communiceren en objecten uitwisselen. Dat is ook weer een vorm van geautomatiseerde integratie.

5.2 Marketing- en Managementinformatiesystemen

Vooraf bij de marketing- en managementinformatiesystemen ziet men veel aspecten van een geautomatiseerde integratie samen en tegelijkertijd optreden. In veel gevallen zullen eerst de interne gegevensbronnen met gestructureerde gegevens aangeboord en geaggregeerd worden om bijvoorbeeld per klant de totale artikelafname over een afgebakende periode in beeld te brengen. We kunnen ook denken aan de totale omzet van alle divisies, onderverdeeld naar regio, klantencategorie en artikelgroep voor het management. Later komen daar nog externe bronnen bij. Het bestand van de Kamer van Koophandel wordt toegevoegd om een indruk te krijgen van de potentiële klanten binnen specifieke regio's. Externe gegevens over concurrenten worden in het managementinformatiesysteem naast de interne cijfers van de eigen omzet gelegd om de groei van het eigen marktaandeel te bewaken. In feite ligt voor de opslag van de grote hoeveelheden aan verzamelde en geaggregeerde gegevens het concept van een Data Warehouse gereed. Van daaruit zal een selectie ter analyse en ter presentatie in een marketing- of een managementinformatiesysteem terecht komen (zie figuur 5). Moderne softwarepakketten (voor het ontwikkelen van een EIS, een Executive Information System) beschikken over speciale multidimensionale database-structuren voor Online Analytical Processing (OLAP: het online analyseren van grote groepen van gelijksoortige gegevens). Die multidimensionale database is te vergelijken met een grote kubus

waarbij in elke coördinaat een gegeven is opgeslagen en elke ribbe van de kubus een dimensie (klantencategorie, artikelgroep, tijd) voorstelt. De gegevens in de kubus zijn via pointers met elkaar verbonden en zijn daardoor zeer snel te benaderen.



Figuur 5
OLAP en Data
Warehouse (DWH)

Voor de opbouw van dergelijke systemen kiest men in de praktijk veelal de weg van prototyping, zodat eerst kleinschalig een aantal integratieproblemen uit de weg kan worden geruimd. Zodra er een goede infrastructuur is aangelegd en de grootste problemen zijn opgelost, ziet men het gehele ontwikkelingsproces pas echt in een stroomversnelling komen. Deel 6 van deze Praktijkgids biedt ook voorbeelden van projecten, waarbij voor een bepaalde aanpak is gekozen, wat geresulteerd heeft in een succesvolle implementatie van een marketing- of een managementinformatiesysteem.

Maar het hoger management is niet alleen geïnteresseerd in gestructureerde gegevens. Er komen nieuwe EIS-pakketten op de markt, waarin tevens ongestructureerde gegevens (teksten, afbeeldingen, geluid) kunnen worden opgenomen, wel of niet gekoppeld aan de aanwezige gestructureerde gegevens.

6 Bijzondere aspecten

6.1 Privacy

Een geautomatiseerde integratie is natuurlijk niet alleen een automatiseringstechnische aangelegenheid. In sterke mate spelen hier privacy-gevoelige zaken in mee. Zeker als gegevens over personen en organisaties de grenzen van de eigen organisatie of het eigen land overschrijden. Te denken valt aan de Airmiles-actie, die vanuit meerdere leveranciers het koopgedrag van de consument in beeld kan brengen. De registratie van de Airmiles is trouwens in een juridisch zelfstandige onderneming ondergebracht, waardoor dus een betere waarborg voor de privacy van de consumenten gegarandeerd wordt. Ondanks de wettelijke voorzieningen op dit terrein blijft deze materie nog volop in beweging, omdat een recentelijk promotie-onderzoek heeft uitgewezen, dat de wetgeving onvoldoende rekening houdt met de nieuwste ontwikkelingen binnen de automatisering. Reden te meer om in deel 6 van deze Praktijkids aan dit aspect aandacht te schenken.

6.2 Methoden en technieken

Het is de vraag of voor het (geautomatiseerd) integreren van informatie officiële methoden en technieken bestaan. Misschien is het makkelijk kunnen integreren een (toevallig) bijproduct van een bepaalde aanpak of zienswijze uit de praktijk (door het gebruik van object-georiënteerde programmeertalen bijvoorbeeld). Niettemin bestaat er aan goede methodes en technieken inclusief een degelijke theoretische onderbouwing duidelijk behoefte, omdat veel organisaties zich zelfstandig op dat terrein van geïntegreerde informatievoorziening gaan begeven. Sommige universiteiten in Nederland hebben dit onderkend en beginnen nu onderzoeksprojecten voor promovendi. Er zijn momenteel nog weinig adviesbureaus gespecialiseerd in het probleemgebied om met behulp van een bepaalde methode een bestaande situatie met vele 'losse' informatiesystemen om te zetten tot een geïntegreerde informatievoorziening.

Beide benaderingen, de vanuit de praktijk gevolgde succesvolle aanpak en de theoretische onderbouwde methoden en technieken, zullen in deze Praktijkids uitgebreid aan bod komen.

7 Conclusies

Uit het voorafgaande is gebleken dat een (geautomatiseerde) integratie van informatie veel invalshoeken kent. Vanuit de historie zijn intern geïsoleerde informatiesystemen gecreëerd, die geleide-

lijk aan een groot deel van het operationele bedrijfsproces geautomatiseerd hebben. Nu heeft men in bepaalde organisaties, organisaties die de eerste fases in de automatiseringsontwikkeling reeds hebben doorlopen, de behoefte om al die systemen op een efficiënte wijze te integreren en beter af te stemmen op het totale bedrijfsproces dat eventueel door een herontwerp een nieuwe structuur krijgt. In een aantal gevallen ligt hier de impuls om de systeem- en gegevensarchitectuur ook onder handen te nemen.

De integratie kent veel dimensies die vanuit de eigen organisaties vorm krijgen op het organisatorische vlak (bijvoorbeeld richtlijnen en voorschriften), de bedrijfsprocessen, de informatiesystemen, de data, de hardware en de communicatienetwerken.

De noodzaak om te integreren kan tevens vanuit de omgeving van de organisatie worden opgedrongen. Het uitwisselen van (gestandaardiseerde) elektronische berichten alsmede de levering van externe bestanden vergen aanpassingen in de interne systemen en gegevensverzamelingen.

Het integreren van gestructureerde en ongestructureerde data, de hardware, het netwerk, het aggregeren als bijzondere vorm van integreren en het gebruik van een datadictionary richten zich met name op de problematiek van de data-integratie. Vaak is hierin onverwacht veel werk verborgen. Meestal denkt men bij integreren van informatie alleen aan systeemintegratie. In de praktijk is de data het medium dat de integratie werkelijk tot stand moet brengen. De gegevens zijn de impliciete communicatietaal binnen de informatievoorziening. Pas als de systemen via de data dezelfde taal 'spreken' (met andere woorden: dezelfde syntaxis en semantiek van de data) is er pas echt sprake van integratie.

Mede op het vlak van de hardware en de netwerken kan een verdere integratie van de informatievoorziening plaatsvinden. Deze faciliteiten maken het mogelijk om systemen en databases een ander platform te geven binnen de technische infrastructuur. Hierdoor kan een bredere en betere beschikbaarheid van data en systemen worden bewerkstelligd.

Marketing- en managementinformatiesystemen brengen met name verschillende facetten van de systeem- en data-integratie bijeen in één probleemgebied, zodat de ontwikkeling van dergelijke systemen een speciale aanpak nodig maakt.

Het integreren van informatie kent naast deze dimensies nog een paar bijzondere aspecten, zoals de privacy van personen en organisaties, alsmede het gebruik van methoden en technieken, die speciaal voor het integreren zijn bedoeld.

Uit het perspectief van het management is het goed te weten wat de voordelen van en de redenen voor een goede integratie zijn:

- een snellere en soepele doorstroming van informatie door de organisatie;
- een verbeterde afstemming tussen de bedrijfsprocessen onderling;
- de diversiteit aan hardware en software en dergelijke wordt teruggebracht naar aanvaardbare proporties, zodat men met minder leveranciers hoeft te onderhandelen over prijs, prestaties en services;
- alle systemen verschijnen op eenzelfde wijze en met een gestandaardiseerde schermopbouw (zoals bij Windows bijvoorbeeld) aan de gebruikers;
- de redundantie van gegevens vermindert en dus ook de omvang van de databases. Gegevens worden eenmalig bij het begin van het bedrijfsproces geregistreerd en meerdere malen gebruikt in de volgende fases van het bedrijfsproces. Dat verhoogt de betrouwbaarheid van de gegevens.

Daartegenover staan de nadelen c.q. bezwaren:

- 1 grotere afhankelijkheid van één of een beperkt aantal (hard- of software)leveranciers, een programmeertaal of een databasepakket met al hun specifieke (on)mogelijkheden;
- 2 grotere complexiteit en een sterkere onderlinge gevoeligheid in relaties tussen de op elkaar afgestemde systeemfuncties, de database en het netwerk;
- 3 de soms forse investeringen om een systeem- en data-integratie te bereiken;
- 4 bij een aangekocht geïntegreerd standaard softwarepakket zal de organisatie zich aan het pakket moeten aanpassen en niet andersom.

Deel 6 van deze Praktijkids gaat in op de vele invalshoeken van een geautomatiseerde integratie van informatie. Een zaak die op dit moment voor veel organisaties een volgende, een vanzelfsprekende en een dringende stap vooruit zal zijn in het groeipad van de automatisering. Dit deel wil in praktische zin deze organisaties een helpende hand bieden.

Auteursgegevens

Drs. Z.F.M. Hamers werkt thans als informatiemanager bij Rabobank Nederland. In deze functie is hij verantwoordelijk voor de architectuur en ontwikkeling van (geautomatiseerde) bestuurlijke informatiesystemen. Hij is tevens werkzaam als bestuurslid van de NGI-afdeling Management & Informatie en als coördinator van de NGI-werkgroep EIS-evaluatie. Hij is tevens lid van de redactie van de *Praktijkids Bedrijfsinformatiesystemen*.

