

Gegevensmanagement

Een handig gegeven

versie 2.0

februari 2006

INHOUDSOPGAVE

1	Inleiding.....	4
2	Trends en ontwikkelingen	7
3	Introductie van Gegevensmanagement	11
3.1	Inleiding Gegevensmanagement.....	11
3.2	Gegevens als bedrijfsmiddel.....	12
3.2.1	Rol van informatie in de organisatie.....	12
3.2.2	Gegevens en informatie	12
3.2.3	Gegevens; het vierde bedrijfsmiddel.....	14
3.3	Wat is Gegevensmanagement ?.....	14
3.3.1	Relatie met het informatiemanagement en architectuur.....	14
4	Beleid van Gegevensmanagement.....	18
4.1	Doelstelling Gegevensmanagement	18
4.2	Strategie	19
4.2.1	Strategie en beleid.....	19
4.2.2	Topdown strategie.....	20
4.2.3	Bottom-up strategie	21
5	Competence Centre Gegevensmanagement	23
5.1	Plaats in de organisatie.....	23
5.2	Relatie met architectuur	25

Gegevensmanagement

1. Inleiding

1 Inleiding

De gegevens zijn de bronnen van waaruit de informatie ten behoeve van de uitvoering en besturing van een organisatie geproduceerd wordt. Informatie over de bedrijfsvoering, over cliënten en producten en over de ontwikkelingen en veranderingen in de omgeving van het bedrijf, zoals de concurrentiepositie, is van strategisch belang om als organisatie succesvol te kunnen blijven opereren.

De gegevens die voor de informatieverzorging nodig zijn, zullen beschikbaar moeten zijn voor de totale organisatie. Dit stelt hoge eisen aan de kwaliteit van de gegevens en vraagt tevens om goede afspraken over de betekenis en samenhang tussen de gegevens en over het gebruik en de beveiliging ervan.

Vandaar dat steeds meer het inzicht groeit dat gegevens een bedrijfsmiddel vormen, dat in een organisatie adequaat bestuurd moet worden.

Na dertig jaren automatiseren hebben (vrijwel) alle bedrijven vele verschillende databronnen. Mede door de noodzaak voor het meer uitwisselen van gegevens binnen een organisatie, maar ook buiten de organisatie wordt steeds duidelijker dat gegevens generiek moeten worden gedefinieerd.

Bij de gegevensgerichte aanpak is standaardisatie van methodes, technieken en hulpmiddelen nodig om onder andere vele synoniemen en homoniemen bij de definities te voorkomen. Door een gemeenschappelijk gebruikte Repository wordt synergie en hergebruik bereikt van gegevens-definities en -modellen.

In de wereld van data warehousing is gebleken dat de "traditionele aanpak" van data integratie middels programmeren en talloze "losse" producten een moeizame weg is. Een andere manier van software ontwikkeling is nodig, waarin metadata en metamodellen een centrale rol spelen. Metadata beschrijven alles wat nodig is om data te kunnen gebruiken.

Het is zo langzamerhand duidelijk geworden dat we data pas efficiënt kunnen beheren en beheersen als we de metadata begrijpen en controleren. Metadata kan van alles zijn: documenten, databases met schema's en inhoud, programma code, business modellen enz. Het begrijpen van de metadata en het kunnen uitwisselen van de metadata tussen verschillende systemen is essentieel: locatie, interpretatie, integratie. Daarom zijn spelregels nodig om dat te bereiken.

De kwaliteit van de gegevens bepaalt de kwaliteit van de informatie.
De gegevens vormen de gemeenschappelijk te gebruiken standaard bouwstenen voor services, berichten en databases om een efficiënte ontsluiting en uitwisseling van gegevens te bereiken.

De problemen ten aanzien van gegevens openbaren zich in diverse situaties:

- Gebruik van gegevens ten behoeve van het kunnen uitvoeren van primaire en ondersteunende business processen.
- Onderhoud en ontwikkeling van informatiesystemen ten behoeve van wijziging van de business processen.
- ICT vernieuwing ten behoeve van strategiewijziging (andere markten, klanten, afnemers, distributiekkanalen, samenwerking in de keten, etc.).
- Beveiliging en privacy van gegevens.
- Beheer en exploitatie van de gegevens.
- Eigenaarschap van gegevens.
- Eisen aan gegevens als gevolg van externe wet en regelgeving.

Gegevensmanagement

2. Trends en ontwikkelingen

2 Trends en ontwikkelingen

De volgende trends en ontwikkelingen zijn relevant bij een gegevensgerichte aanpak.

- Binnen de verzekeringswereld wordt een ketenintegratie gerealiseerd door de administratieve werkprocessen van tussenpersoon tot verzekeraar te uniformeren met een eenmalige invoer van gegevens aan of door de bron (intermediair, consument, werkgever, werknemer). Het doel van deze ketenintegratie is het aantal administratieve handelingen en gebruikte media sterk te reduceren en de concurrentiekracht te vergroten. Voor de maatschappijen en tussenpersonen betekent dit een sterke kostenreductie (15% tot 35% van het operationeel kostenniveau), terwijl de kwaliteit van dienstverlening aanzienlijk verbetert.

Dit doel wordt bereikt door het standaardiseren van *non-competitieve* proceselementen, de uniformering van de infrastructuur en afspraken over te gebruiken technologie.

Standaarden faciliteren de coördinatie van werkprocessen die zich uitstrekken over meerdere organisaties en applicaties (ketenintegratie).

De volgende standaarden zijn ontwikkeld en geïmplementeerd:

- Processtandaarden – Procesatlas
 - Transactiestandaarden - Generieke Interface Manager
 - Gegevensstandaarden - All Finance Model (5000 standaard attributen)
 - Presentatiestandaarden – Stijlgids
 - Transportstandaarden - XML/SOAP/http
- Van statische data warehouses naar "real-time" data warehouses. Warehouses moeten zo ingericht worden, om zo snel mogelijk nieuwe data te kunnen verwerken. Dit is onder andere van belang bij de inzet van Business Intelligence technieken in de e-Commerce.
 - Het samengroeien van operationele systemen en data warehouses, waarbij de inzet van pakketten een grote rol gaat spelen. Dit speelt ook bij het kunnen realiseren van Webtoepassingen en portals. Ook nieuwe database technieken om steeds sneller enorme hoeveelheden data te kunnen verwerken zijn van belang.

- Organisaties besteden ICT-onderdelen uit (outsourcing), zoals de systeemrealisatie en zelfs de exploitatie en het beheer. Standaardisatie van de producten is voor het succes bij een dergelijke beweging essentieel.
- Organisaties ontwikkelen applicaties en/of implementeren pakketten onder een service-architectuur om flexibiliteit en efficiency door ontkoppelpunten in de procesketen te bereiken. Standaardisatie van berichten, services en gegevens is bij deze ontwikkeling van belang voor een efficiënte ontsluiting en uitwisseling van gegevens. Bovendien zal het hergebruik worden bevorderd.
- Gebruik van open standaarden. Enkele voorbeelden zijn:
 - XML (eXtensible Markup Language): een meta taal die elke vorm van informatie definieert en beschrijft, ontwikkeld door W3C¹.
 - XMI (XML Metadata Interchange): XMI is de standaard voor uitwisselen van data en metadata tussen tools, middleware en repositories.
 - CWM (Common Warehouse Metamodel): CWM definieert een gemeenschappelijk metamodel. Het beschrijft zowel de business- als technische metadata. Het kan gebruikt worden voor de uitwisseling van metadata tussen de vele in gebruik zijnde systemen bij data warehousing en business intelligence.
- Model Driven Architecture (MDA); wil applicatie bouw vereenvoudigen door zoveel mogelijk te automatiseren op basis van metadata. Het doet dat op basis van modelleren, waarbij eerst een platform onafhankelijk model van de business wordt gemaakt, een PIM. De PIM wordt omgezet naar een platform specifiek model, een PSM. Dat wordt tenslotte omgezet naar een uiteindelijk systeem. MDA is een nu sterk in ontwikkeling zijnde standaard die de inzetbaarheid van softwaretoepassingen vergroot en de beheerkosten verlaagt. MDA is een uniforme basis voor de ontwikkeling van open, platform- en omgevingsonafhankelijke toepassingen. MDA definieert een nieuwe software architectuur dat integratie en interoperabiliteit mogelijk moet maken, niet alleen tussen afzonderlijke objecten en componenten, maar ook over de hele life cycle van applicaties. MDA gebruikt hiervoor de kracht van Modelling, Metadata, Middleware en Mappings ("the 4 M's").
- Instanties die standaarden beheren en opstellen worden opgericht, als antwoord op de noodzaak van uitwisseling van gegevens op diverse niveau's binnen en buiten organisaties.

¹ World Wide Web Consortium, een organisatie dat zich tot doel stelt technologieën op de markt te brengen die de ontwikkeling van het Web bevorderen ter bevordering van interoperabiliteit

Voorbeelden op nationaal niveau zijn:

- Het SIVI (Standaardisatie Instituut voor Verzekeringen in de Intermediairbranche) een voorbeeld van op nationaal niveau
- OSOSS is de naam van het programma “open standaarden en open source software voor de overheid”. Open standaarden en open source software zijn voor de Nederlandse overheid van belang vanwege de keuzevrijheid, interoperabiliteit en efficiency van haar informatiesystemen. Het programma OSOSS stimuleert overheidsorganisaties open standaarden (ICT) toe te passen in hun informatiesystemen en informeert hen over de mogelijkheden van open source software. Het programma biedt concrete ondersteuning in de vorm van voorlichting, kennisuitwisseling en instrumenten, waarmee elke overheidsorganisatie zelf open standaarden en open source software kan toepassen. Het programma OSOSS wordt uitgevoerd door ICTU in opdracht van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties en het ministerie van Economische Zaken. De begroting van het programma bedraagt 3 miljoen euro voor een periode van 3 jaar. Het programma loopt vanaf 2003 tot en met 2005. Het Programma OSOSS richt zich op twee terreinen; ten eerste open standaarden en ten tweede open source software

Voorbeeld op internationaal niveau is:

- De OMG (Object Management Group, een internationale onafhankelijke non-profit organisatie dat zich tot doel stelt om integratie problemen op te lossen door open, leverancier onafhankelijke specificaties / standaarden te leveren. Bedrijven en kennisinstellingen kunnen lid worden voor de productie en het onderhoud van de specificaties en standaardisering voor de computerindustrie ter bevordering van de interoperabiliteit tussen bedrijfstoepassingen.

De OMG richt zich als non-profit organisatie op de ontwikkeling en het onderhoud van standaarden – zoals CORBA², UML³, MOF⁴, CWM, XMI en ook MDA– voor de software-industrie. Momenteel zijn zo'n 800 grote ondernemingen uit de computerindustrie en honderden kennisinstellingen wereldwijd lid van de OMG. De meeste toonaangevende bedrijven voor de ontwikkeling van het computergebruik voor het internet en in organisaties, zijn vertegenwoordigd in de raad van bestuur.

² Common Object Request Broker Architecture, voor communicatie tussen gedistribueerde objecten

³ Unified Modeling Language, een object georiënteerde analyse en ontwerptaal

⁴ Meta Object Facility, een meta-metamodel, dat als basis dient voor zaken als CWM, UML, XMI, CORBA, enz.

Gegevensmanagement

3. Introductie van gegevensmanagement

3 Introductie van Gegevensmanagement

3.1 Inleiding Gegevensmanagement

Onder trends is duidelijk geworden dat gegevensmanagement van belang is. Maar wat is het precies?

De gegevens zijn de bronnen van waaruit de informatie ten behoeve van de uitvoering en besturing van een organisatie geproduceerd wordt. Informatie over de bedrijfsvoering, over cliënten en producten en over de ontwikkelingen en veranderingen in de omgeving van het bedrijf, zoals de concurrentiepositie, is van strategisch belang om als organisatie succesvol te kunnen blijven opereren.

De gegevens die voor de informatieverzorging nodig zijn zullen beschikbaar moeten zijn voor de totale organisatie. Dit stelt hoge eisen aan de kwaliteit van de gegevens en vraagt tevens om goede afspraken over de betekenis en samenhang tussen de gegevens en over het gebruik en de beveiliging ervan.

Vandaar dat steeds meer het inzicht groeit dat gegevens een bedrijfsmiddel vormen, dat in een organisatie adequaat bestuurd moet worden.

Veelal wordt ook het modelbeheer als onderdeel van gegevensmanagement gezien. Bij het modelbeheer gaat het om het onderhoud van de architecturen (bedrijfsarchitectuur, systeemarchitectuur, technische architectuur en de organisatie architectuur). De modellen die op basis van de architectuur worden vervaardigd, moeten consistent blijven. Wijzigingen zullen noodzakelijk zijn, maar het modelbeheer zorgt ervoor dat de wijzigingen in alle modellen consistent wordt doorgevoerd.

Er zijn problemen met gegevens in de huidige situatie:

- § inefficiëntie en inconsistentie bij gegevensonderhoud en gebruik;
- § de slechte toegankelijkheid en beschikbaarheid van gegevens;
- § de matige overdraagbaarheid en uitwisseling van gegevens;
- § gebrek aan standaardisatie van gegevens;
- § gegevensopslag kan een hinderpaal vormen bij het vernieuwen van de IT.

Al deze problemen hangen met elkaar samen. Oorzaak van deze problemen is dat gegevens beschouwd worden als een onderdeel van een bepaalde functie of toepassing die de gegevens nodig heeft. De gegevens worden dan afgeleid van de toepassing, zonder dat er gekeken wordt of deze gegevens ook door andere toepassingen gebruikt worden.

De gegevens worden gedefinieerd en gestructureerd binnen de context van de toepassing. Dit wordt ook wel de functiegerichte benadering genoemd. Zoals in het hoofdstuk trends beschreven maakt dat het kantelen van een informatiehuishouding (de klant centraal stellen) niet gemakkelijk. Oplossing voor dit probleem is de gegevens vanuit de organisatie en onafhankelijk van de toepassing te beschouwen en centraal te plaatsen. Dit wordt ook wel de gegevensgerichte benadering genoemd.

3.2 Gegevens als bedrijfsmiddel

3.2.1 Rol van informatie in de organisatie

Om te overleven dient een organisatie zich voortdurend te veranderen. De genoemde trends geven daar een opsomming van. De trend zich meer op de klant te willen richten is een reden voor ingrijpende veranderingen.

Bij een toename van het aantal en de complexiteit van informatiesystemen ontstaat er een tweeledige behoefte. Op het operationele vlak wil men de systemen koppelen of integreren zodat een snellere informatie uitwisseling kan plaatsvinden. Daarnaast ontstaat er een toenemende vraag naar besturingsinformatie, zowel voor de operationele processen als voor de consolidatie en verdere uitbouw van de organisatie. Beide aspecten typeren voor een belangrijk deel de rol van de informatie in een organisatie. Bovendien blijft, gesteund door de technische ontwikkelingen, de haast klassieke vraag bestaan naar een snellere wijze van informatieverstrekking. Informatie wordt, mede geïnspireerd door Internet ervaringen, geacht overal meteen beschikbaar te zijn.

Teveel wordt echter stilgestaan bij het kwantitatieve aspect van de informatievoorziening in de organisatie. Dit is gezien de meetbaarheid niet geheel onlogisch, maar toch wel eenzijdig te noemen.

Het belangrijkste obstakel in het normeren van de kwaliteitsaspecten van informatie is niet gelegen in de problematiek van doorlooptijden en dergelijke, maar in de afhankelijkheid van de kwaliteit van de bouwstenen van informatie: *de gegevens*.

3.2.2 Gegevens en informatie

Tussen de begrippen 'GEGEVENS' en 'INFORMATIE' wordt in het normale spraakgebruik vaak geen verschil gemaakt. Ter begripsvorming nemen we aan dat gegevens binnen een bepaalde context een objectieve representatie zijn van feiten. Voor de communicatie zullen deze gegevens eenduidig gedefinieerd moeten zijn. Deze gegevensdefinities worden metagegevens genoemd; om gegevens bruikbaar te representeren kent men er waarden aan toe, die op een geformaliseerde

wijze op een bepaald medium worden vastgelegd. Wij hebben dan te maken met de actuele gegevens. Hier worden de termen, 'gegevens' en 'metagegevens' gebruikt waarbij met gegevens de actuele gegevens worden bedoeld. Informatie is noodzakelijk voor de sturing van de processen binnen de organisatie. Deze informatie is het resultaat van een veredelingsproces dat gebruik maakt van gegevens.

Dit gegevensproces is analoog aan het fabricageproces, waarin grondstoffen getransformeerd worden tot producten. Deze producten kunnen echter halffabrikaten zijn, die dan weer de grondstof vormen voor een verdere veredeling.



Figuur 1. Van gegevens naar informatie en kennis

Kennismanagement (en de kenniseconomie) is hiermee nauw verbonden. Er is een duidelijke tendens in de markt aanwezig steeds meer uit de klantgegevens te halen. Bedrijven bezitten vaak een schat aan niet-toegankelijke gegevens over hun klanten. Deze gegevens kunnen toegankelijk zijn, als ze goed zijn opgezet.

De kwaliteitsafhankelijkheid die de informatie vertoont ten opzichte van zijn bouwstenen (de gegevens), kent een aantal invalshoeken.

De *integratiegraad* van de gegevens ofwel het vermogen de relatie met andere gegevens te onderkennen is zo'n factor. In logistieke zin vormt de toegankelijkheid door de interactiegraad van de gegevens een andere belangrijke factor. De *betrouwbaarheid* van de gegevens vormt de inhoudelijke graadmeter van de kwaliteit van de informatie.

Gegevens zijn in tegenstelling tot informatie stabiel. Informatie is sterk persoonsgebonden en zal afhankelijk van tijd, plaats en functie verschillend zijn. De stabiliteit is van essentieel belang voor het ontwikkelen en beheren van informatiesystemen.

3.2.3 Gegevens; het vierde bedrijfsmiddel

‘ De gegevens in een organisatie vormen een stabiele basis en zijn, tezamen met de andere bedrijfsmiddelen, van strategisch belang voor een adequate bedrijfsvoering. Zij vormen *niet* het exclusieve bezit van een enkele bedrijfsfunctie.’

In steeds meer organisaties treft men een dergelijke beleidsvisie aan die heeft geleid tot de opzet van een Gegevensmanagement functie naast de reeds bestaande administratieve eenheden voor personeel, kapitaal en middelen. Evenzeer als voor de andere bedrijfsmiddelen, wordt gestreefd naar een optimaal gebruik en is een organisatie waarin het gegevensbeheer een centrale rol speelt noodzakelijk. Dit vraagt, zeker in een sterk geautomatiseerde omgeving, een bepaalde mate van gegevensgericht denken.

Zeker voor gegevens wordt beschikbaarheid en kwaliteit vereist om besluitvorming op de diverse organisatieniveaus mogelijk te maken.

Om de stabiliteit van dit bedrijfsmiddel te garanderen is een gegevensbeleid noodzakelijk. Het gegevensbeleid en de procedures met betrekking tot de gegevens kunnen slechts gestalte krijgen indien ze worden toevertrouwd aan een functie in de organisatie: *Gegevensmanagement*.

3.3 Wat is Gegevensmanagement ?

Gegevensmanagement is de besturing en inrichting van de bedrijfsgegevens in een organisatie. Gegevens-management zal vanuit drie invalshoeken worden besproken, te weten de componenten die het Gegevens-management vormen, de relatie met de rest van de organisatie en de relatie met de informatiefunctie.

3.3.1 Relatie met het informatiemanagement en architectuur

Binnen het informatiemanagement, dat verantwoordelijk is voor de informatievoorziening in de organisatie, wordt een aantal aspecten en taakgebieden onderscheiden

De wijze waarop het beleid tot en met het beheer van de gegevens zo goed mogelijk gestalte kan krijgen wordt gerealiseerd en uitgedragen door Gegevensmanagement.

Taakgebieden Gegevensmanagement

Beleid

Beleid is het geheel van doeleinden, uitgangspunten, kaders en randvoorwaarden waarbinnen het ontwikkelen, gebruik en beheer van de gegevens zich in de komende jaren zal richten. In het beleid dient een onderscheid gemaakt te worden tussen functies en gegevens teneinde het gemeenschappelijk gebruik van gegevens te bereiken.

Planning

Planning omvat het informatieplan met de gegevensarchitectuur en een invoeringsplan met de daarin beschreven database architectuur en migratiepaden, uitmondend in op te starten projecten.

Ontwikkeling

Het op basis van de gegevens en database architectuur ontwikkelen en definiëren van gegevens en gegevens-structuren en ontwerpen en bouwen van gegevensverzamelingen.

Bij het ontwikkelen van systemen en gegevensverzamelingen wordt er onderscheid gemaakt tussen:

§ infrastructurale database omgeving

§ applicatieve database omgeving.

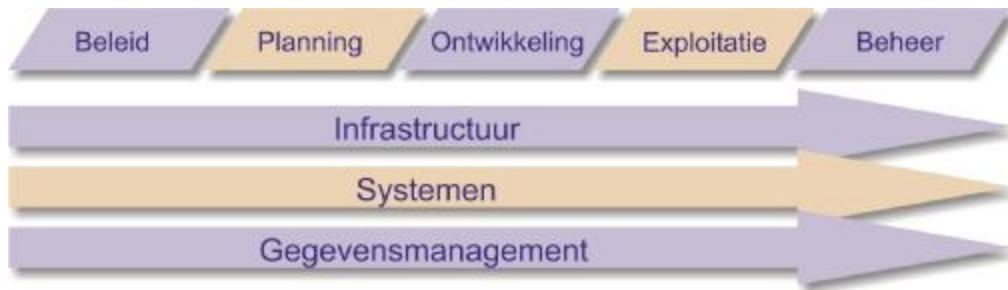
De infrastructurale databases vormen een onderdeel van de infrastructuur van de informatievoorziening. De applicaties zijn verplicht van deze infrastructurale voorziening gebruik te maken.

Beheer

Voor het beheer van gegevens kan men over het algemeen een 3 tal niveaus onderscheiden, te weten het 'corporate', divisie en systeem niveau. Op het 'corporate' niveau zijn de normen en richtlijnen belegd. Afhankelijk van de mate van het gemeenschappelijk gegevensgebruik zal het gegevensbeheer op één van de niveaus georganiseerd zijn.

Om het besturen van dit totale beheer mogelijk te maken is een administratie van de gegevens noodzakelijk. Hierbij wordt het hulpmiddel Data Dictionary ingezet om de beschrijvingen van en de relaties tussen de gegevensgerichte producten tussen de verschillende niveaus te registreren en informatie hierover beschikbaar te stellen. Ook kunnen producten tijdens het ontwikkeltraject met behulp van dit hulpmiddel gegenereerd worden.

Gegevensmanagement speelt dus een rol in alle fasen en taakgebieden van informatiemanagement. In onderstaande figuur is dat schematisch weergegeven.



Figuur 2. De relatie tussen informatie- en gegevensmanagement

Gegevensmanagement

4. Beleid van gegevensmanagement

4 Beleid van Gegevensmanagement

4.1 Doelstelling Gegevensmanagement

Een zo groot mogelijke mate van gemeenschappelijk gegevens gebruik staat centraal, met als resultaat een aantal betrouwbare gegevensbanken die als onderdeel van de database infrastructuur openstaan voor alle geautoriseerde systemen en eindgebruikers.

Hierdoor wordt bereikt:

1. productiviteitsverbetering:
zowel bij het ontwikkelen als bij het onderhoud van systemen;
2. kwaliteitsverbetering:
van de informatieverzorging ten behoeve van de uitvoering en besturing van de bedrijfsvoering door betrouwbare en consistente gegevens;
3. kennismanagement wordt mogelijk
opgeslagen gegevens worden toegankelijk gemaakt;
4. het wordt mogelijk gegevens gedistribueerd ter beschikking te stellen.

De productiviteit bij het ontwikkelen en onderhouden van informatiesystemen is laag. Men spreekt zelfs van een 'software crisis', dat wil zeggen:

- § geringe effectiviteit van het ontwikkelproces;
- § systemen voldoen niet meer aan de gebruikers eisen;
- § onderhoud vergt veel tijd.

Oplossingen hiervoor zijn nieuwe en krachtiger talen.

Een belangrijke voorwaarde bij deze oplossingen is vaak een goed gedefinieerde gegevensstructuur, zoals ook in de inleiding van deze publicatie weergegeven. Een goed ingevoerd Gegevensmanagement is een noodzakelijke voorwaarde zodat een van de bekendste problemen in de organisatie namelijk doublures van gegevensopslag (redundantie) zo veel mogelijk beperkt wordt. Deze doublures in gegevensopslag leiden niet alleen tot extra kosten voorontwikkeling en onderhoud, maar ook tot inconsistentie van gegevens.

Een kwaliteitsverbetering van de informatieverzorging wordt bewerkstelligd door:

- § integratie van gegevens;
- § gegevens en de toepassingen te scheiden;
- § uniformiteit van gegevens;
- § standaardisatie van coderingen.

De eindgebruikers zullen steeds meer te maken krijgen met automatisering en de afstand tussen gebruiker en automatisering zal steeds kleiner worden. Er zullen eenvoudige en gebruikersvriendelijke hulpmiddelen komen, die de gebruiker zelf in staat stellen zijn eigen ontwikkelingen te maken om aan zijn informatiebehoefte te voldoen. Dit zal leiden tot een groeiende vraag naar gegevens en tot een eigen ontwikkelomgeving. Copy Management zal voor de uitwisseling van de gegevens tussen beide ontwikkelomgevingen zorgen.

De doelstelling van Gegevensmanagement heeft zeker niet alleen betrekking op de geautomatiseerde omgeving. In feite richt de doelstelling zich in het algemeen op het verbeteren van de communicatie in de organisatie. Bij communicatie tussen mensen en/of machines is het belangrijk dat men dezelfde taal spreekt. De betekenis van de gegevens moet voor iedereen duidelijk zijn en dezelfde uitleg hebben. Gegevensmanagement heeft een stimulerende en coördinerende rol om tot een éénduidig begrippenkader binnen een bedrijf te komen. Hierdoor kunnen misverstanden zoveel mogelijk worden voorkomen. Babelonische spraakverwarring doordat begrippen in een organisatie niet juist geïnterpreteerd worden, leidt helaas vaak tot vervelende en kostbare fouten

4.2 Strategie

4.2.1 Strategie en beleid

In het gegevensbeleid dient een aantal strategieën benoemd te zijn om het beleid te realiseren. Deze strategieën zijn gerelateerd aan het korte en lange termijn beleid. Voor het lange termijn beleid kan een Top Down Strategie worden gehanteerd, terwijl het korte termijn beleid in een Bottom Up Strategie kan worden vervat.

4.2.2 Topdown strategie

Bij het formuleren van de informatiestrategie en de informatieplanning moet er voldoende helder inzicht zijn in de betekenis en samenhang van de gegevens in de organisatie of een bepaald gebied ervan.

Een belangrijk gegevensgericht product, dat tijdens het informatieplanningstraject opgeleverd wordt, is de gegevensarchitectuur. De belangrijkste doelstellingen van een gegevensarchitectuur zijn:

- § het verkrijgen van een overzicht van de gegevens, die in de organisatie of een deel ervan worden gebruikt, en hoe dit gebruik is;
- § het verkrijgen van inzicht in de betekenis en samenhang van de gegevens, los van specifiek gebruik door functies of afdelingen;
- § het onderkennen van het gemeenschappelijk gegevensgebruik in de organisatie;
- § het weergeven van de wijze, waarop gegevens functioneel en geografisch gedistribueerd zijn;
- § het gebruik van een gegevensarchitectuur als middel om de infrastructuur van gegevensopslag voor nieuwe systemen te kunnen realiseren volgens een planmatige aanpak;
- § het scheppen van een kader, waarvan projecten gebruik kunnen maken bij het voortraject van de systeemontwikkeling.

Naast deze doelstellingen heeft een dergelijke architectuur ook andere voordelen, namelijk het toepassen van een subjectgerichte benadering.

Gegevens zullen bij een subjectbenadering naar de belangrijkheid van onderwerpen worden gepland en ingericht, zodat eenieder, die daartoe gerechtigd is en iets over een betreffend onderwerp wil weten, gebruik kan maken van de gemeenschappelijke gegevensverzameling.

Uit onderzoeken is gebleken dat 80% van de gegevens gemeenschappelijk gebruikt kan worden door verschillende informatiesystemen. Een subjectbenadering stelt echter hoge eisen aan de flexibiliteit van de technische infrastructuur, vanwege allerlei gebruiksmogelijkheden. Zaken zoals gegevensdistributie naar decentrale computersystemen, verstrekken van ad hoc informatie, etc. dienen met behulp van één of meer infrastructurele databases gerealiseerd te worden. Dit impliceert dat de eisen, die gesteld worden aan de structuur van de gegevens en aan de definities van de gegevens, zeer hoog zijn.

Ook de informatiearchitectuur met daarin weergegeven de conceptuele informatiesystemen en de onderlinge relaties is een zinvol hulpmiddel om het gemeenschappelijk gegevensgebruik te kunnen vaststellen. Uiteraard moeten de genoemde producten wel opgeleverd worden tijdens de informatieplanning en is de vorm en de naamgeving van deze producten sterk afhankelijk van de methode en techniek, die men hanteert bij het informatieplanningsproces.

4.2.3 *Bottom-up strategie*

De Bottom Up Strategie wordt gebruikt om het korte termijnbeleid te verwezenlijken, waarbij het op een snelle wijze verkrijgen van rendement de belangrijkste doelstelling is. Vanuit bestaande gegevensverzamelingen wordt het gemeenschappelijk gegevensgebruik en de mate van overtolligheid van gegevensopslag vastgesteld.

Dezelfde gegevens zullen in meerdere gegevensverzamelingen opgeslagen zijn, hetgeen een enorm productieverlies kan veroorzaken. Productieverlies in de vorm van schijfopslag, computertijd, ontwikkelings en onderhoudscapaciteit.

Bij de analyse van de huidige gegevensverzamelingen kan een Data Dictionary/Directory Systeem (DD/DS) een uitstekend hulpmiddel zijn. Zeker wanneer bestaande systemen via dit hulpmiddel gedocumenteerd zijn en van waaruit de sources van de gegevensverzamelingen, zoals databasebeschrijvingen (afhankelijk van het type DBMS) en Data Divisions worden gegenereerd. Ook wanneer een DD/DS pakket aanwezig is en dit nog niet gebruikt is, kan dit hulpmiddel als kladblok zinvol worden ingezet ten behoeve van de scanning van de huidige gegevensverzamelingen. Aangezien de DD/DS cross-reference-faciliteiten heeft kan een analyse van de gegevensverzamelingen betrekkelijk snel worden uitgevoerd. Bij een dergelijke activiteit moet men alert zijn op synoniemen en homoniemen, omdat veelal een standaard naamgeving niet werd gehanteerd.

Zodra het gemeenschappelijk gegevensgebruik is vastgesteld, wordt voor een infrastructurele database gekozen, die zodanig gefaseerd ontwikkeld en geïmplementeerd wordt, dat deze flexibel uit te breiden is. Dit betekent dat er noodvoorzieningen (interfaces tussen conventionele bestanden en databases v.v.) getroffen moeten worden om de huidige systemen met hun applicatieve gegevensverzamelingen te migreren naar de subjectbenadering. Dergelijke oplossingen worden veelal toegepast om bijvoorbeeld een cliënten-gegevensverzameling stapsgewijs naar een infrastructurele database te migreren.

Gegevensmanagement

5. Competence Centre Gegevensmanagement

5 Competence Centre Gegevensmanagement

5.1 Plaats in de organisatie

De gegevensmanagementfunctie kan op verschillende manieren in de organisatie worden belegd. Waar dat gebeurt, is sterk afhankelijk van het beleid en vooral van het type bedrijf.

De ervaring leert dat er meestal sprake is van een groeitraject, waarbij de plaats van gegevensmanagement samenhangt met de ontwikkeling van de informatiefunctie in een bedrijf.

De stadia kunnen de volgende zijn:

1. Toepassingsgerichte bestanden

In deze fase concentreert men zich op de technische aspecten van de gegevens. Binnen projecten worden bestanden ontwikkeld en geïmplementeerd. Er is weinig gemeenschappelijk gegevensgebruik, zodat er redundantie is. Gegevens en toepassingen worden door een systeembeheerder beheerd.

2. Databases

In deze fase wordt geprobeerd een aantal problemen uit fase 1 op te lossen. Er wordt een databasemanagementsysteem geïntroduceerd. Er worden dan toepassingsgerichte database ontwikkeld, die qua structuur sterk afhangen van de betreffende toepassing. Definities en standaarden worden op beperkt niveau ingevoerd. De database administrator (DBA) doet zijn entree.

3. Data administration

In deze fase wordt getracht boven het systeemniveau uit te stijgen. Informatieplanning wordt gestart en gegevensgericht denken wordt gestimuleerd. De behoefte aan een data-administrator ontstaat. Soms, afhankelijk van de grootte van een bedrijf, wordt ook een data dictionary administrator ingevoerd.

Het is hierbij van belang DA en DBA goed met elkaar te laten samenwerken. De beide functies kunnen in concurrentie met elkaar treden. In deze fase van ontwikkeling kan dat leiden tot het weer verdwijnen van de DA functie.

Daarom is het creëren van een samenwerkingsverband tussen DA en DBA van belang. Dat kan ertoe leiden dat er een afdeling gegevensmanagement wordt gecreëerd die bestaat uit DA DBA en DDA. Voorwaarde is wel dat de organisatie hier aan toe is.

4. Subjectbenadering

In deze fase wordt de subjectbenadering nagestreefd, waardoor het inhoudelijk beheer in de gebruikersorganisatie moet worden geregeld. Er kan een indeling in resultaatgebieden worden opgezet. Deze komt dan in de plaats van de traditionele indeling in DA, DBA, DDA.

5. Corporate en decentraal gegevensmanagement

In dit stadium wordt een onderscheid gemaakt in het beheer van metagegevens op corporate- en op bedrijfszonderdeelniveau. Vanuit het corporate niveau worden normen en richtlijnen voor de totale organisatie uitgevaardigd. Op dat niveau vindt tevens het beheer van de corporate metagegevens plaats. Gegevensmanagement wordt op gelijk niveau gesteld met de automatiseringsafdeling binnen een informaticadivisie. Voorwaarde voor deze inbedding is dat de organisatie gegevensmanagement als volwaardig accepteert.

6. Competence centre

Uiteindelijk kan er een competence centre voor gegevensmanagement worden ingericht. Van daaruit vindt centrale sturing plaats.

Bij het doorlopen van de verschillende stadia is het belangrijk dat de interne structuur van gegevensmanagement een afspiegeling is van de structuur van de totale organisatie. Wanneer een organisatie sterk gedecentraliseerd is in bijvoorbeeld profit centres, zal gegevensmanagement in dezelfde mate gedecentraliseerd georganiseerd moeten worden opgezet. Dit betekent dat de ontwikkeling van gegevensmanagement binnen de organisatie meer als een trendvolger dan een trendsetter plaatsvindt.

5.2 Relatie met architectuur

Door Don Tapscott en Art Caston zijn enkele architectuurprincipes opgesteld. Sommige daarvan zijn van invloed op het onderwerp gegevensmanagement:

- § klant gericht. De architectuur beslissingen moeten gericht zijn op het opleveren van klant georiënteerde systemen. Dat zijn systemen die een antwoord zijn op een klant-wens;
- § consistentie: als er gemeenschappelijke eisen zijn, dan worden die steeds hetzelfde behandeld;
- § modulariteit: er zullen steeds standaard modules worden gehanteerd, met standaard interfaces, zodat er flexibiliteit en herbruikbaarheid ontstaat;
- § er worden open standaards gebruikt;
- § er wordt een architectuur opgezet die het integreren van standaardpakketten mogelijk maakt;
- § gebruik een gestandaardiseerde user interface;
- § alle informatie wordt door gegevensmanagement onderhouden om gemeenschappelijke definities te creëren, alsook om reden van consistentie.

Deze principes zijn in lijn met de trends die in hoofdstuk 2 van deze publicatie zijn genoemd.



Figuur 4. Centrale positie gegevensmanagement