

EEN RAAMWERK VOOR INFORMATIEDIENSTEN

Theo Huibers, Paul van der Vet en Bernd Wondergem

1 Inleiding

De opkomst van Internet eind jaren negentig van de vorige eeuw leek het bestaansrecht en de toegevoegde waarde van een uitgeverij ernstig te ondermijnen. In de nieuwe informatie-economie was 'the medium not the message' (Negroponte 1995), maar ging het om de content! Om de digitale klant optimaal te kunnen voorzien in zijn zakelijke en private informatiebehoefte gingen bedrijven uit verschillende disciplines samenwerken. Uitgeverijen, kabelbedrijven, de entertainment-industrie, technologische innovators, internetbedrijven, iedereen leek met iedereen samen te gaan. Allen met dezelfde digitale droom om te convergeren naar een nieuw type organisatie waarin de content, de services en de distributie optimaal op elkaar zijn afgestemd.

Helaas bleek de praktijk weerbarstiger en vele acquisities en investeringen werden binnen korte tijd ongedaan gemaakt. Elk bedrijf had op zich een toegevoegde waarde in de informatie-economie maar het businessmodel van een gecombineerde aanpak was kennelijk niet de som van de autonome toegevoegde waarden. Het 'een plus een is drie' verhaal ging hier niet op; sterker nog 'een plus een werd minder dan twee'.

Na de woelige jaren van de internethype lijkt de rust wedergekeerd. Uitgeverijen denken, ieder jaar weer, met de nodige ambitie na over hun internettoekomst maar nemen slechts gereserveerd stappen in deze richting (zie Huibers en Timmers, 2004). De recessie in de ICT-markt heeft ervoor gezorgd dat ook ICT-dienstverleners niet staan te springen om nieuwe initiatieven en de distributiebedrijven kijken vooral de andere kant op (voornamelijk richting de energiemarkt).

Maar toch... in de huidige informatiemaatschappij draait alles om het vergaren en verwerken van informatie, informatie die betrouwbaar, tijdig, volledig maar vooral relevant moet zijn. De technologische barricades worden langzaam maar zeker één voor één geslecht. De penetratiegraad van het computergebruik loopt gestaag richting de 100% en de gebruikersmarkt voor digitale informatie en digitale dienstverlening via internet beperkt zich niet meer tot enkele studentikoze computernerds. Bedrijven zoals Google, Overture en Amazone spinnen op dit moment het meeste garen in de nieuwe informatiemarkt. Maar wat is de rol van de uitgeverij? Welke keuzemogelijkheden zijn er voor een uitgeverij om zijn positie in de informatiemaatschappij te handhaven of te

versterken? En, omdat allerlei bedrijven en instellingen die traditioneel niet als uitgever worden gezien zelf materiaal beschikbaar stellen: wie of wat is een uitgeverij?

In dit artikel introduceren wij een nieuw kader, gebaseerd op nieuwe uitgangspunten van de informatie-economie. De waarde van dit kader ligt in een duidelijke definitie van de begrippen rondom informatiediensten. Daarbij worden de onderlinge verbanden helder aangegeven. Dit is belangrijk in het werkveld van informatiediensten, waarin immers vaak sprake is van begripsverwarring. Door de toenemende samenwerking tussen organisaties, ook bij het leveren van informatiediensten, is een model van belang dat gebaseerd is op de aangeboden diensten. De diensten vormen immers een belangrijke (externe) schakel tussen de organisaties die samen een informatiedienst leveren.

In de praktijk kan het kader fungeren als discussiemodel voor nieuwe actoren die daarin een rol willen spelen of voor bestaande actoren (zoals de uitgeverijen) die willen of moeten transformeren om een nieuwe rol te spelen. Het doel van het kader is het helpen van de verschillende actoren bij het bepalen van een passende rol. Dit kan eventueel in samenwerking met andere actoren gedaan worden.

In dit artikel verkennen we de mogelijkheden van een dergelijk raamwerk vanuit een economisch perspectief. Op welke wijze kunnen bestaande of nieuwe informatie-intermediairs hun positie kiezen aan de hand van het raamwerk. De toepasbaarheid van het kader wordt aangetoond door de bestaande discussie rondom uitgeverijen en hun toegevoegde waarde erin te schetsen en daaruit conclusies te trekken.

In hoofdstuk 2 schetsen we eerst de kaders en de definities van informatiediensten. In hoofdstukken 3 en 4 worden de begrippen informatiedomein, taken, toepassingen en informatiedienst neergezet. In hoofdstuk 5 wordt het raamwerk toegepast in een economisch domein. Hoofdstuk 6 beschrijft de ICT die nodig is voor het realiseren van informatiediensten.

2 Informatiediensten

Informatie is van alle tijden. De mens heeft altijd beschikking gehad over informatie en heeft altijd informatie gebruikt. De manier waarop informatie wordt gebruikt – wat er met de informatie wordt gedaan – is in de loop van de tijd aanzienlijk veranderd, onder andere door ontwikkelingen in de ICT. Door deze ontwikkelingen kan niet alleen de eigenaar van informatie diensten op zijn content kan aanbieden, ook andere, eventueel concurrerende, partijen kunnen dat. Bijvoorbeeld www.kranten.com aggregeert kranteninformatie en biedt het resultaat vervolgens als unieke dienst aan.

Dit betekent dat er nieuwe spelers in de informatiemaatschappij ontstaan (Huibers 2004). Zowel nieuwe als oude spelers moeten twee belangrijke vragen beantwoorden: welke

informatie bezitten we c.q. willen we bezitten en wat willen we ermee doen? (Bezitten is overigens, zoals we verderop zullen zien, in een digitale omgeving een flexibel begrip; bereikbaarheid is vaak al een voldoende voorwaarde.) De twee vragen gelden zowel voor de aanbod- als voor de vraagzijde van informatie. De dialoog tussen aanbieder en gebruiker van informatie richt zich dus op deze twee aspecten.

We introduceren daarom het begrip informatiedienst, waarmee we beschrijven welke informatie aangeboden wordt en met welke functionaliteit. In de volgende hoofdstukken ontleden we het model van een informatiedienst.

3 De bouwstenen: informatiedomein, taken en toepassingen

De basis van ons raamwerk wordt gevormd door het informatiedomein. Hiermee wordt de (inhoudelijke) informatie bedoeld waarover de informatiedienst beschikt, bijvoorbeeld in de vorm van content in intellectueel eigendom of content die bereikt kan worden, al dan niet tegen kosten. Het gaat om de opgeslagen informatie, het medium doet er op zich niet toe. Het kan een papieren stapels memo's zijn of een computersysteem met databestanden. In een digitale netwerkomgeving is de exacte locatie van informatie van geen belang. Vanzelfsprekend zullen de aangeboden diensten op het informatiedomein in hoge mate afhankelijk zijn van de vorm en de toegankelijkheid van het informatiedomein.

Een informatiedomein zullen we in ons model als volgt grafisch weergeven:



Figuur 1: Een informatiedomein

Relevante vragen over het informatiedomein zijn:

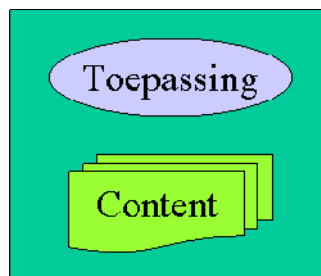
- wie is de eigenaar van de collectie?
- wie is de eigenaar van de content?
- is het informatiedomein duurzaam houdbaar?
- wat is het kwaliteitsniveau van het domein?
- wie is hiervoor verantwoordelijk?

Antwoorden op deze vragen zijn bepalende succesfactoren voor aanbieders van diensten van informatie. Stel een reisbureau koopt haar informatie bij een uitgeverij. Een klant boekt een reis bij het reisbureau op basis van de informatie in brochure. Deze informatie blijkt op de plek van bestemming niet te kloppen. Wie is nu eindverantwoordelijk en gaat de rekening uiteindelijk betalen? Het reisbureau of de uitgever?

Gebruikmakend van een informatiedomein kunnen met behulp van taken verschillende informatiediensten in de lucht gebracht en gehouden worden. Een eenvoudige opzet waarin de aangeboden informatie niet wordt bewerkt vergt al een scala aan taken om de dienst te (blijven) bieden, zowel technisch als inhoudelijk. In een WWW-omgeving kan dan gedacht worden aan de volgende zaken: de informatie moet opgeslagen zijn op een drager, vaak zal er duplicatie plaatsvinden om problemen bij uitval te voorkomen, de drager moet benaderd kunnen worden via een webbrowser en er moet onderhoud plaatsvinden om te voorkomen dat derden met verkeerde data aan de slag gaan.

Een iets ingewikkelder opzet verrijkt de data, bijvoorbeeld door metadata toe te voegen, door data doorzoekbaar te maken, of zelfs door een interface toe te voegen. Ambitieuzer is een opzet waarbij data van derden beschikbaar worden gesteld. Dit vergt altijd selectie, die niet beperkt hoeft te blijven tot een selectie van externe databronnen maar ook binnen de bronnen zelf bepaalde data wel en andere niet beschikbaar stelt. Weer een stap verder is het aanbieden van programmatuur die bewerkingen op data mogelijk maakt. De reisplanner van de NS is een voorbeeld van zo'n informatiedienst. De data komen deels van de NS zelf (het digitale spoorboekje), deels van derden. Bij de reisplanner van de NS zijn de mogelijkheden tot bewerken van de data voorgekookt. Nog een stap verder is het aanbieden van een informatiedienst die verschillende databronnen en programma's beschikbaar stelt, waarbij de gebruiker de mogelijkheid heeft om zelf een dataflow te definiëren, bijvoorbeeld door middel van een grafische interface.

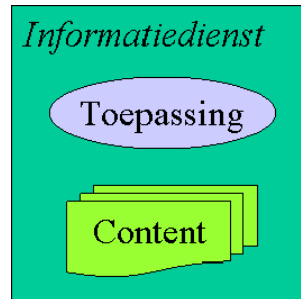
De meeste diensten worden dus gerealiseerd door meerdere taken uit te voeren. Taken zullen in hoofdstuk 6 verder uitgewerkt worden. We groeperen taken om aan te geven dat een pakket van meerdere, met elkaar samenhangende taken een dienst op een informatiedomein vormt. Dit samenhangend pakket van taken op een dienst noemen we een 'toepassing'. Toepassingen op een informatiedomein worden als volgt weergegeven:



Figuur 2: Toepassing op een informatiedomein

4 Informatiediensten

Een informatiedienst kan nu gepresenteerd worden als een toepassing op een informatiedomein, en een toepassing als een pakket van één of meerdere taken. Dit wordt als volgt grafisch weergegeven.

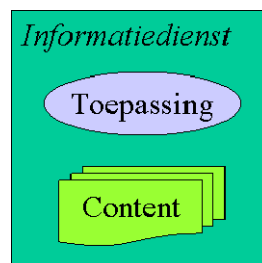


Figuur 3: Een informatiedienst

Voordat we een en ander concretiseren, bekijken we eerst de gelaagdheid van informatiediensten, welke zijn heeft wortels in de informatietechnologie. Deze heeft namelijk drie belangrijke veranderingen teweeggebracht:

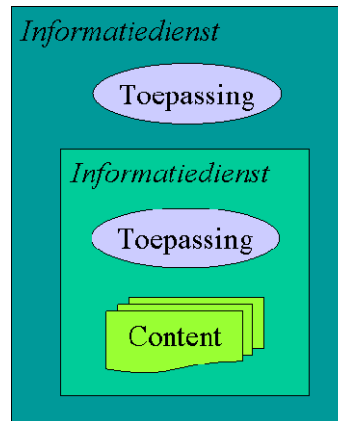
1. Informatie is veelal (ook) digitaal beschikbaar, of kan eenvoudig gedigitaliseerd worden.
2. Toepassingen kunnen door middel van eenvoudige technologie op het internet ingezet worden (bijvoorbeeld via ASP).
3. Het beschikbaar zijn van een 'digitaal platform' voor info-uitwisseling. Dit kan het internet zijn, maar dat hoeft niet. Aldrich noemt dit een digital function platform (Aldrich 1999).

Dit betekent dat er eenvoudig diensten op een bestaande informatiedienst gebouwd kunnen worden. We spreken dan van een gelaagd model. Een actor die een informatiedienst wil maken heeft nu drie opties. Ten eerste kan hij een nieuwe informatiedienst starten door een informatiedomein van een toepassing te voorzien.



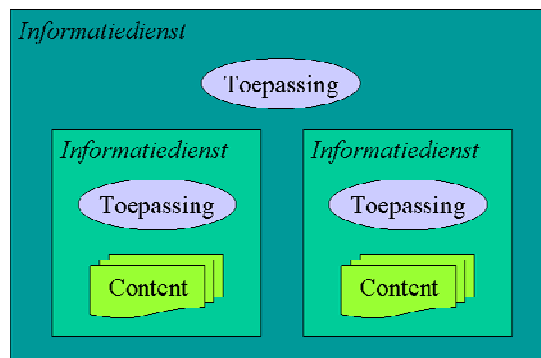
Figuur 4: Een informatiedienst als toepassing op een informatiedomein

Ten tweede kan hij een (aantal) bestaande informatiediensten uitbreiden. Hierin zijn twee varianten te onderkennen. Hij kan een toepassing op de bestaande informatiedienst bouwen en daarmee een nieuwe informatiedienst creëren. We spreken dan van een verticaal uitgebreide (of een verdiepte) informatiedienst.



Figuur 5: Een verdiepte informatiedienst

Of hij kan een nieuwe informatiedienst naast de bestaande informatiedienst op hetzelfde domein of dienst aanbieden. In dit geval spreken we van een verbreding.



Figuur 6: Een verbrede informatiedienst

Een voorbeeld van een verbrede informatiedienst is Google. Google combineert namelijk verschillende informatiediensten en biedt daar nieuwe / eigen taken op aan. Ook de ontwikkelingen in de bioinformatica zijn hier interessant, omdat de beoefenaren van deze discipline geconfronteerd worden met een onvoorstelbare toevloed van data die zonder geavanceerde informatiediensten niet te hanteren valt. Op dit moment worden door bioinformatici de eerste federated information systems opgezet die letterlijk giga-hoeveelheden data uit heterogene bronnen gedistribueerd over de hele wereld vanaf een plek ter beschikking stellen. De volgende stap, het aanbieden van een geïntegreerde

toegang tot data en programmatuur om die data te bewerken wordt nu gezet (Stein, 2003 en Gray & Kemp, 2000).

5 Het raamwerk toegepast in een bedrijfseconomische omgeving

Het gepresenteerde raamwerk kent dus vier entiteiten: informatiedomein, taken, toepassingen en informatiediensten. In dit hoofdstuk zullen we het raamwerk toepassen om bedrijfseconomische consequenties van veranderende informatiespelers te bestuderen. Bedrijfseconomische consequenties zien we in de praktijk in twee aspecten van informatiediensten. Ten eerste zal bepaald moeten worden via welk business model een informatiedienst aangeboden gaat worden. Ten tweede is er de belangrijke vraag hoe een concurrentievoordeel opgebouwd en gehandhaafd kan worden. Dit wordt uiteen gezet in de volgende twee secties.

5.1 Business modellen

Laten we uitgaan van een eigenaar van een toepassing. De vraag rijst dan welk businessmodel hanteerbaar is. Er zijn vier typen te onderscheiden:

	Omschrijving	Voorbeeld
Klassiek	De toepassing vindt plaats op een eigen informatiedomein	www.nrc.nl die via een betaalde toegang informatie uit eigen archieven beschikbaar stelt
Aggregator	De toepassing vindt plaats op een informatiedomein van een andere organisaties	www.nu.nl die informatie beschikbaar maakt afkomstig uit andere bronnen
Full-service	De toepassing vindt plaats op een eigen informatiedienst	www.froogle.com een informatiedienst voor producten gebouwd op een eigen zoekdienst (Google)
Services Syndication	De toepassing vindt plaats op een informatiedienst van een andere organisatie	www.overture.com een handelsysteem voor het verkopen van internetadvertenties gebouwd op en voor internetzoeksystemen

Ieder type kent zijn eigen diversiteit aan businessmodellen en issues omtrent deze businessmodellen.

Klassiek: Voor het opzetten van deze toepassing geldt de klassieke investeringsvraag: wegen de investeringen op tegen de verwachte opbrengsten op termijn? Een belangrijke tweede vraag is of de nieuwe informatiedienst geen kannibalisatie van de bestaande producten en diensten oplevert.

Aggregator: Hiervoor geldt de afhankelijkheidsvraag. Hoe afhankelijk is de aanbieder van de toepassing van het onderliggende informatiedomein? Is de content op een andere manier te vergaren? Kan de eigenaar van het informatiedomein op eenvoudige wijze de toepassing zelf op gaan zetten? Dit gebeurde bijvoorbeeld in www.rechtspraak.nl waarbij eerst de rechtsuitspraken werden verspreid door de SDU; op zeker moment stelde Justitie de informatie zelf beschikbaar via WWW.

De vraag rijst dus wat de toegevoegde waarde van de toepassing voor de eigenaar van het informatiedomein is. De meeste informatiediensten die een applicatie hebben op een vreemd informatiedomein doen dit op publiek toegankelijke en gratis domeinen. Dit is dus een enorme vechtmarkt omdat deze markt een lage drempel voor toetreders kent.

Full-service: Dit is vergelijkbaar met het klassieke type maar hier geldt dat er een informatiedienst bestaat waarop een additionele toepassing (dus dienst) wordt gelegd. Hier moeten de make-or-buy en outsourcingvraagstukken worden beantwoord. In de internethype ging menige informatiedienst zijn eigen zoekstelsel ontwikkelen, om deze als extra dienst aan te bieden op de eigenlijke dienstverlening, bijvoorbeeld in de reiswereld waar men klanten in staat wilde stellen op verschillende manieren te zoeken naar een reisbestemming. Vaak wordt de keuze gemaakt om de core-businessfunctionaliteit of de functionaliteit met het meeste klantcontact in eigendom te nemen en alle andere toepassingen uit te besteden.

Service Syndicator: Dit type kent soortgelijke vragen als het aggregator maar hier is veel meer van belang de vraag naar het customer-ownership. Wie bezit de klant? Een toepassing invoeren tussen de klant en de bestaande informatiedienst (bijvoorbeeld de eerder gememoreerde www.kranten.com) leidt tot felle concurrentiestrijd. Lukt het immers om op deze wijze de markt te penetreren dan bewerkstelligt de nieuwe toepassing een belangrijk concurrentievoordeel.

Men kan hier de ontwikkelingen van de reisbranche weer in herkennen. Bijna iedere reisorganisatie heeft haar eigen website. Maar door het bestaan van de vele service syndicators zoals bijvoorbeeld internetzoekmachines gaat de klant niet meer rechtstreeks naar het internetreisbureau. De zoekmachine stuurt (al dan niet bewust) de klant naar bepaalde reisbureaus. Dergelijke ontwikkelingen hebben we bijvoorbeeld ook waar kunnen nemen in de verkoop van huizen (www.funda.nl).

5.2 Concurrentievoordeel voor informatiedienst

Voor de eigenaren van een (commerciële) informatiedienst is het van belang om een voordeel ten opzichte van de concurrentie op te bouwen en te behouden. Hiervoor zijn vanuit ons raamwerk twee basismanieren aan te geven.

Ten eerste kan de informatiedienst op grond van zijn informatiedomein een concurrentievoordeel boeken. Dit hangt af van de mate waarin het informatiedomein uniek is. Een monopolie op het informatiedomein vormt voor de informatiedienst een goede basis voor een concurrentievoordeel. Dit zal echter niet vaak voorkomen, omdat steeds meer informatie door meerdere aanbieders voor publiek toegankelijk wordt gemaakt. In de slimme/unique combinatie van een aantal bronnen liggen ook mogelijkheden voor het realiseren van een concurrentievoordeel. Maar ook hier geldt dat veel informatie gedekt worden door verschillende bronnen. In de regel is het dus riskant om het informatiedomein te zien als de basis voor (duurzaam) concurrentievoordeel.

Veel informatiediensten zullen daarom op basis van hun toepassingen een concurrentievoordeel moeten realiseren. Ook de mogelijkheden hiervan om blijvend voordeel te bieden zijn echter beperkt. Veel toepassingen zijn namelijk dupliceerbaar; succesvolle toepassingen worden na verloop van tijd gekopieerd door de concurrenten. Alleen innovatieve informatiediensten zullen de concurrentie steeds een stap voor kunnen blijven. Innovatie is voor informatiediensten daarom een noodzakelijke voorwaarde voor een duurzaam concurrentievoordeel.

Daarnaast is het voor het opbouwen van een concurrentievoordeel noodzakelijk dat de informatiedienst over goede klantcontacten beschikt. Alleen dan kan de informatiedienst namelijk gericht inspringen op de informatiebehoeften van de klant. De vraag 'wat wil de moderne informatiek klant' is daarom voor informatiedienst een essentiële.

In het digitale tijdperk gaan de behoeften van klanten verder dan de drie traditionele aspecten van prijs (verkoopprijs, prijsstructuur,), kwaliteit (tijdig, juist, volledig, ...) en merk (betrouwbaarheid, imago, naamsbekendheid, ...). In het raamwerk van Aldrich worden ook content (extra informatie(diensten) bij een product of dienst) en tijd (snelheid zoals zoektijd en publicatietijd), en gemak (bijvoorbeeld door gebruiksvriendelijkheid) als valuedrivers genoemd (Aldrich 1999). Omdat de belevingseconomie steeds belangrijker wordt, is ook de value driver 'fun' van belang. Informatiediensten zijn namelijk ook emotie. Innovatieve services hebben een grotere 'wow factor' en trekken daarmee extra klanten. Deze observatie, die overigens niet nieuw is, kan door informatiediensten gebruikt worden bij de innovatie van de dienst.

Voor de eigenaar van een informatiedienst is het namelijk van belang om een continu proces van klantgerichte innovatie op gang te brengen. Daarbij wordt telkens beoordeeld hoe de informatiedienst op bovengenoemde value drivers verbeterd kan worden. Alleen dan kan een informatiedienst langere tijd succesvol zijn.

6 ICT oplossingen voor informatiediensten

De beschreven informatiediensten kunnen worden aangeboden door middel van ICT-oplossingen waarvan er vele “off the shelf” verkrijgbaar zijn, terwijl andere zich in de ontwikkelfase bevinden of onderwerp zijn van onderzoek. Waar het aan ontbreekt is een samenhangende architectuur die alle componenten op een zinvolle wijze laat samenwerken. Zo'n architectuur is zelf weer het resultaat van een visie op de implementatie van webdiensten. We zullen hier geen architectuur schetsen omdat we er van overtuigd zijn dat er vele visies en vele business modellen zijn die ieder een op maat gesneden architectuur vragen. We zullen de betrokken processen kort karakteriseren en daarbij ook aandachtspunten identificeren die bij de keuze van architectuur een rol spelen.

Eerst willen we echter kort ingaan op een kwestie die in het verleden tot discussie heeft geleid: data warehouse (DW) of federated information system (FIS). Kort gezegd is een DW een database systeem dat in huis wordt onderhouden. Een FIS daarentegen laat de data waar ze zijn en haalt ze selectief op op het moment dat ze gevraagd worden. Als een aggregator kiest voor een DW, zal hij van tijd tot tijd informatie van derden op moeten halen en integreren in het eigen warehouse voordat de data aan de eigen klanten kunnen worden aangeboden. Kiest de aggregator voor een FIS, dan wordt informatie van derden pas opgehaald op het moment dat de klant van de aggregator erom vraagt. In een opzet waarin we abstraheren van locatie, timing van processen, en allocatie van actoren aan processen zijn beide architecturen vrijwel identiek. In een omgeving waarin breedband-internet beschouwd wordt als een normale voorziening mag locatie geen rol meer spelen. Dan blijven over de timing van processen en de allocatie van actoren aan processen. Zulke beslissingen zijn lokaal en kunnen niet in algemene zin, zonder specifieke omstandigheden er in te betrekken, besproken worden. Dit betekent dat de architectuur ontworpen zal moeten worden als onderdeel van de organisatie en haar strategische doelstellingen (zie onder meer Huibers, Kooper, Manschot en Shahim, 2004).

Keren we nu naar de processen die door een bedrijf uitgevoerd moeten worden om informatiediensten te bieden. We gaan uit van een web-omgeving, de lezer denke er de analoge processen voor een papieren uitgeverij zelf bij.

- **Archiveren.** Data opslaan in een leesbaar formaat op een duurzame drager. Dit is minder makkelijk dan het klinkt omdat dragers verouderen (ook tapes en CD's) en operating systems en software voortdurend veranderen. De Koninklijke Bibliotheek heeft samen met Elsevier en IBM een archief opgezet met de garantie dat de inhoud ook na 50 jaar nog leesbaar is. 50 jaar is een enorm lange periode: wie van ons kan zonder meer een document in WordPerfect 3.0 lezen? Met andere woorden, zelfs als we deselectie buiten beschouwing laten is archiveren in een digitale omgeving een continu proces. De verschillen tussen DW en FIS spelen hier een rol omdat de

verantwoordelijkheid voor het archief in de twee oplossingen bij verschillende actoren wordt gelegd.

- **Acquireren.** Als de dienst is gebaseerd op data van derden zal er een actief acquisitiebeleid gevoerd moeten worden. Daarbij speelt ook een rol welke relatie er bestaat met de leverancier van de data. Moet er worden betaald, zijn alle data toegankelijk, mag de aanbieder kiezen in welke vorm data van derden worden aangeboden. Ict kan een rol spelen in de vorm van dedicated spiders die het web doorzoeken op sites die interessante data of diensten bieden en om die reden geacquireerd zouden kunnen worden.
- **Kwaliteitscontrole.** Data kunnen onjuist of onvolledig zijn, en daarnaast verouderen data zelf ook, bijvoorbeeld door nieuwe inzichten. In de wetenschappelijke wereld verliezen data repositories die niet onderhouden worden in deze zin hun waarde binnen twee à drie jaar. In de bioinformatica wordt tegenwoordig vaak gekozen voor een FIS om de eenvoudige reden dat de primaire data, vergaard in empirisch onderzoek, voortdurend onderhoud vergen. Zulk onderhoud kan het best worden uitgevoerd door de instituten die de data hebben vergaard, omdat zij er het meest van weten. Ook hier verschillen DW en FIS in de zin, dat de verantwoordelijkheden op verschillende wijze bij de deelnemende actoren worden gelegd.
- **Beschrijven.** Data moeten worden beschreven, bijvoorbeeld door middel van metadata, om er in te kunnen zoeken. Systemen die automatisch metadata toevoegen op basis van een analyse van de data bevinden zich nog in het laboratorium. Lerende systemen, die toevoegen van metadata leren op basis van een corpus inclusief handmatig toegevoegde metadata, zijn een alternatief maar ook zulke systemen zijn het stadium van onderzoeksprototype nog niet ontgroeid. Hier ligt een vruchtbaar onderzoeksterrein dat bovendien vrij snel praktisch nut oplevert.
- **Distribueren.** Toegankelijk maken van de data door ze op te slaan op een webserver of door een gewaarborgde toegang tot webserver van anderen te realiseren. Hierbij hoort ook het kostenaspect: zijn de data vrij toegankelijk of moet er betaald worden? Als er betaald moet worden, welke vorm wordt dan gekozen: abonnement, *pay-per-view*, of een combinatie van beide? Iedere vorm vereist specifieke ICT-ondersteuning, al was het maar omdat de verschillende vormen gevolgen hebben voor de administratieve lasten van de betrokken partijen. De contracten die nu gesloten zijn tussen wetenschappelijke uitgeverijen en universiteiten zijn gebaseerd op abonnementen. Voor de uitgeverijen is dat aantrekkelijk omdat deze vorm een deel van de administratieve last naar de klant verschuift. Omdat klantenwensen steeds meer gedifferentieerd zullen worden, verwachten wij dat het abonnement zijn langste tijd gehad heeft. *Pay-per-view*, al dan niet in combinatie met een abonnement (vergelijkbaar met een belbundel bij een telecom-aanbieder), heeft wat ons betreft de toekomst. De administratieve last verschuift weer naar de uitgever, die daarvoor gespecialiseerde software in huis zal moeten hebben.

- **Rechtenbeheer.** Dit omvat registratie van de herkomst van data en regelen van toegangsrechten. Er zijn DRM (*Digital Rights Management*) systemen op de markt die de ICT-aspecten van deze processen ondersteunen.
- **Converteren.** Of er nu een DW of een FIS wordt gebouwd, converteren van data naar een ander formaat is vrijwel onontkoombaar. Een informatiedienst zal er niet aan ontkomen om over een vrij uitgebreid park van *wrappers* te beschikken. Wrappers vergen voortdurend onderhoud omdat formaten voortdurend veranderen.
- **Combineren.** De toegevoegde waarde van een aggregator en, meer nog, van *service syndication* zit hem in de combinatie van informatie tot een nieuw en zinvol geheel. Ook hier zijn wrappers noodzakelijk, maar het combinatieproces is primair gericht op het samenstellen van zinvolle combinaties.
- **Inhoudelijk bewerken van data.** De kern van *service syndication* is naar onze mening de mogelijkheid om data te bewerken tot nieuwe data, bijvoorbeeld door data afkomstig uit verschillende bronnen te aggregeren tot een nieuwe dataset die conclusies toelaat die op grond van de oorspronkelijke data niet getrokken kunnen worden. De analogie is hier met de medische wereld. Primaire informatie, bijvoorbeeld over neveneffecten van geneesmiddelen, is vaak verkregen met relatief kleine populaties en daarom is de betrouwbaarheid niet zo groot als we willen. In een meta-analyse worden vele van zulke kleinschalige onderzoeken gecombineerd tot één onderzoek. De software die zulk werk ondersteunt is overigens, alweer, voornamelijk in het laboratorium te vinden. Het is echter duidelijk dat het gemak waarmee informatie beschikbaar kan worden gesteld zal leiden tot actief onderzoek naar software die meta-analyses ondersteunt.
- **Interactie.** Dienstverleners bieden informatie aan omdat ze menen dat er een markt voor is. Gebruikers van de diensten zullen die dienst moeten activeren en naar hun wensen richten door middel van een dialoog met het systeem, meestal via een webinterface. Het ontwerpen van zulke interacties is van vitaal belang voor het slagen van een informatiedienst. Daarvoor moeten vragen worden beantwoord als: wat weet onze gebruiker van de dienst, hoe zal hij of zij willen navigeren, waartoe wordt de dienst gebruikt, hoe wil de gebruiker de informatie gepresenteerd krijgen. Omdat iedere niet-triviale dienst de klant een omvangrijk en vaak ook gedifferentieerd pakket zal aanbieden, is Information Retrieval (IR) een sleuteltechniek voor iedere dienst aanbieder. Wetenschappelijke uitgeverijen zijn voortdurend bezig om de zoekmogelijkheden op hun eigen sites te verfijnen (voorlopig alleen voor abonnees, overigens) terwijl Elsevier daarnaast, en voordat Google Scholar er was, de speciale zoekdienst Scirus heeft opgezet. Ook archieven van onderzoeksgroepen, faculteiten, universiteiten zullen alleen interessant zijn voor niet-ingewijden wanneer een geavanceerde IR-functionaliteit wordt geboden. Een stap verder en de gebruiker is zelf in staat om diensten te definiëren die bestaan uit combinaties van meer eenvoudige diensten die via de informatieleverancier toegankelijk zijn. Zulke omgevingen zijn voorwerp van onderzoek in het nationale Bsic-project VL-E (*Virtual*

Laboratory Environment), waarin data en programmatuur in een GRID-omgeving beschikbaar zijn.

De wijze waarop dergelijke informatiediensten kunnen worden geïmplementeerd vergt meer dan slechts een technologische aanpak. De veel gehanteerde gebalanceerde aanpak met de aandachtsgebieden People, Process en Technology dient op zijn minst uitgebreid te worden met specifieke aandacht voor Content.

7 Conclusies en verder onderzoek

In dit artikel werd een kader gepresenteerd waarin de elementen informatiedomein, taken, toepassingen en informatiediensten kunnen worden geordend en besproken. De bruikbaarheid werd aangetoond door actuele bedrijfseconomische discussies van informatiespelers aan de hand van het raamwerk te bespreken.

In verder onderzoek willen wij het raamwerk uitwerken waarbij wij ons vooral willen richten op de diverse ICT-toepassingen passend binnen het raamwerk. Wij zijn er namelijk van overtuigd dat vele nieuwe en bestaande ICT-toepassingen zoals Information Retrieval en Question/Answering-tools een nog belangrijkere rol gaan spelen in de informatie-economie. Met name de bedrijfseconomische beslissingen van de inzet van dergelijke toepassingen behoeft verder onderzoek.

8 Auteurs

Prof. dr. Theo Huibers MMC is directeur bij KPMG Business Advisory Services en tevens hoogleraar Information Retrieval aan de Universiteit Twente. Zijn specialisaties zijn het (her)inrichten van organisaties en processen op basis van een strategische of tactische heroriëntatie. Specifiek ligt zijn ervaring op de gebieden management & organisatie, informatiestrategie en -management, sourcingsvraagstukken, financieel management en verandervraagstukken. Email: Huibers.Theo@kpmg.nl

Dr. Paul van der Vet is onderzoeker bij de afdeling Informatica van de Universiteit Twente op het gebied van digitale informatievoorziening. Hij richt zich onder andere op text mining, information retrieval, ontologieën, gebruikersinterfaces voor complexe natuurwetenschappelijke informatie en business modellen voor webuitgevers. Email: p.e.vandervet@utwente.nl

Dr. Bernd Wondergem MMO is adviseur bij KPMG Business Advisory Services. Zijn specialisaties zijn informatiemanagement, procesmanagement, risicomanagement en performance management. Email: Wondergem.Bernd@kpmg.nl

9 Referenties

- Aldrich, 1999 Mastering the digital marketplace. D.F. Aldrich, John Wiley & Sons, 1999.
- Gray & Kemp, 2000 Federated database technology for data integration: lessons from bioinformatics. P.M.D. Gray en G.J.L. Kemp. In, Electronic collaboration in science, editors S.H. Koslow en M.F. Huerta, pp 45–72, Mahwah NJ: Lawrence Erlbaum, 2000.
- Hexagoon 2, 1990 De schijn van zekerheid, een pleidooi voor herbezinning op auteurscont(r)acten. Wolters Kluwer, 1990.
- Huibers, 2004 Nieuwe spelers in de informatiemaatschappij. T.W.C. Huibers. In 'de kracht van vernieuwing, visies op ICT', editor Abbas Shahim, Academic Services, 2004.
- Theo Huibers, Michiel Kooper, Kathelijn Manschot en Abbas Shahim, 2004 Flexibiliteit sleutelfactor in alignmentproces. In Informatie, februari 2004.
- Huibers en Timmers, 2004 De Uitgever aan het woord. T.W.C. Huibers en C. Timmers, KPMG, 2004.
- Kist, 1996 Bibliodynamica –Slaag en faalkansen bij innovatie van informerende systemen, in het bijzonder in het uitgeversbedrijf–. J. Kist, Otto Cramwinkel Uitgever, 1996.
- Negroponte, 1995 Being Digital. N. Negroponte, Hodder & Stoughton, London, 1995.
- Roosendaal, 2002 Informatie en wetenschap: een keten van waarde. H.E. Roosendaal. Oratie, Universiteit Twente, mei 2002.
- Stein, 2003 Creating a bioinformatics nation. L. Stein. Nature 417, pp 119–120, 2003.