

# ***Integratie van text mining-technieken in de ontwikkeling van een e-learningcursus***

**Sven EVENS**

promotor :  
Prof. Jeanne SCHREURS

## INHOUDSOPGAVE

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| INHOUDSOPGAVE.....                                                    | 2  |
| SAMENVATTING .....                                                    | 5  |
| WOORD VOORAF .....                                                    | 8  |
| HOOFDSTUK 1 : INLEIDING TOT HET ONDERZOEK .....                       | 9  |
| 1.1 Inleiding .....                                                   | 9  |
| 1.2 Probleemstelling .....                                            | 9  |
| 1.3 Methode en aanpak .....                                           | 10 |
| 1.4 Overzicht van de eindverhandeling .....                           | 10 |
| HOOFDSTUK 2 : EVOLUTIE NAAR NIEUWE ONDERWIJSCONCEPTEN IN E-LEARNING.. | 12 |
| 2.1 Inleiding .....                                                   | 12 |
| 2.2 Het 'pre-e-learning'-tijdperk .....                               | 12 |
| 2.2.1. Nieuwe media in afstandsleren .....                            | 12 |
| 2.2.2. Computer-based training .....                                  | 13 |
| 2.3 E-learning .....                                                  | 15 |
| 2.3.1. De oorzaken.....                                               | 15 |
| 2.3.1.1 Maatschappelijke veranderingen .....                          | 15 |
| 2.3.1.2 Internethype .....                                            | 16 |
| 2.3.2. Ontwikkelingsstadia van e-learning .....                       | 17 |
| 2.3.2.1 Fase van overdreven enthousiasme.....                         | 19 |
| 2.3.2.2 Fase van desillusie .....                                     | 20 |
| 2.3.2.3 Fase van geleidelijke verbetering .....                       | 23 |
| 2.3.3. Definitie e-learning .....                                     | 24 |
| 2.3.4. Voor- en nadelen van e-learning .....                          | 26 |
| 2.3.4.1 Voordelen .....                                               | 26 |
| 2.3.4.2 Nadelen .....                                                 | 26 |
| 2.4 Trends in e-learning .....                                        | 27 |
| 2.4.1. Algemene trends .....                                          | 27 |
| 2.4.2. Web 2.0 .....                                                  | 28 |
| 2.4.2.1 Sociale software .....                                        | 28 |
| 2.4.2.2 Folksonomieën.....                                            | 29 |
| HOOFDSTUK 3 : LEEROBJECTEN IN EEN E-LEARNINGCURSUS .....              | 30 |
| 3.1 Inleiding .....                                                   | 30 |
| 3.2 Leerobjecten .....                                                | 30 |
| 3.2.1. LEGO of atomen? .....                                          | 30 |
| 3.2.2. Conceptueel model voor leerobjecten .....                      | 31 |
| 3.2.3. Definitie van leerobjecten .....                               | 33 |
| 3.2.4. Leerobject als psychologisch hulpmiddel in het leerproces..... | 35 |
| 3.3 Ontwikkeling van een e-learningcursus .....                       | 36 |
| 3.3.1. Decompositie .....                                             | 37 |
| 3.3.2. Scenario-model .....                                           | 39 |
| HOOFDSTUK 4 : TEXT MINING EN NATURAL LANGUAGE PROCESSING .....        | 42 |
| 4.1 Inleiding .....                                                   | 42 |
| 4.2 Text mining versus Natural Language Processing .....              | 42 |
| 4.2.1. Natural Language Processing.....                               | 43 |

|                                         |                                                                     |    |
|-----------------------------------------|---------------------------------------------------------------------|----|
| 4.2.2.                                  | Text mining.....                                                    | 44 |
| 4.2.3.                                  | Overlapping tussen text mining en natural language processing ..... | 44 |
| 4.3                                     | Text mining-kader.....                                              | 46 |
| 4.3.1.                                  | Information retrieval .....                                         | 46 |
| 4.3.2.                                  | Information extraction.....                                         | 47 |
| 4.3.2.1                                 | Named Entity Recognition .....                                      | 48 |
| 4.3.2.2                                 | Coreference resolution.....                                         | 49 |
| 4.3.2.3                                 | Template Element construction .....                                 | 50 |
| 4.3.2.4                                 | Template Relation construction .....                                | 50 |
| 4.3.2.5                                 | Scenario Template production .....                                  | 51 |
| 4.4                                     | i.Know .....                                                        | 51 |
| 4.4.1.                                  | Uitgangspunt.....                                                   | 51 |
| 4.4.2.                                  | Technologie .....                                                   | 52 |
| 4.4.2.1                                 | De multidimensionale kennisvoorstellingsmodule .....                | 52 |
| 4.4.2.2                                 | De component voor de gebruikersfunctionaliteit.....                 | 55 |
| 4.4.2.3                                 | De interface voor de kennisuitwisseling.....                        | 58 |
| HOOFDSTUK 5 :                           | GEVALSTUDIE .....                                                   | 59 |
| 5.1                                     | Doel van de gevalstudie .....                                       | 59 |
| 5.2                                     | Methode 1 .....                                                     | 59 |
| 5.2.1.                                  | Eerste categorisatie .....                                          | 60 |
| 5.2.2.                                  | Ontwikkeling taxonomie en aanpassing ontologie i.Know .....         | 62 |
| 5.2.3.                                  | Verfijning categorisatie.....                                       | 63 |
| 5.2.4.                                  | Output .....                                                        | 64 |
| 5.2.5.                                  | Topics .....                                                        | 66 |
| 5.2.6.                                  | Resultaten topic-herkenning.....                                    | 67 |
| 5.3                                     | Methode 2 .....                                                     | 68 |
| 5.3.1.                                  | Uitgangspunt.....                                                   | 69 |
| 5.3.2.                                  | Concepten .....                                                     | 70 |
| 5.3.3.                                  | Relaties .....                                                      | 71 |
| 5.3.4.                                  | CRC's .....                                                         | 72 |
| 5.3.5.                                  | Topics .....                                                        | 72 |
| 5.3.6.                                  | Resultaten topic-herkenning.....                                    | 73 |
| 5.3.7.                                  | Tekst 2.....                                                        | 74 |
| 5.3.8.                                  | Controleteksten topic-herkenning .....                              | 77 |
| 5.3.9.                                  | Samenvattingen .....                                                | 83 |
| 5.3.10.                                 | Sleutelwoorden.....                                                 | 84 |
| 5.4                                     | Conclusie gevalstudie .....                                         | 85 |
| HOOFDSTUK 6 :                           | CONCLUSIES.....                                                     | 89 |
| 6.1                                     | Algemene conclusies.....                                            | 89 |
| 6.2                                     | Suggesties voor verder onderzoek.....                               | 91 |
| 6.3                                     | Persoonlijke conclusies .....                                       | 92 |
| LIJST VAN DE GERAADPLEEGDE WERKEN ..... |                                                                     | 93 |

## BIJLAGEN

|              |                                                                     |
|--------------|---------------------------------------------------------------------|
| Bijlage I    | Taxonomie                                                           |
| Bijlage II   | Tekst 1: 'Outsourcing Collaborative Partnerships'                   |
| Bijlage III  | Schema tekst 1: 'Outsourcing Collaborative Partnerships'            |
| Bijlage IV   | Tekst 2: 'Outsourcing Information Systems Functions'                |
| Bijlage V    | Schema tekst 2: 'Outsourcing Information Systems Functions'         |
| Bijlage VI   | Tekst 3: 'Outsourcing as an Economic Strategy'                      |
| Bijlage VII  | Tekst 4: 'Stategic dimensions of IT outsourcing'                    |
| Bijlage VIII | Samenvattingen Tekst 1: 'Outsourcing Collaborative Partnerships'    |
| Bijlage IX   | Samenvattingen Tekst 2: 'Outsourcing Information Systems Functions' |
| Bijlage X    | Samenvattingen tekst 3: 'Outsourcing as an Economic Strategy'       |
| Bijlage XI   | Samenvattingen tekst 4: 'Stategic dimensions of IT outsourcing'     |

## SAMENVATTING

De kennismaatschappij van vandaag wordt gekenmerkt door een dagelijks groeiende hoeveelheid aan informatie. Bovendien is met het internet de beschikbaarheid van deze informatie enorm uitgebreid. Met de steeds sneller toenemende informatie daalt de beschikbare tijd, waarin mensen deze kunnen verwerken. Men heeft dus nood aan technologische hulpmiddelen en dit zowel in bedrijven als in onderwijsinstellingen, waar men tegenwoordig meer en meer gebruik maakt van één of andere vorm van e-learning.

De ontwikkeling van een e-learningcursus is voor een docent een uitgebreid proces, waarin nog een aantal tijdrovende manuele stappen voorkomen. Misschien kunnen de steeds verder evoluerende technieken uit *text mining* en gerelateerde domeinen hierbij een hulpmiddel zijn? Dit is de vraag die centraal staat in deze eindverhandeling. Concreet bestaat het opzet eruit om te kijken in welke mate en op welke manier bepaalde aspecten van de ontwikkeling van een e-learningcursus kunnen worden geautomatiseerd met behulp van technieken uit het text mining-gebied.

Het eerste luik van deze eindverhandeling neemt het e-learning-gebeuren onder de loep. De snelle opkomst ervan is toe te schrijven aan een aantal maatschappelijke veranderingen in combinatie met de internethype rond 1995, die als technologische hefboom fungeerde. De hoge vlucht van e-learning stakte nadat men vaststelde dat de hoge verwachtingen niet ingelost konden worden. Door vele opduikende praktische obstakels, gevolgd door financiële problemen, verzeilde de hele sector zelfs in een ware crisis. Sinds enkele jaren is er opnieuw sprake van een hernieuwd en ditmaal gezond optimisme en kent e-learning dan ook een geleidelijke groei. Recent zorgen trends als *social computing*, de toenemende mobiele toepassingen, de stijgende mate van personalisatie en meer aandacht voor *collaboration* in het onderwijs voor een nieuwe wind binnen e-learning.

De nieuwe onderwijsconcepten in e-learning worden mogelijk gemaakt door de zogenaamde *learning objects* of leerobjecten. Dit zijn relatief kleine, maar zelfstandige digitale componenten die binnen het leerproces als een herbruikbaar, psychologisch hulpmiddel dienst doen. De leerobjecten zijn belangrijk, aangezien ze in de creatie van een e-learningcursus een centrale rol aannemen.

Twee grote fasen zijn te onderscheiden bij het ontwikkelingsproces van een e-learningcursus. In een eerste fase wordt de leerinhoud gedecomposeerd. Hierbij wordt de leerstof opgesplitst in kleine afzonderlijke elementen, die de naam 'atomische leerobjecten' meekregen. In de tweede fase worden deze atomische leerobjecten volgens een uitgekozen scenario opnieuw samengevoegd tot een volwaardige e-learningcursus. De eerste stap hierbij is de creatie van informatieblokken, die worden opgebouwd uit een inhoudelijke component en enkele presentatie-elementen. De informatieblokken worden vervolgens gecombineerd tot een leerobject. Een volledige cursus

tenslotte bestaat uit een pakket van lessen die op hun beurt een verzameling zijn van één of meerdere leerobjecten.

Het tweede luik van de eindverhandeling maakt een verkenning binnen de domeinen van *text mining* en *natural language processing*. Beide gebieden trachten kennis te destilleren uit de berg aan gegevens, die voornamelijk uit schriftelijke documenten bestaat. De technologieën van text mining vinden hun oorsprong bij data mining en de methoden van natural language processing kennen een taalkundige invalshoek, maar beide domeinen overlappen elkaar steeds meer. Na het schetsen van een kader voor text mining, waarin voornamelijk *information extraction* uitgediept wordt, komt de door het bedrijf i.Know ontwikkelde technologie uitvoerig aan bod. De relationele semantische benadering van i.Know is een voorbeeld waarbij technieken uit text mining en natural language processing gecombineerd worden. De kern van hun technologie is het gebruik van concept-relatie-concept-structuren, die het mogelijk maken om de gebruikte concepten in een tekst niet los, maar in hun context te zien. Hierdoor kan een tekst op een meer correcte manier geïnterpreteerd worden. De functionaliteit, die de software van i.Know op basis van deze technologie bezit, wordt vervolgens aan de hand van de kennislevenscyclus besproken.

In een praktijkgeval is een poging ondernomen om de technologie van i.Know te integreren in het decompositieproces van de ontwikkeling van een e-learningcursus. Eerst is aan de hand van een tweetal teksten rond *outsourcing* het kennisdomein van het systeem uitgebreid. Vervolgens is het systeem geoptimaliseerd om aan paragrafen het juiste topic (binnen outsourcing) toe te kennen. Nadien zijn twee extra teksten over outsourcing ter controle geanalyseerd. Bijkomend zijn er samenvattingen gemaakt en is er een lijst van sleutelwoorden opgesteld. De resultaten zorgden in eerste instantie voor een teleurstelling, maar bleken na een wijziging in de aanpak toch hoopgevend.

Algemeen kan men besluiten dat e-learning aan het evolueren is naar vormen waarbij de 'lerende' op een centrale plaats komt. Daarnaast stelt men vast dat de door gebruikers samen gecreëerde microcontent een fundamenteel kenmerk is van de nieuwste e-learning-toepassingen. Deze microcontent kan op diverse manieren worden herbruikt. Leerobjecten zijn een gelijkaardige technologische oplossing om herbruikbaarheid van elementen binnen de ontwikkeling van een e-learningcursus te realiseren. Zoals reeds gesteld convergeren de domeinen text mining en natural language processing steeds verder naar elkaar. De mogelijkheden nemen steeds toe en worden continu verbeterd; hiervan zijn *topic detection* en de automatische aanmaak van samenvattingen maar twee voorbeelden. De integratie van deze twee voorbeelden en andere text mining-technieken in het decompositieproces van de ontwikkeling in een e-learningcursus biedt zeker mooie perspectieven. Op dit ogenblik kunnen deze methoden enkel een ondersteunende rol aannemen, maar met bijkomende verfijningen en verdere ontwikkelingen moet de graad van

automatisatie van het decompositieproces van een e-learningcursus zeker uitgebreid kunnen worden.

## **WOORD VOORAF**

Deze eindverhandeling vormt het sluitstuk van mijn opleiding tot Handelsingenieur in de Beleidsinformatica. Van deze mooie gelegenheid wil ik gebruik maken om enkele personen te bedanken, zonder hun steun had ik mijn opleiding en dit werkstuk in het bijzonder niet kunnen beëindigen.

In de eerste plaats wil ik mijn promotor bedanken voor de goede opvolging van deze eindverhandeling en om altijd met raad en daad klaar te staan.

Daarnaast wil zeker een speciaal woordje van dank richten naar alle de medewerkers van i.Know. In het bijzonder wil ik Dirk van Hyfte danken om mij de mogelijkheid te bieden voor de gevalstudie met i.Know samen te werken en daarbij uitgebreid te hebben kunnen kennis maken het de technologieën van i.Know. Ook Saskia Debergh mag hier zeker niet ontbreken. Zij heeft prachtig werk geleverd en stond steeds klaar om mijn vragen te beantwoorden.

Als laatste gaat mijn dank uit naar mijn ouders om mij enerzijds de kans gegeven te hebben om deze studies aan te vatten en te voltooien. Anderszijds wil ik hen ook bedanken voor de steun die ze mij gedurende de jaren van mijn opleiding zijn blijven geven.

Sven Evens

*Zonhoven, augustus 2007*



## **HOOFDSTUK 1 : Inleiding tot het onderzoek**

### **1.1 Inleiding**

Het onderwijs is de afgelopen jaren op diverse vlakken veranderd. Zo wordt voor de wijze waarop de leerstof aan de studenten aangeboden wordt, steeds frequenter en intensiever gebruik gemaakt van één of andere vorm van e-learning. De ontwikkeling van een degelijke e-learningcursus is voor de docent een uitgebreid proces. Ruwweg kan men het ontwikkelingsproces opdelen in twee stappen. In eerste instantie dienen één of meerdere bronteksten geanalyseerd te worden, waarbij de essentiële elementen eruit gefilterd worden en vervolgens in een gepaste structuur gegoten worden. De volgende stap dient dan om dit alles in een mooi kleedje te presenteren aan de gebruikers van de e-learningcursus.

De eerste van die hierboven beschreven stappen, dient volledig manueel uitgevoerd te worden. Het is dan ook overbodig om nog te stellen dat dit een tijdrovend karwei is. Gezien de vele mogelijkheden op technologisch vlak zou het interessant zijn om dit proces te automatiseren.

*Text mining* en gerelateerde technologieën kennen ook een steeds verdere opmars en worden in steeds meer toepassingen gebruikt. Misschien kunnen de technologieën in dit domein een oplossing bieden voor het arbeidsintensieve werk bij de ontwikkeling van een e-learningcursus.

### **1.2 Probleemstelling**

In deze eindverhandeling is het dan ook de bedoeling om te kijken, in welke mate en op welke manier bepaalde aspecten van de ontwikkeling van een e-learningcursus geautomatiseerd kunnen worden met behulp van technieken uit het text mining-gebied. De centrale onderzoeksvraag luidt dan ook als volgt:

*"Op welke manier kunnen text mining-technieken geïntegreerd worden bij de ontwikkeling van een e-learningcursus?"*

Deze centrale onderzoeksvraag kan verder worden opgesplitst in een aantal deelvragen:

- Wat houdt e-learning precies in? Welke evoluties heeft e-learning reeds doorgemaakt en welke trends tekenen zich momenteel af?
- Welke stappen dienen er te worden gezet om tot een e-learningcursus te komen? Wat zijn leerobjecten en welke rol speelt zij bij de ontwikkeling van een e-learningcursus?
- Welke technologieën en mogelijkheden groeien er binnen de schot van *text mining* en *natural language processing*?
- Welke zijn de mogelijke toepassingen van text mining-technologieën binnen e-learning en hoe kunnen deze worden gerealiseerd?

### **1.3 Methode en aanpak**

Het uitvoeren van een literatuurstudie vormt een eerste luik van deze eindverhandeling. Aan de hand van de literatuur wordt er getracht een antwoord te formuleren op de centrale onderzoeksvraag en de bijbehorende deelvragen. Dit onderzoek heeft zich toegespitst op een drie domeinen, namelijk:

- e-learning;
- leerobjecten en de ontwikkeling van een e-learningcursus;
- text mining en natural language processing.

Voor de literatuurstudie werden ondermeer de universiteitsbibliotheek van de Universiteit Hasselt te Diepenbeek en de Provinciale Bibliotheek Limburg te Hasselt geraadpleegd. Het aantal relevante werken dat in deze bibliotheken gevonden werd, is beperkt. Daarnaast werd er op het internet gezocht naar relevante artikelen. Door de aard van het onderwerp werden hier wel vele relevante bronnen gevonden. De literatuurstudie is bijgevolg ook grotendeels gebaseerd op bronnen afkomstig van het internet.

Het tweede deel van deze eindverhandeling bestaat uit het uitvoeren van een gevalstudie. Deze bestaat eruit om de aan text mining-gerelateerde technologie van i.Know toe te passen in de ontwikkeling van een e-learningcursus. Het automatiseren van de decompositiefase tijdens de creatie van een e-learningcursus met behulp van de technologie van i.Know is het concrete onderwerp van het praktijkprobleem.

In eerste instantie er kennis gemaakt met het systeem van i.Know enerzijds door het doornemen van bedrijfspapers en nadien met het uitvoeren van enkele eerste stappen. Aan de hand van een interactieve samenwerking met zowel i.Know als prof. Schreurs werd in eerste instantie de kennis van het systeem uitgebreid. Vervolgens werd van het systeem gebruik gemaakt in enkele stappen van het decompositie-proces van een e-learningcursus. Tot slot werden hieruit conclusies getrokken.

### **1.4 Overzicht van de eindverhandeling**

De literatuurstudie vangt aan in het tweede hoofdstuk met een bespreking van de evolutie naar nieuwe onderwijsconcepten in e-learning. Eerst wordt er een terugblik gemaakt naar het pre-e-learning tijdperk. Vervolgens komen de recente evoluties binnen e-learning uitgebreid aan bod. Het hoofdstuk sluit af met een overzicht van enkele nieuwe trends binnen e-learning.

Hoofdstuk drie formuleert een antwoord op de tweede deelvraag. In een eerste deel worden de *learning objects* of leerobjecten uitgediept. Het tweede luik van dit hoofdstuk overloopt de ontwikkeling van een e-learningcursus en beschrijft tevens welke rol leerobjecten hierin spelen.

De overeenkomsten en verschillen tussen *text mining* en *natural language processing* worden in hoofdstuk vier behandeld. Vervolgens wordt er een kader van text mining geschetst. Het derde gedeelte van dit hoofdstuk bekijkt de uitgangspunten en functionaliteiten van de technologie van i.Know.

De gevalstudie wordt uitgewerkt in hoofdstuk vijf. Na de bespreking van het doel worden de gevolgde stappen uitgewerkt om topics te herkennen, samenvattingen te genereren en een lijst van sleutelwoorden op te stellen. Het hoofdstuk wordt afgesloten met enkele conclusies.

In het laatste hoofdstuk wordt een overzicht gegeven van de conclusies die uit deze eindverhandeling getrokken kunnen worden.

## **HOOFDSTUK 2 : Evolutie naar nieuwe onderwijsconcepten in e-learning**

### **2.1 Inleiding**

De exponentiële groei aan informatie ligt aan de basis van de kenniseconomie van vandaag. Deze informatie en kennisbronnen moeten door ons verwerkt worden; daardoor wordt de nood aan efficiënte leersystemen groter dan ooit. Niet alleen krijgen we een groter volume aan informatie te verwerken, daarnaast moeten we deze kennis en vaardigheden ook steeds sneller verwerven. De omgeving verandert voortdurend en daarenboven volgen de veranderingen elkaar steeds sneller en sneller op. Hierdoor vermindert de tijd die beschikbaar is om nieuwe zaken te leren.

Oude leermodellen hebben het moeilijk om om te gaan met deze massa aan informatie en de steeds toenemende snelheid waarbij deze informatie op ons af komt. Het is dus noodzakelijk om nieuwe leerconcepten te ontwikkelen om deze evoluties te kunnen bijbenen. Vooraleer de nieuwe onderwijsconcepten in e-learning aangekaart worden, volgt eerst een kort historisch overzicht van het gebruik van 'nieuwe' technologieën in het leerproces en vervolgens wordt het begrip e-learning onder de loep genomen.

### **2.2 Het 'pre-e-learning'-tijdperk**

Sinds het begin van de twintigste eeuw hebben verschillende technologische vernieuwingen een grote impact gehad op de samenleving in het algemeen, maar ze hebben ook ons dagdagelijkse leven –ingrijpend– veranderd. De technologische evoluties gingen ook voor het onderwijs niet ongemerkt voorbij. In het niet-klasseke onderwijs, zoals onder andere het afstandsleren, sprong men steeds snel op de kar. De mogelijkheden van de nieuwe technologieën werden telkens snel verkend om later eventueel zelfs te adopteren (Harting & Erthal, 2005).

We bekijken eerst beknopt de opkomst van de nieuwe media en vervolgens de invloed van de computer.

#### **2.2.1. Nieuwe media in afstandsleren**

Rond 1928 begonnen scholen de radio te gebruiken in het volwassenen-onderwijs. De uitzendingen werden wel aangevuld met een bijbehorende cursus en boekenlijst. Gedurende de jaren '50 was de British Broadcasting Corporation zeer actief bezig met opvoedkundige uitzendingen (Holmberg, 1995, in Harting, 2005).

Tijdens de Tweede Wereldoorlog zocht het Amerikaanse leger een middel om een grote doelgroep, verspreid over de hele wereld, op een consistente manier op te leiden. De film bracht hier de oplossing. Vervolgens werden er een hele reeks educatieve films gemaakt: gaande van wapenonderhoud tot films over persoonlijke hygiëne. Na de oorlog vloeiende hier een heuse opvoedkundige filmmarkt uit voort.

Een vervolg hierop vormde de televisie. In combinatie met de videorecorder leek de job van onderwijzer overbodig te worden. In de jaren '60 kenden de educationele televisiestations dan ook een grote opmars. Maar de aanschaf van de technologische componenten consumeerde een grote hap uit het budget zodat er voor het creëren van de software niet meer veel overbleef. Het gevolg was, dat er saaie programma's uitgebracht werden die helemaal niet aansloegen bij de studenten. Het grootste probleem met dit medium was echter het ontbreken van ook maar enige vorm van interactie met de gebruiker (Rosenberg, 2001).

### **2.2.2. Computer-based training**

De voorgaande mediavormen bleken wel geschikt om een groot publiek op verspreide locaties te bereiken. Maar het feit, dat er geen interactiviteit mogelijk was, beperkte de lessen tot het overbrengen van de instructies of leerstof.

Vaak denken we dat *computer-based training* (CBT<sup>1</sup>) een recent begrip is. Niets is echter minder waar; het gebruik van de computer voor leerdoeleinden dateert al van ruim dertig jaar geleden. In het begin van de jaren zestig had men reeds een zeer eenvoudig leersysteem ontwikkeld. Met behulp van een terminal, die gelinkt was aan een mainframe, kon men eenvoudige meerkeuzevragen op de gebruiker afvuren. Op honderde ponskaarten stond de programmacode beschreven, die met 'juist', 'fout' en 'probeer opnieuw' uiterst minimale feedback aan de gebruiker geven kon.

De noodzaak aan interactiviteit zorgde voor een vernieuwde aandacht voor *computer-based training* in de jaren zeventig en tachtig. Het bekendste voorbeeld uit die periode is zonder twijfel PLATO<sup>2</sup>. Voor dit op een mainframe gebaseerde leersysteem werden goede educatieve programma's ontwikkeld; zodoende werd PLATO al snel het dominante CBT-platform.

Met de opkomst van de *personal computer* was het mogelijk om CBT-programma's op een *floppydisk* op de markt te brengen. Deze leerprogramma's op een floppydisk waren uiteraard goedkoper en gebruiksvriendelijker en bovendien evenaarden ze de kwaliteit van PLATO. Daarnaast

---

<sup>1</sup> CBT ook wel de afkorting voor computer-based teaching.

<sup>2</sup> PLATO is origineel de afkorting voor *Programmed Logic for Automated Teaching Operations*.

waren deze programma's aantrekkelijker, omdat ze gebruik maakten van een grafische gebruikersinterface, kleuren en multimedia. PLATO, dat bestond uit zeer dure gepatenteerde hardware en software, die enkel voor leerdoeleinden gebruikt kon worden, werd bijgevolg al snel onttroond en de pc wierp zich op als het nieuwe dominante platform voor CBT<sup>3</sup>.

De floppydisk werd als mediadrager al snel vervangen door de cd-rom. De grote opslagcapaciteit van de cd-rom liet toe om gebruik te maken van verschillende vormen van digitale media: video-beelden, gesproken instructies, muziek, afbeeldingen, animaties, ... Hieronder was het mogelijk om cursussen aan te bieden die zowel innovatief, interactief als effectief waren.

Alhoewel de vooruitzichten voor kant-en-klare CBT-pakketten met de introductie van de pc veelbelovend leken, kwamen er halverwege de jaren '90 een aantal problemen aan de oppervlakte.

De personal computer raakte steeds meer verspreid en bood aldus een uitzicht op een universele markt, maar de gebrekkige compatibiliteit verhinderde dit. Zo moest men bijvoorbeeld verschillende versies ontwikkelen voor Apple, Unix en IBM systemen, hetgeen uiteraard een zeer kostelijke zaak was. Toen men deze problemen overwonnen had, maakten de snel op elkaar volgende technologische evoluties de software al vlug overbodig. Dit was bijvoorbeeld het geval bij de overgang van de 5,25 inch naar de 3,5 inch floppydisks.

Een volgend probleem was het feit, dat de snel veranderende kennisbank de inhoud van de CBT-software snel verouderde. Technologische evoluties, productaanpassingen, reorganisatie van de ondernemingen, marktevoluties, ... lagen aan de basis van deze constant veranderende kennisbank. Om te vermijden dat de inhoud reeds achterhaald was alvorens de software goed en wel de deur uit was, werd de stabiliteit van de inhoud een criterium dat eerst goed overwogen werd.

Daarnaast groeide de behoefte aan industrie- of domein-specifieke cursussen. Aangezien deze uiteraard veel meer middelen vragen om te ontwikkelen en bovendien een gevoelig kleinere afzetmarkt hebben, hadden deze een negatieve invloed op de massa-productie van CBT-cursussen. Bovendien werd er door bedrijven als Macromedia met een software-tool als *Authorware* handig op ingespeeld. Met *Authorware* is het voor CBT-ontwikkelaars met weinig ervaring mogelijk om toch aantrekkelijke cursussen te creëren.

---

<sup>3</sup> PLATO werd overgenomen door The Roach Organization, Inc., (TRO). Het PLATO-platform werd vervolgens helemaal omgevormd tot een flexibel en open platform gebaseerd op internationale standaarden. In 2000 veranderde de naam in PLATO Learning, Inc. Momenteel is PLATO Learning, Inc., de voornaamste aanbieder van educatieve software. (<http://www.plato.com>)

Tot slot werden er grote vooruitgangen geboekt in het begrijpen van hoe mensen leren. Het onderzoek in de onderwijspsychologie heeft aangetoond, dat leren vooral plaatsvindt via sociale interactie (Rubens, 2003). Deze en andere nieuwe inzichten en principes uit de 'zachte' wetenschappen, werden ook in het CBT-domein geïntroduceerd. Maar de implementatie van de meeste efficiënte leermethoden werd gehinderd door de beperkingen van de op dat moment beschikbaar zijnde computertechnologieën.

## **2.3 E-learning**

Stilaan maakt e-learning zijn opwachting. Vooraleer de ontwikkelingsstadia van e-learning uitvoerig besproken worden, wordt er eerst even stilgestaan bij de factoren die de opkomst van e-learning mogelijk maakten. Nadien wordt het begrip e-learning gedefinieerd en worden enkele voor- en nadelen ervan opgesomd.

### **2.3.1. De oorzaken**

#### **2.3.1.1 Maatschappelijke veranderingen**

In de paragraaf over *Computer-Based Training* werden de veranderingen in de opvattingen over leren en opleiden gedurende de jaren negentig reeds aangehaald. Deze evoluties zijn slechts één aspect van een aantal maatschappelijke veranderingen die hebben plaatsgevonden, waardoor een sterke opmars van e-learning mogelijk werd.

In het laatste decennium van de 20ste eeuw is onze maatschappij geëvolueerd naar een kennis- en informatiemaatschappij. Parallel met dit proces is ook de economische wereld aangrijpend veranderd. De alomtegenwoordige massaproductie van weleer is overgeschakeld naar een economie die inspeelt op de vraag en wensen van de consumenten. Deze wensen zijn niet alleen gevarieerd maar bovendien ook onderhevig aan schommelingen. Bolhuis en Simons (1999, in Rubens, 2003) stellen dan ook dat deze veranderingen in de samenleving en het bedrijfsleven een toenemende en bovendien permanente behoefte aan leren teweeggebracht hebben.

Aan de noodzaak voor 'levenslang leren' kunnen de klassieke onderwijs- en opleidingsvormen volgens Rubens (2003) niet in voldoende mate tegemoetkomen. Enerzijds hebben de traditionele scholingsvormen een te groot tijdsbestek nodig om adequate werknemers te vormen. Anderzijds zijn de vervolgopleidingen te algemeen-vormend voor de veelal zeer specifieke omgeving in het bedrijfsleven.

Het is voor bedrijven dan ook een must om jonge werknemers een degelijke opleiding te bieden. Maar daar houdt het echter niet bij op. Ook meer ervaren werknemers hebben nood aan bijkomende vorming om op een verantwoorde manier de marktevoluties te kunnen volgen en hier

vervolgens op te kunnen inspelen. Gezien het belang en de kwantiteit van de scholing toeneemt, gaan de ondernemingen de toegevoegde waarde van opleidingen met een steeds meer kritische blik bekijken. Rosenberg (2001) stelt zelfs dat:

*"[...] training is accountable for the same primary measure as any other function: business value."*

Een opleiding moet dus niet alleen de werknemer iets bijbrengen, maar moet ook de organisatie in het geheel iets opleveren. Daarenboven moet, gezien training als elke andere bedrijfsfunctie behandeld wordt, die opleiding zo efficiënt mogelijk zijn in zowel tijd als kosten.

Naast de maatschappelijke en bedrijfs-economische ontwikkelingen, spelen de nieuwe inzichten op het vlak van onderwijspsychologie een rol voor het aanslaan van e-learning. Zo stapte men af van het klassikale systeem van onderwijs, waarbij de leerlingen enkel passief luisteren naar de kennis en instructies overgebracht door de leerkracht. Men heeft meer aandacht voor een leerproces waarbij de leerling centraal staat en niet de boodschapper. Dit tracht men te verwezenlijken door de eigen verantwoordelijkheid van de lerende te versterken, door het creëren van een leermethode waarbij men moet samenwerken met andere leerlingen en het benutten van de relevante werksituatie als leeromgeving (Rubens & Admiraal, 2003).

#### **2.3.1.2 Internethype**

De maatschappelijke veranderingen op zich zijn niet voldoende om de snelle opkomst van e-learning te verklaren. Zoals in het 'pre-e-learning'-tijdperk fungeerde ook de opmars van een nieuwe technologie als hefboom, in dit geval het world wide web of gewoonweg het internet.

Halverwege het laatste decennium van de vorige eeuw kreeg het internet een enorme boost. Op korte tijd kende het digitale wereldnetwerk een zeer brede acceptatie door de samenleving en ontketende het zowaar een (nieuwe) digitale revolutie. De tot stand gekomen wereldwijde standaarden, zoals TCP/IP en HTML, maken het mogelijk dat informatie van op elke plaats en bovendien ook op elk moment opvraagbaar is. Communicatie kan vanaf nu dus *any time and any place* plaatsvinden (Rubens, 2003).

Voor bedrijven en overheden bracht het internet een waaier aan nieuwe mogelijkheden met zich mee. In het bedrijfsleven kan men door de opkomst van *e-business* en *e-commerce* zelfs spreken van een totaal nieuw economisch perspectief. Ook de gewone burgers ervoeren de effecten die de doorbraak van de internettechnologie teweegbracht. Waarschijnlijk zijn er vandaag veel mensen die hun elektronische brievenbus veel frequenter controleren dan de klassieke versie aan de rand van weg. Kortom, e-mail is gewoon niet meer uit ons dagelijks leven weg te denken en dit is nog maar één voorbeeld. Dus zonder ook maar enige twijfel had internet –en heeft het misschien nog steeds– een immense invloed op de volledige samenleving.



Voor de educatieve sector was er eveneens geen ontkomen aan. De brede acceptatie schepte mogelijkheden om 'Informatie- en Communicatietechnologie' (ICT) op een brede schaal toe te passen, stelt Rubens (2003). Bij de maatschappelijke veranderingen die in de vorige paragraaf besproken werden, is reeds geopperd, dat men afstapt van de traditionele klassikale onderwijsvormen en men zich meer richt naar opleidingsstructuren waar leren kan plaatsvinden door sociale interactie. Rosenberg (2001) geeft aan, dat de internettechnologie deze veranderingen in de onderwijsmethode kan ondersteunen. Door internettechnologie is het mogelijk om de interactiemogelijkheden bij gebruik van ICT bij leerprocessen aanzienlijk te verbeteren. Door de nieuwe communicatiemogelijkheden kunnen de betrokkenen tijds- en plaatsonafhankelijk kennis onder elkaar uitwisselen. Merk hierbij bovendien op, dat deze interactie tussen mensen plaatsvindt. Voorheen, met CBT-pakketten, was er enkel sprake van een interactie tussen mens en machine (computer).

De internethype en de ontwikkeling van e-learning maakte het dus mogelijk om de 'verbeterpunten' van het traditionele leren aan te pakken. Het internet creëerde also een aantal nieuwe verwachtingen met betrekking tot het 'leren'. Ten eerste zou leren sneller, efficiënter en veel interactiever worden (Dobbs, 2002, in Rubens, 2003). Daarnaast dacht men, dat het mogelijk zou worden om complexe taken binnen 'authentieke' leeromgevingen te onderwijzen. Tot slot beweren Rubens en Admiraal (2003) dat het internet het inrichten van leerprocessen, waarbij de lerende en het leerproces centraal staan, zou faciliteren. Bij deze leerprocessen spelen de interactie en communicatie een centrale rol en internet zou dit zeker vergemakkelijken.

De veranderingen in de samenleving brachten dus nieuwe en andere noden betreffende leren aan de oppervlakte. Het bedrijfsleven zocht nieuwe kanalen om hun werknemers –over de hele wereld– zo effectief en efficiënt mogelijk te kunnen opleiden. Het internet bracht hier met mogelijkheden tot tijds- en plaatsonafhankelijke communicatie soelaas. In het onderwijs had men behoefte aan meer middelen om interactie en samenwerking tussen en met de lerenden tot stand te brengen, en ook hier bracht het internet een oplossing. De combinatie van de evoluties in de samenleving waardoor nieuwe 'vragen' ontstonden en de opkomst van het internet als belangrijk deel van het antwoord, zorgde ervoor dat er een rijke voedingsbodem was voor e-learning. Deze rijke voedingsbodem zou een snelle groei van e-learning mogelijk maken, misschien een te gezwinde opgang, zoals blijkt uit het vervolg.

### **2.3.2. Ontwikkelingsstadia van e-learning**

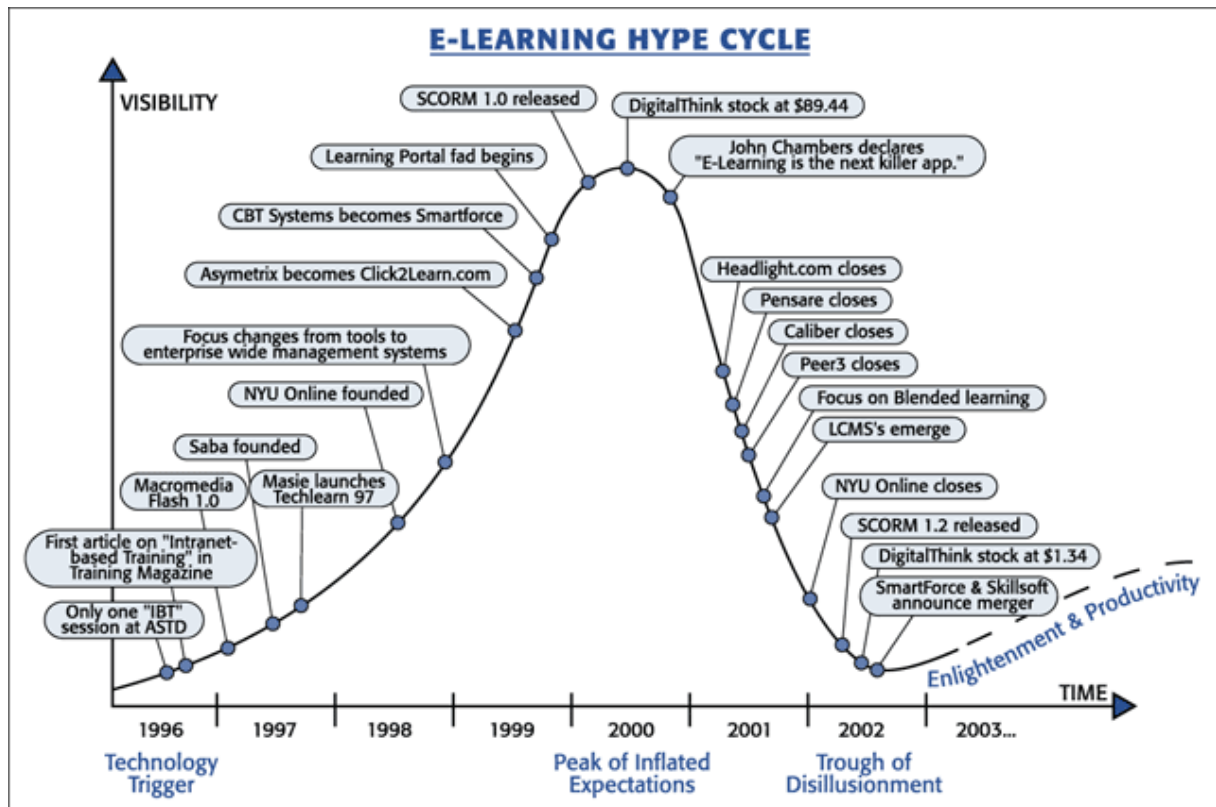
Zoals reeds gesteld, liet de introductie van nieuwe technologieën in het onderwijs nooit lang op zich wachten. In het verleden creëerde elke technologisch evolutie steeds weer grote verwachtingen. Maar na een hoopvolle periode, moest men deze verwachtingen telkens opnieuw terugschroeven, aangezien er een aantal belemmerende factoren aan de oppervlakte kwamen. De

opkomst van het internet en e-learning beantwoordt ook aan dit patroon. Deze vaststelling wordt bevestigd door Williams en Goldberg (2005):

*"E-learning has gone through several evolutionary phases, some of which have been highly optimistic, while others have been characterised by considerable pessimism."*

Om deze opeenvolgende fasen te beschrijven heeft Kruse (2002) de *Gartner's Hype Cycle for Emerging Technologies* op de geschiedenis van e-learning geprojecteerd. Figuur 1 toont de E-learning Hype Cycle van Kruse. Deze Gartner's Hype Cycle for Emerging Technologies, ondermeer ontwikkeld door Linden, stelt dat nieuwe technologie altijd volgens een zelfde patroon evolueert. In dit patroon kan men drie fasen onderscheiden:

1. fase van overdreven enthousiasme (overenthusiasm);
2. fase van desillusie (desillusionment);
3. fase van geleidelijke verbetering (gradual improvement).



**Figuur 1: E-learning hype cycle<sup>4</sup> (Bron: Kruse, 2002)**

In de volgende paragrafen worden de verschillende stadia in de e-learning evolutie verder uitgediept.

<sup>4</sup> Figuur overgenomen van Kruse (2002).

### **2.3.2.1 Fase van overdreven enthousiasme**

Eén van de twee aangehaalde oorzaken voor de snelle opkomst van e-learning is, de internetboom die vanaf 1995 plaatsvond. Kruse (2002) plaatst de concrete aanleiding bij een druk bijgewoonde workshop met als titel "*Intranet-based training*" tijdens de ASTD Conference van 1996. Naar aanleiding van deze workshop zijn er tal van artikels verschenen, nieuwe ideeën gecreëerd en bovendien belangrijke contracten afgesloten.

De opkomst van het internet maakte dat men ging experimenteren met de internettechnologie in leertoepassingen. Dit schepte hoge verwachtingen, zoals aanzienlijke kostenbesparingen, grotere flexibiliteit en een verbetering van de kwaliteit van de leeromgeving.

In deze fase werd er veel geld geïnvesteerd in onderzoek naar en ontwikkeling van naar mogelijk nieuwe toepassingen van de internettechnologie in de educatieve sector. Het resultaat is dan ook een brede waaier aan toepassingen, gaande van klant- tot leveranciersgerichte applicaties. De meeste toepassingen zijn echter wel bedoeld om de medewerkers efficiënter en effectiever te laten functioneren, stelt Rubens (2003). Een voorbeeld hiervan is een toepassing, waarbij handelsvertegenwoordigers via e-learning een nieuwe CRM-toepassing aanleren.

Daarenboven geeft de *dot.com*-boom aanleiding tot het ontstaan van vele e-learningportalen, zoals ondermeer KnowledgePlanet en Click2Learn. Bij deze portalen kan men tegen betaling een heel uitgebreid gamma aan cursussen verkrijgen. Deze portalen beloven de gebruikers een goedkope, onbeperkte toegang tot onderwijs. Ze laten hen geloven dat leren vanaf nu zó efficiënt is, dat het als het ware vanzelf gaat. Bovendien profileren deze leerportalen zich als veel flexibeler en eenvoudiger dan de traditionele onderwijsvormen (Williams & Goldberg, 2005). De talrijke leerportalen hadden weliswaar een zeer uitgebreid gamma, maar de kwaliteit was vaak ondermaats (Cross & Hamilton, 2002).

Aangezien men eenvoudig investeerders kan vinden die het vereiste kapitaal willen verschaffen, reizen nieuwe e-learningbedrijven als paddenstoelen uit de grond. Zo worden rond 1996-1997 ondermeer Saba, door ex-directieleden van Oracle, en TechLearn, door Elliot Masie, opgericht. Al vlug stappen deze bedrijven naar de Amerikaanse technologiebeurs NASDAQ, waar hun koersen snel een hoge vlucht nemen. De reden voor deze hoge koersen is, dat ze gebaseerd zijn op de verwachte in plaats van op de gerealiseerde winstcijfers (Rubens, 2003). John Chambers, CEO<sup>5</sup> van CISCO, deed daar nog een schepje bovenop door e-learning "*the next big killer application*" van het internet te noemen.

---

<sup>5</sup> CEO is de afkorting van Chief Executive Officer, oftewel algemeen directeur.

Kenmerkend voor deze fase is de opkomst van het *Learning Management System* (LMS) (Cross & Hamilton, 2002). Het LMS is een samenvloeiing van twee verschillende systemen, namelijk *Computer-Managed Instruction* (CMI) systemen en *Training Management Systems* (TMS). Het CMI-systeem, ook wel *Course Management System* (CMS) genoemd, was een WAN<sup>6</sup>-, intranet- of webgebaseerd systeem, dat centraal de administratie en de vorderingen van de studenten registreerde. Dit systeem werd in de eerste plaats gebruikt voor cd-rom of CBT-cursussen en later ook voor web-based training. Marcormedia, Asymetrix en Allen Communication waren de eerste verdelers van deze CMI/CMS-systemen.

Het Training Management System focuste ondermeer op de volgende twee functionaliteiten: het modelleren van de vaardigheden van de werknemers en het meten van ontbrekende vaardigheden door online testen. Voorbeelden van fabrikanten van dergelijke TMS zijn KnowledgeSoft<sup>7</sup>, Syscom<sup>8</sup> en DKSystems.

Gezien de functionaliteiten van de CMI/CMS- en TMS-toepassingen elkaar steeds meer begonnen te overlappen, evolueerden beide systemen eind jaren '90 naar één geheel, namelijk het *Learning Management System*. Deze op Oracle-gebaseerde systemen waren een uitgebreide poging om het beheer, het management, de opvolging en de rapportering van vaardigheden, de klassikale cursussen en CBT doorheen het hele bedrijf te centraliseren. Saba, Docent en Plateau Systems zijn enkele leveranciers van dergelijke LMS-systemen.

### **2.3.2.2 Fase van desillusie**

Tijdens de jaren 1999 en 2000 rezen de verwachtingen van e-learning tot een ongekennde piek. Zo stegen ook de aandelen van de ondernemingen actief in e-learning tot recordhoogtes. De aandelenkoers van Digital Think noteerde in 2000 een recordhoogte van maar liefst \$ 89.44 (Kruse, 2002).

Sinds 1999 verdwenen de cd-romcursussen grotendeels en werden ze vervangen door webgebaseerde trainingssessies (Cross & Hamilton, 2002). De cursussen kon men dan via het intranet van de onderneming of vaak ook gewoon via het world wide web opvragen. Deze overgang van cd-rom naar het web leverde een aantal voordelen op. E-learning kon veel sneller geïmplementeerd worden aangezien het slechts op één server geïnstalleerd moest worden. Het was dus niet meer nodig om talrijke kopieën te verdelen en vervolgens nog eens bij te houden op welke locaties deze zich bevonden. Bijkomend voordeel van het feit, dat het bronmateriaal slechts op één locatie aanwezig was, is dat men ook maar één update diende uit te voeren. Een update van een

---

<sup>6</sup> WAN is de afkorting van Wide Area Network.

<sup>7</sup> Nu KnowledgePlanet.

<sup>8</sup> Syscom veranderde eerst in TrainingServer en heet nu THINQ Learning Solutions.

cd-rom was echter niet mogelijk, waardoor de informatie op de cd-rom al vlug verouderde. Web-gebaseerde cursussen hadden dus een langer 'leven' aangezien een update eenvoudig uitgevoerd kon worden. Daarnaast kon men van op elke plaats en ook op elk moment het e-learningportaal raadplegen. Men diende enkel over een pc met een webbrowser en een connectie tot de *host server* te beschikken. Tot slot kon men met de implementatie van e-learning over het intranet, gedeeltelijk de hoge kosten van de installatie van het intranet verantwoorden.

De gevolgen van e-learning leken veelbelovend, maar al gauw doken na de eerste praktijkervaringen rond eind 2000 de eerste problemen op.

De smalle bandbreedte vormde een aanzienlijke beperking voor het gebruik van verschillende media in een e-learningcursus. De cd-romcursussen waren rijkelijk geïllustreerd met figuren, video, geluid, ... Als men gebruik maakte van deze media in de eerste e-learningcursussen, dan duurde het, door de smalle bandbreedte, redelijk lang alvorens deze op de computer van de gebruiker ingeladen waren. De lange wachttijden veroorzaakten bij de gebruikers een negatieve leerervaring (Cross & Hamilton, 2002).

Men verwachtte dat e-learning het zou mogelijk maken om in het leerproces de lerende persoon centraal te stellen en via sociale interactie het leren te stimuleren. Maar de beperkingen van de beschikbare webtechnologieën lieten immers geen of in beperkte mate enige vorm van interactiviteit toe (Rubens, 2003). Het gevolg was, dat de e-learningcursussen opnieuw voornamelijk instructiegericht waren en er dus eigenlijk een stap terug in de tijd gezet werd. Sarcastisch werd gezegd dat de 'e' uit e-learning stond voor eenzaam. Deze eenzaamheid leidde samen met andere negatieve ervaringen er vaak toe, dat de gebruiker de cursus niet voltooide.

Hoewel men met e-learning op technologisch vlak een grote vooruitgang had geboekt in vergelijking met de cursussen op cd-rom, betekende dit op didactisch vlak eigenlijk een achteruitgang. In theorie had men met het internet als nieuw platform veel meer mogelijkheden. Maar de praktijk leerde, dat de webtechnologieën en de technische capaciteiten niet voldoende ontwikkeld waren om de theoretische ideeën in werkelijkheid om te zetten.

Naast de negatieve ervaring voor de gebruikers, liep het aan de kant van de ontwikkelaars van e-learningtoepassingen ook helemaal niet zoals verwacht. De implementatie van een e-learningapplicatie en daarbovenop de ontwikkeling van e-learningleerstof van een degelijk niveau brachten zeer hoge kosten met zich mee. Bovendien haalde de omzet niet de streefcijfers, vermits het aandeel van e-learning in de educatieve markt onder de verwachtingen bleef. De hoge kosten voor de ontwikkelingen en implementatie, gecombineerd met de lage omzet, brachten de producenten onvermijdelijk in grote financiële problemen. Deze financiële moeilijkheden kenden uiteraard hun weerslag op de aandelenkoersen, met als gevolg dat de hele sector in een

economische crisis verzeilde. Vanaf 2001 sloten een aantal ondernemingen, actief in de e-learningsector, hun deuren en vele anderen werden verkocht of fuseerden, zoals SmartForce en Skillsoft in 2002.

Door de financiële crisis in de sector werden de investeringen in onderzoek en ontwikkeling teruggeschroefd, met als logisch gevolg dat er geen echte innovaties meer plaatsvonden. Hier bestaat echter één uitzondering op, namelijk het *Learning Content Management System* (LCMS). Met LCMS is het mogelijk om leerobjecten te ontwikkelen, opnieuw te gebruiken, te beheren en centraal aan te bieden. Met het gebruik van dergelijke systemen verwachtte men grote voordelen op het vlak van efficiëntie.

In het onderwijs, waar financiële aspecten een veel kleinere rol spelen, kon e-learning ook niet echt een voet aan de grond zetten. Onderzoek van E-Learning Plaza toont aan, dat e-learning in het onderwijs geen fundamentele verandering teweeggebracht heeft, maar dat er eerder sprake is van automatisering. De elektronische leeromgevingen worden voornamelijk gebruikt om informatie uit te wisselen, maar er is geen sprake van een vernieuwing van het onderwijs. De redactie van E-learning Plaza haalt als voornaamste reden aan, dat een onderwijsvisie afgestemd op de nieuwe technologie nog ontbreekt. (E-Learning Plaza, 2003 in Rubens, 2003)

Vanaf 2001 stelt Kruse (2002) vast dat men eerder focust op *blended learning* in plaats van op e-learning. Wat dus betekent dat e-learning één van de middelen is om de leerstof over te brengen op de lerenden.

In de vorige fase kenden de leermanagementsystemen een sterke opkomst, maar ook deze systemen voldeden uiteindelijk niet aan de hoge verwachtingen. Cross en Hamilton (2002) halen hiervoor onder meer de volgende oorzaken aan:

- te lange implementatietijd;
- enkel administratieve aspecten werden behandeld;
- geen *plug-and-play* met de diverse softwarepakketten (standaarden ontbraken);
- de administratieve processen waren niet aanpasbaar aan de specifieke situaties.

Deze vaststellingen tonen aan, dat er nood was aan één technologie-platform voor e-learning. Er waren verschillende technologieën en toepassingen op de markt, maar het grote probleem bestond erin, dat er geen éénduidige technologie of universele standaarden waren (Rubens, 2003). Hoewel SCORM 1.0<sup>9</sup> in januari 2000 reeds uitgebracht werd, werd deze nog niet algemeen geaccepteerd om als standaard te kunnen fungeren.

---

<sup>9</sup> SCORM is de afkorting voor Sharable Content Object Reference Model.

E-learning had in het begin hoge verwachtingen geschapen, maar na een tijdje bleek al gauw dat deze niet helemaal ingelost konden worden. De ene tegenvaller na de andere zorgde ervoor dat de sector in een fase van desillusie was geraakt. Op technologisch vlak werd er wel vooruitgang geboekt, maar de technologische mogelijkheden werden niet ten volle benut, zodat er op didactisch vlak een stap achteruit gezet werd.

### **2.3.2.3 Fase van geleidelijke verbetering**

Kruse (2002) stelde, dat er aan de leverancierszijde een verdere consolidatie nodig was, alvorens er een periode van *enlightenment* en productiviteit zou intreden. Kruse schatte, dat we deze fase rond 2004 konden bereiken. Boehle (2006) beweert, dat deze marktregulering inmiddels voltooid is en stelt dat met de overname van THINQ door Saba, de zwakke spelers uit de markt verdwenen zijn. De overblijvende ondernemingen draaien een voldoende grote omzet om hun positie in de markt te handhaven. Deze 'opruimingsactie' aan de aanbodzijde heeft eveneens tot gevolg gehad, dat de aandelenkoersen nu aan een realistische waarde genoteerd staan (Rubens, 2003).

Williams en Goldberg (2005) beamen dat we momenteel in een periode van 'vernieuwd optimisme' aanbeland zijn en argumenteren dat er bewijzen zijn, dat er een periode van continue groei mogelijk is. De grote uitdaging in deze fase is volgens hun, om de beschikbare technische capaciteiten (met name bandbreedte), de webtechnologieën en de pedagogische kennis succesvol te combineren. Als men hierin lukt, dan kan er werkelijk sprake zijn van een revolutie. De visie van Rubens (2003), dat er in deze fase duidelijkheid ontstaat over de echt zinvolle applicaties, sluit hierop aan.

In het bedrijfsleven stellen Rubens (2003) en Boehle (2006) dat e-learning zich integreert met de bedrijfskritische applicaties. We zien reeds dat grote spelers als SAP en Peoplesoft hun pijlen meer en meer richten op e-learning. Adkins (2003, in Rubens, 2003) ziet deze integratie als een noodzakelijke voorwaarde om tot tweede generatie producten te komen. In ondernemingen wordt er getracht een sterke verbinding tot stand te brengen tussen e-learning enerzijds en de bedrijfsdoelstellingen anderzijds. Dit betekent dus, dat elke afdeling voortaan niet meer naar eigen willekeur e-learningtoepassingen kan implementeren, maar dat e-learning een vaste plaats krijgt binnen het centrale *human resources*-beleid voor de gehele onderneming. De werknemers zien e-learning ook opduiken in de concrete werkmethoden (Boehle, 2006).

In het onderwijs verschuift het zwaartepunt van de technologie –in de vorige fase– naar het didactische aspect. Docenten en onderwijzers gaan na welke functionaliteiten van e-learning voor welk groep van lerenden ingezet kan worden en hoe online begeleiding hierbij mogelijk is (Rubens, 2003). Men creëert dus een leerproces waarbij e-learning en de traditionele onderwijsvorm, namelijk het klassikaal leren, elkaar aanvullen en zelfs versterken. E-learning maakt het mogelijk om de leerstof in verschillende formaten aan te bieden, voorziet in interactie met collega-studenten

door middel van discussiefora of chat-toepassingen. Het gevolg is, dat men minder tijd dient te besteden aan het 'gewoon' overbrengen van de inhoud en er dus meer tijd overblijft voor reflectie en kritische analyse (Williams & Goldberg, 2005).

Met de opkomst van het internet ging ook voor de educatieve sector een nieuwe wereld open. Een wereld die bijna het 'leerparadijs' op aarde beloofde. De verwachtingen waren zeer groot en e-learning nam dan ook meteen een zeer hoge vlucht. Maar vanaf de jaren 2000-2001 begon deze droom stilaan in elkaar te stuiken. Sinds 2003-2004 heeft e-learning een nieuwe geleidelijke start genomen en is het geleidelijk aan fundamentele veranderingen aan het teweegbrengen.

### **2.3.3. Definitie e-learning**

Niettegenstaande e-learning reeds een hele weg afgelegd heeft, is het toch nog altijd een relatief jong fenomeen. De term e-learning dook pas in 1999 voor het eerst op, eerder deden benamingen als *webbased*-leren, teleleren of online leren de ronde (Rubens, 2003). Uiteindelijk werd de term e-learning populairder, waarschijnlijk omwille van de analogie met de uitdrukkingen *e-mail*, *e-business* en *e-commerce*. In deze benamingen duidt de letter 'e' op het aanwezig zijn van een elektronische component. In dit geval zou men kunnen zeggen, dat de letter 'e' staat voor het gebruik van de elektronische snelweg, het internet met andere woorden, en de bijbehorende technologie.

Gezien e-learning dus nog een redelijk jong begrip is, valt het eigenlijk moeilijk te definiëren. Bijgevolg circuleren er dan ook talrijke, verschillende en veranderende definities.

In 2001, hanteerde Huisman (2001) de volgende definitie voor e-learning:

*"vormen van leren door personen, die individueel of in groepsverband, door één of meer e-learning-technologieën ondersteund worden."*

In bovenstaande definitie verstond Huisman onder e-learning-technologieën het volgende:

*"Informatie- en communicatietechnologie die geschikt is om een leerinteractie met of tussen individuen te realiseren, te besturen, te volgen of te administreren."*

De definiëring van Huisman is niet helemaal correct, want mediavormen als radio en televisie vallen volgens Huisman ook onder e-learningtechnologieën. Nu kan men e-learning wel beschouwen als een opvolger van het zogenaamde afstandsleren, want onderwijs is eveneens plaats- en zelfs tijdsonafhankelijk mogelijk. Maar het is wel belangrijk om aan te geven dat er gebruik gemaakt wordt van internettechnologieën, zoals Fallon en Brown (2003) wel doen in hun omschrijving van het begrip e-learning:

*"any learning, training or education that is facilitated by the use of well-known and proven computer technologies, specifically networks based on Internet technology".*



Rosenberg (2001) definieert e-learning als volgt:

*"E-learning refers to the use of Internet technologies to deliver a broad array of solutions that enhance knowledge and performance."*

De definitie van Rosenberg steunt op drie fundamentele criteria:

1. E-learning vindt plaats in een netwerk, hetgeen inhoudt dat het updaten, opslaan, opvragen, distribueren en delen van instructies of informatie onmiddellijk mogelijk is. Dit is een absolute voorwaarde voor e-learning en sluit media als cd-roms en dvd's uit.
2. De inhoud wordt aan de eindgebruiker geleverd via een computer en bij de communicatie wordt gebruik gemaakt van standaard internettechnologieën. Het begrip computer moet hier in de brede betekenis van het woord opgevat worden. Een personal digital assistant, gsm, enz. vallen dus ook onder de benaming computer. Belangrijker is het feit dat de communicatie plaatsvindt met behulp van internettechnologie. Onder internettechnologie verstaat men dus de webbrowser, het TCP/IP-protocol, enz.
3. De interpretatie van de definitie is niet beperkt tot het traditionele leeraspect. Met andere woorden is e-learning dus niet gelimiteerd tot het overbrengen van instructies, maar ook toepassingen die informatie leveren en hulpmiddelen die de prestaties verbeteren, vallen onder de vlag e-learning.

De definitie van Rosenberg is volgens mij de meest complete. Ze onderstreept ten eerste duidelijk dat er gebruik gemaakt moet worden van internettechnologieën. Via welke medium de gebruiker de e-learningtoepassingen kan raadplegen, is daaraan ondergeschikt, vermits de gebruikte hardware sterk aan veranderingen onderhevig is. Tegenwoordig beschikken de meeste gsm's naast een bel- en sms-functie over een connectie met het internet, via GPRS of via de nieuwste 3G-connecties. Zelfs draagbare spelconsoles, zoals de PlayStation Portable (PSP) van Sony, bezitten de mogelijkheid om een verbinding te leggen met het web. Vandaar dat het accent moet gelegd worden op het gebruik van internettechnologieën en niet op het medium dat de gebruiker hanteert. Deze voorwaarde omvat duidelijk het tweede fundamentele criterium van Rosenberg, maar impliceert eveneens het eerste. Indien internettechnologieën gebruikt worden, dan houdt dit eveneens in, dat men verbonden is met een netwerk.

De definities van Huisman en Fallon en Brown focussen op het directe leeraspect. Maar Rosenberg verruimt hier de blik en laat alle toepassingen die kennis verruimen en de prestaties verbeteren, onder de e-learning noemer vallen. Volgens mij is het zeer belangrijk om voor de toekomst zo een ruime definitie te hanteren. Er zijn immers vele toepassingen in de maak die eerder een faciliterende functie in het leerproces hebben en ook onder e-learning thuishoren. Deze definiëring van Rosenberg zorgt er eveneens voor, dat zowel vormen van formeel als informeel leren tot e-learning behoren. Terwijl men bij de overige begripsomschrijvingen eerder geneigd zou zijn om

enkel formele leerprocessen te aanvaarden. In dit werk zullen we voor e-learning de definitie van Rosenberg gebruiken.

#### **2.3.4. Voor- en nadelen van e-learning**

In de volgende paragrafen worden enkele voordelen en nadelen van e-learning op een rijtje gezet.

##### **2.3.4.1 Voordelen**

**Flexibiliteit**—Eén van de grootste voordelen van e-learning is, dat het tijds- en plaatsonafhankelijk gebeuren kan. Men kan dus 7 dagen op 7 en vanop elke plaats een e-learningtoepassing gebruiken, mits men uiteraard een browser en een connectie tot het internet voorhanden heeft. Dit maakt dat de gebruiker kan leren op het moment en de plaats dat hij er behoefte aan heeft, met andere woorden leren kan *just in time* plaatsvinden.

**Sneller inspelen op marktveranderingen**—Enerzijds kan men met e-learning zeer veel werknemers in een relatief korte tijdsspanne bereiken. Anderzijds is het mogelijk om snel een update van de e-learninginhoud uit te voeren. De combinatie van deze twee factoren maakt, dat men zeer snel kan inspelen op veranderingen in de markt. Nieuwe productontwikkelingen of een gewijzigde wetgeving kunnen snel en efficiënt doorgespeeld worden aan de betrokken medewerkers. Dit kan een onderneming een competitief voordeel opleveren ten opzichte van haar concurrenten, waar dit leerproces misschien veel trager verloopt.

**Universeel of gepersonaliseerd**—Alle werknemers krijgen identiek dezelfde e-learning cursus. Nieuwe procedures worden bijgevolg door iedereen consequent op dezelfde manier aangeleerd. Bij klassikale trainingen, gegeven door verschillende instructeurs, was het mogelijk dat er verschilpunten waren tussen de groepen. Daarnaast kan een e-learningcursus ook vrij eenvoudig afgestemd worden op de noden van een specifieke groep cursisten of zelfs op individuen. Tot slot kan iedereen op zijn eigen tempo leren en zo meer aandacht besteden aan de voor ieder persoonlijk belangrijkere delen.

**Kostenbesparingen**—Aangezien er geen klassikale sessies meer nodig zijn, dienen de cursisten zich niet te verplaatsen. Men spaart dus de vervoersonkosten en ook de reistijd uit. Bovendien zijn er minder instructeurs of begeleiders nodig voor een grote groep.

##### **2.3.4.2 Nadelen**

**Discipline**—De werknemers en studenten dienen de nodige discipline aan de dag te leggen om de e-learningcursus op eigen houtje te doorlopen. De gebruiker moet dus een hoge inzet en motivatie tonen. Bovendien is een basiskennis van pc en internet noodzakelijk. Anderzijds argumenteert

Clarke (2007) dat het gebruik van technologie in het leerproces juist motiverend werkt en mensen meer zelfrespect geeft.

**Voorkeur voor leervorm**—Vele bedienden zitten al gedurende de hele dag aan hun computer gekluisterd en zullen allicht de voorkeur geven aan klassikale cursussen. Mensen leren trouwens het liefst via persoonlijke contacten en die ontbreken natuurlijk met e-learning. Hoewel men dit tracht op te vangen met discussiefora en dergelijke, maar feit is, dat de gebruiker toch maar alleen –en eenzaam– achter zijn computerscherm zit.

**Andere aanpak**—E-learning vraagt van instructeurs een totaal andere aanpak dan de traditionele vormen. Dit betekent dat docenten nieuwe leermethoden en nieuwe begeleidingsvormen dienen uit te werken en dit kan uiteraard wel voor enige wrevel zorgen.

**Geschiktheid**—E-learning is niet voor elk topic even geschikt. Bepaalde zaken zullen omwille van hun delicateit of moeilijkheidsgraad beter in een *face-to-face*-cursus behandeld worden, denk bijvoorbeeld maar aan de medische sector.

## 2.4 Trends in e-learning

### 2.4.1. Algemene trends

In The 2006 Horizon Report<sup>10</sup> (2006) worden een viertal belangrijke evoluties in e-learning weergegeven. Op de eerste plaats stelt het rapport dat de dynamische aanmaak van kennis en *social computing*-toepassingen en -processen steeds meer wijd verspreid en geaccepteerd worden. Allerlei toepassingen om vanop afstand samen te kunnen werken? komen steeds meer frequent voor en deze worden daarenboven makkelijker om te gebruiken. Een voorbeeld hiervan zijn de *wiki's*, met Wikipedia als het bekendste project. Daarnaast ondersteunt deze trend ook de evolutie naar *personal broadcasting*.

“Mobiele en persoonlijke technologie wordt steeds meer beschouwd als een afleveringsplatform voor diverse diensten”, is een volgende ontwikkeling die in The 2006 Horizon Report aangehaald wordt. Nagenoeg iedereen beschikt tegenwoordig over een mobiele telefoon en bovendien wordt deze gsm meer en een meer een manusje-van-alles. Vandaar dat het eigenlijk niet meer dan logisch is, dat men e-learningtoepassing ook op de mobiele telefoon beschikbaar maakt.

---

<sup>10</sup> The 2006 Horizon Report is de derde editie van een onderzoek van The New Media Consortium en EDUCAUSE Learning initiative. Dit onderzoek tracht de nieuwe technologieën die –naar alle waarschijnlijkheid– een aanzienlijke invloed zullen hebben op het lesgeven, leren of creative expressie in het hoger onderwijs te identificeren.

Als derde belangrijke tendens stelt het onderzoek, dat consumenten alsmaar meer gepersonaliseerde diensten, toepassingen en ervaringen verwachten. Bovendien wensen gebruikers een ongelimiteerde toegang tot verschillende media, kennis, informatie en leren.

Tot slot stelde men vast dat samenwerking meer en meer als kritisch beschouwd wordt binnen educatieve activiteiten.

#### **2.4.2. Web 2.0**

Web 2.0 is eigenlijk een verzamelnaam voor een nieuwe generatie van toepassingen die online samenwerken mogelijk maken (Clarke, 2007). In relatie tot e-learning worden deze nieuwe applicaties ook wel samengevat onder de noemer E-Learning 2.0. Over de correctheid van de benaming kan gediscussieerd worden, maar feit is wel dat Web 2.0 projecten beslaat die een nieuwe weg inslaan. Met sociale software en folksonomieën worden twee van deze nieuwe generatie toepassingen behandeld.

##### **2.4.2.1 Sociale software**

Alexander (2006) duidt de sociale software aan als een belangrijke component van de Web 2.0-evolutie. Blogs, wiki's, videoblogs, ... worden immers alle benoemd als sociale software. The 2006 Horizon Report definieert sociale software als de toepassing van computertechnologie om samenwerking en groepsprojecten te faciliteren. Door sociale software zijn spontaan online gemeenschappen met mensen die dezelfde interesses delen, ontstaan.

Vele studenten maakten in hun privé-leven reeds gebruik van deze sociale software. De talrijke persoonlijke blogs zijn hier het beste voorbeeld van. Maar steeds meer studenten gebruiken de sociale toepassingen nu ook voor hun studie (The Horizon Report 2006, 2006). Voor het (hoger) onderwijs maken deze sociale toepassingen het mogelijk om op een effectievere manier kennis te generen en te delen. Daarnaast wordt het veel eenvoudiger om online aan projecten te werken.

Een kenmerk van deze nieuwe toepassingen is, dat zij fundamenteel niet opgebouwd worden uit webpagina's, zoals bij de klassiek websites, maar eerder uit '*microcontent*' (Alexander, 2006). Blogs zijn namelijk een opeenvolging van berichten en wiki's zijn teksten die geschreven en bewerkt zijn door verschillende personen. De sociale toepassingen zijn dus alle opgebouwd uit kleine elementen. Deze micro-elementen kun dan weer worden bewaard, geciteerd, gekopieerd, ... in andere projecten. Strikt genomen kan men e-mails, forumberichten en dergelijke ook als microcontent beschouwen. Maar de inhoud van de Web 2.0-generatie wordt nog sterker door gebruikers samen gecreëerd en is bovendien toegankelijk voor iedereen. Deze openheid wordt door Alexander (2006) gezien als een elementair kenmerk voor Web 2.0-toepassingen.

#### **2.4.2.2 Folksonomieën**

Een tweede belangrijke component van Web 2.0 zijn de *folksonomies*. De grotere openheid en vrijheid die er heerst bij het gebruik van blogs en dergelijke, maakt dat de traditionele hiërarchische structuur van de metadata moeilijker te hanteren valt. Folksonomieën zijn hier een antwoord op. Wikipedia<sup>11</sup> (2007) geeft er de volgende definitie voor: “*een vorm van ordening, op basis van een consensus, van data door het 'volk'*”. Het komt er op neer, dat de internetgebruikers hun *blogpostings*, websites, foto's, filmpjes, ... voorzien van eigen trefwoorden. Volgens The Horizon Report (2006) zouden deze folksonomieën de wijze waarop wij categoriseren en zoeken, wel eens drastisch kunnen veranderen.

*Del.icio.us*<sup>12</sup> is een voorbeeld van een website die eigenlijk puur op folksonomieën steunt. Del.icio.us is een website voor het zogenaamde *social bookmarking*. Gebruikers kunnen, na registratie, hyperlinks naar websites toevoegen. Deze links worden dan voorzien van een korte beschrijving en van zelf gekozen trefwoorden. Daarnaast kan men dus ook zoeken naar webpagina's, die voorzien zijn van dezelfde tags of trefwoorden.

---

<sup>11</sup> Wikipedia - *Folksonomie*. (2007). Opgevraagd op 15 mei, 2007, via <http://nl.wikipedia.org/wiki/Folksonomie>.

<sup>12</sup> <http://del.icio.us>

## **HOOFDSTUK 3 : Leerobjecten in een e-learningcursus**

### **3.1 Inleiding**

Met de overgang van het oude naar het nieuwe millennium heeft de maatschappij geleidelijk talrijke veranderingen ondergaan. Vandaar dat de *millenial students*, geboren na 1982, grote verschilpunten vertonen met hun voorgangers, de *Generation X*. De millennium-studenten zijn opgegroeid met de nieuwe technologieën en gebruiken deze ook voor zowel hun dagdagelijkse bezigheden als voor hun studie-activiteiten. Daarnaast dienen zij over totaal nieuwe vaardigheden te beschikken; één van de voornaamste is, de bekwaamheid om goed in team te kunnen functioneren. (Cramer, 2007)

Vandaar ook de noodzaak dat het onderwijs nieuwe onderwijsconcepten dient te introduceren, zoals besproken in het vorige hoofdstuk. Leerobjecten zijn een uitstekend middel om te beantwoorden aan de veranderde noden van de millennium-studenten.

In dit hoofdstuk worden de leerobjecten zelf in de eerste plaats onder de loep genomen. Daarna wordt er besproken hoe deze leerobjecten in een e-learningcursus gebruikt worden.

### **3.2 Leerobjecten**

In de ontwikkeling van een e-learningcursus staan de *learning objects* (LO) of leerobjecten centraal. Het fundamentele idee dat schuilgaat achter de leerobjecten is, dat leerstof opgesplitst kan worden in kleine componenten. Deze componenten kunnen vervolgens geannoteerd worden met metadata en opgeslagen worden in een database, zodat ze later opnieuw gebruikt kunnen worden (Churchill, 2005).

Om tot een definitie van een leerobject te komen, worden eerst de gebruikte metaforen voor leerobjecten besproken. Vervolgens wordt met het contextueel model voor leerobjecten het kader verduidelijkt waarin LO acteren. Na de definitie wordt de positie van een leerobject in het volledige leerproces aangegeven.

#### **3.2.1. LEGO of atomen?**

In de literatuur worden LEGO-blokken vaak als metafoor voor leerobjecten gebruikt (Varlamis & Apostolakis, 2006 en Taylor, 2002 in Rubens, 2003). De LEGO-blokken hebben een standaardformaat en kunnen op ontelbare wijzen met elkaar gecombineerd worden tot een groter

geheel: een huis, een brug, enz. Wiley (2000) beweert dat deze vergelijking echter niet helemaal correct is. Voor de speelgoedblokjes geeft hij de volgende eigenschappen:

- Elk LEGO-blokje kan gecombineerd worden met elk ander LEGO-blokje.
- LEGO-blokjes kunnen naar eigen willekeur samen gemonteerd worden.
- LEGO-blokjes zijn zo eenvoudig, dat zelfs kinderen er mee kunnen werken.

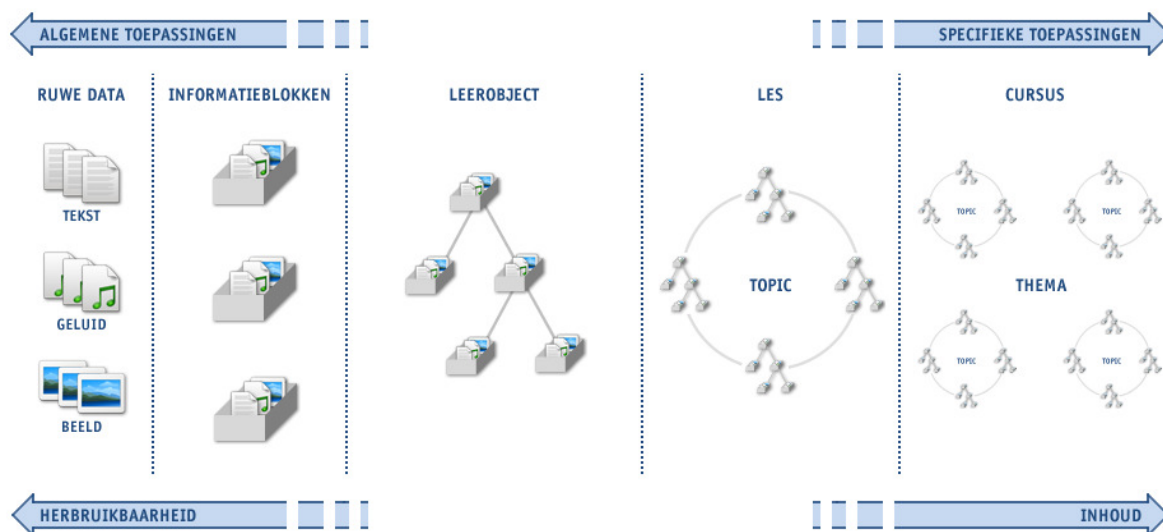
Uiteraard zijn dit alle drie evidente eigenschappen voor LEGO, aangezien het als speelgoed voor kinderen ontworpen is. Maar door een LO te vergelijken met een LEGO-blokje wordt er gesuggereerd, dat deze eigenschappen ook voor leerobjecten gelden, maar hier gaat de vlieger niet helemaal op. Als een alternatieve metafoor voor de leerobjecten introduceert Wiley (2000) het atoom. Een atoom is ook een 'klein' geheel dat met andere atomen gecombineerd kan worden tot een groter geheel, maar er zijn toch enkele essentiële verschillen met LEGO. Ten eerste is het niet mogelijk om elk atoom te combineren met elk ander atoom. Bovendien kunnen atomen enkel in bepaalde structuren samengevoegd worden, afhankelijk van hun eigen innerlijke structuur. Tot slot is, om atomen samen te voegen, enige ervaring en opleiding vereist. Deze eigenschappen gelden ook voor leerobjecten. Leerobjecten kan men ook niet 'zomaar' met elkaar combineren. Om relevant te zijn binnen e-learning, dienen de leerobjecten met elkaar verband te houden en moeten ze passen in het grotere geheel van de e-learningcursus. Dit impliceert dus dat men wel degelijk enige ervaring of domeinkennis nodig heeft, om op een zinvolle manier met leerobjecten te werken en een e-learningpakket te ontwikkelen.

### **3.2.2. Conceptueel model voor leerobjecten**

Vroeger werd een cursus opgedeeld in lessen rond verschillende context-specifieke doelstellingen, maar met het gebruik van leerobjecten wordt deze ruwe opdeling nog verder verfijnd (Masie, 2002). De leerstof wordt uiteindelijk in veel kleinere delen opgesplitst, namelijk zelfstandige leerobjecten. Deze leerobjecten kan men dan apart of in combinatie met andere opnieuw gebruiken.

Varlamis en Apostolakis (2006) ontwikkelden een eigen conceptueel model voor leerobjecten dat visueel weergegeven wordt in figuur 2: conceptueel model voor leerobjecten. Voor de ontwikkeling ervan steunden Varlamis en Apostolakis ondermeer op het model van Masie (2002). Het ontwikkelde model is een hiërarchische structuur van *korrelige* componenten. Een bepaalde component is telkens een verzameling van elementen uit het lagerliggende niveau, die samen aan een objectief beantwoorden. Net als in de klassieke opvattingen beslaan cursussen en lessen de twee hoogste niveaus. Hierbij behandelt een cursus een centraal thema en is opgebouwd uit verschillende lessen, die elk één topic van het centrale thema beslaan. Een les is vervolgens een verzameling van een aantal leerobjecten. Eén leerobject is een combinatie van een aantal informatieblokken, die bijdragen tot één gemeenschappelijk objectief of doel. De informatieblokken

zelf zijn op hun beurt een collage van een aantal gegevens of pure data, zoals tekst en figuren, die bijvoorbeeld een bepaald feit beschrijven.



**Figuur 2: Conceptueel model voor leerobjecten<sup>13</sup>**

De verschillende niveaus kunnen op een continuüm geplaatst worden, gaande van algemeen tot specifiek toepasbaar. Aan het ene uiterste, links op figuur 2, vinden we de pure data, de teksten, geluids- en videofragmenten, enz. Deze gegevens zijn algemeen en kunnen bijgevolg voor diverse doeleinden gebruikt worden. De informatieve waarde van deze elementen is eerder laag, aangezien er geen enkele context aanwezig is. Hoe verder we ons naar het andere uiteinde verplaatsen, des te meer context krijgen de componenten. De componenten worden aangepast aan het gebruik in specifieke situaties. Bijgevolg bevat een cursus veel inhoudelijke informatie, maar daalt de mate waarin hergebruik mogelijk is.

Zoals uit figuur 2 duidelijk blijkt, nemen de leerobjecten in dit model een centrale plaats in. Als verzameling van een aantal informatieblokken vormen zij de bouwstenen voor de lessen. Door hun centrale positie bezitten de leerobjecten een optimale mix tussen inhoud en herbruikbaarheid. Aangezien enerzijds de basisblokken, namelijk de pure gegevens, reeds in enigszins beperkte context geplaatst werden, bezitten zij over voldoende inhoud om waardevol te zijn. Anderzijds is de mate van aanpassing nog beperkt, zodat hergebruik nog in voldoende mate mogelijk is.

De grenzen tussen de verschillende niveaus in het model zijn niet altijd even strikt, vandaar dat deze ook in een stippellijn weergegeven zijn. Een leerobject zou eigenlijk ook uit slechts één informatieblok kunnen bestaan of aan het andere uiteinde zou het ook een combinatie kunnen zijn

<sup>13</sup> Deze figuur is een eigen bewerking van 'Figure 2. Content Object Model' van Varlamis en Apostolakis (2006, p. 63).



van verschillende groepen van informatieblokken. De opdeling in niveaus, gehanteerd in het model, is maar suggestief; desgewenst kan men nog een extra niveau toevoegen.

### 3.2.3. Definitie van leerobjecten

Zoals het model van leerobjecten illustreerde, zijn de grenzen van een leerobject niet strikt afgebakend. Er is dus ruimte voor interpretatie, met het gevolg dat er in de literatuur ook verschillende definities circuleren. We starten het overzicht met een begripsomschrijving voor een leerobject van Rubens (2003):

*"Afgeronde op zichzelf staande hoeveelheid leerstof die zelfstandig toegevoegde waarde heeft voor het leerproces, binnen maximaal 20 minuten kan worden doorlopen en die hergebruikt kan worden ten behoeve van meerdere leertrajecten."*

In deze definitie van Rubens komen twee essentiële kenmerken van leerobjecten aan het licht. Enerzijds betreft het een zelfstandig geheel en anderzijds kan het opnieuw gebruikt worden. De kenmerken zelfstandigheid en herbruikbaarheid komen eveneens terug in de definitie van de *Learning Object Metadata* (LOM) standaard<sup>14</sup>:

*"Learning Objects are defined here as any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning."*

Opvallend is hier, dat de LOM standaard niet-digitale elementen niet uitsluit van de definitie van leerobjecten. Tot niet-digitale entiteiten behoren onder meer mensen, organisaties, gebeurtenissen, ... De LOM standaard hanteert dus een zeer brede definitie. In de literatuur wordt echter algemeen aangenomen dat leerobjecten wel van digitale aard zijn. Varlamis en Apostolakis (2006) vermelden deze 'digitale' voorwaarde uitdrukkelijk in hun benadering van een leerobject:

*"A digital part of a course ranging in size and complexity from a single graphic to an entire course itself."*

Varlamis en Apostolakis geven aan, dat een leerobject kan variëren in grootte en complexiteit. Dit sluit aan bij de eerder vermelde vaststelling, dat een leerobject niet duidelijk afgebakend is. Aangezien de complexiteit van een leerobject kan verschillen, wijst dit er op, dat leerobjecten op diverse doelen gericht zijn. Cramer (2007) neemt het doel van een leerobject als invalshoek om een beschrijving te geven:

*"Learning Objects are instructional materials found on the internet that can be used to illustrate, support, supplement, or assess student learning."*

Cramer onderscheidt een viertal doelstellingen, Shepherd (2000) echter maakt een indeling van leerobjecten naargelang hun doel in drie categorieën, weergegeven in figuur 3. Een eerste categorie vormen de geïntegreerde leerobjecten. Hiertoe behoren ondermeer de mini-colleges, waarbij verschillende elementen samengebracht werden tot één afgerond geheel. Een tweede groep leerobjecten is van een informatieve aard. Definities, artikels, voorbeelden, ... maken deel uit van

---

<sup>14</sup> De Learning Object Metadata (LOM) standaard is een uitgave van IEEE Learning Technology Standards Committee (LTSC). (IEEE is de afkorting van Institute of Electrical and Electronics Engineers, Inc.)

deze categorie. De laatste opdeling bevat leerobjecten die gericht zijn op toepassing en controle van de onderwezen materie. Hierin vinden we gevalstudies, simulaties, oefeningen en testen terug.

| Geïntegreerd                                                                                                                               | Informatief                                                                                                                                                                                                                                   | Toepassing en controle                                                                                                                                                     |
|--------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"><li>• Mini-werkcolleges</li><li>• Mini-gevalstudies, simulaties, ... met bijkomende informatie</li></ul> | <ul style="list-style-type: none"><li>• Overzicht/samenvatting</li><li>• Definities</li><li>• Demonstratie/model</li><li>• Voorbeeld</li><li>• Gevalstudie</li><li>• Paper/artikel</li><li>• Beslissingsondersteunende hulpmiddelen</li></ul> | <ul style="list-style-type: none"><li>• Problemen/gevalstudies</li><li>• Spellen/simulaties</li><li>• Oefeningen</li><li>• Herhalingsoefeningen</li><li>• Testen</li></ul> |

**Figuur 3: Indeling leerobjecten**

De definitie van leerobjecten volgens Shepherd (2000) is de volgende:

*"A learning object is a small, reusable digital component that can be selectively applied – alone or in combination – by computer software, learning facilitators or learners themselves, to meet individual needs for learning or performance support."*

De twee eerder genoemde eigenschappen, namelijk zelfstandigheid en herbruikbaarheid, vinden we ook hier terug. Daarnaast geeft Shepherd aan, dat een leerobject niet eenduidig op het overbrengen van leerstof gericht kan zijn, maar eveneens een ondersteunende functie aannemen kan. Hier wordt, net als in de definitie van e-learning van Rosenberg (2001), een ruime blik aangenomen. In de definiëring van een 'beter' leerobject, verwijst Churchill (2005) ook naar het ondersteunde karakter:

*"A better learning object may be an interactive, visual representation and a mediating tool in a learning activity."*

In de geciteerde definities komen een vijftal essentiële kenmerken van een leerobject naar voren<sup>15</sup>:

- zelfstandig geheel
- herbruikbaarheid
- digitaal
- variërend in grootte
- ondersteuning van het leerproces

Hierbij valt op te merken dat het laatste kenmerk, de ondersteuning van het leerproces, ruim dient te worden geïnterpreteerd.

---

<sup>15</sup> De volgorde van de kenmerken in de opsomming is willekeurig en houdt dus geen enkele rangorde van belangrijkheid in.

Deze kenmerken vinden we alle terug in de definitie voor leerobjecten van Schreurs en Moreau (2006):

*"A learning object (LO) is an independent content component that can function as the learning content of a course module. It can be defined as any digital content resource that supports learning, that can be reused and that can be delivered on demand across the network, be it large or small."*

### **3.2.4. Leerobject als psychologisch hulpmiddel in het leerproces**

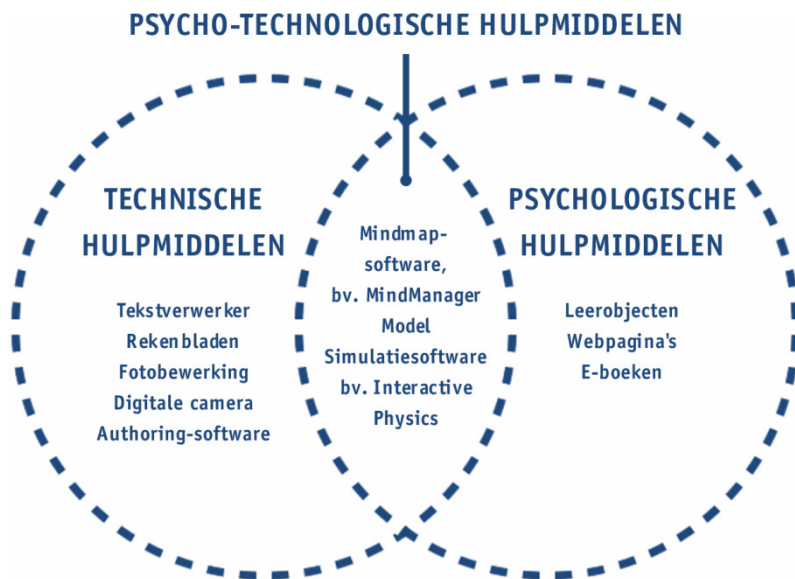
Tijdens een leerproces maakt men gebruik van een aantal hulpmiddelen. Tot deze hulpmiddelen behoort ook de technologie. Met het gebruik van technologische toepassingen kunnen mensen de eigen productiviteit verhogen. Een arbeider die een graafmachine ter beschikking heeft, haalt bijvoorbeeld een veel grotere productiviteit dan een collega die met de traditionele spade te werk mag gaan. In dit voorbeeldje verbetert de technologie de fysische capaciteiten van de mens. Daarnaast kan de technologie ook een uitbreiding zijn van de intellectuele capaciteiten en bovendien een verruiming en verandering van de sociaal-culturele omgeving teweegbrengen (Churchill, 2005).

In het kader van leeractiviteiten verdeelt Churchill (2005) de hulpmiddelen in eerste instantie in twee categorieën, namelijk technische en psychologische hulpmiddelen. De technische hulpmiddelen beïnvloeden en wijzigen de eigenschappen van het voorwerp van een leeractiviteit. Tekstverwerker, Rekenbladen, digitale camera, ... zijn voorbeelden van technische werktuigen. De tweede categorie bestaat uit de psychologische tools. Deze categorie richt zich naar ons eigen bewustzijn, waar de interactie tussen verschillende mentale functies, zoals redeneren, het geheugen, *inner speech*, ... , plaatsvindt. Als voorbeeld hiervan haalt Churchill een interactieve visuele toepassing aan, voor het bestuderen van een rechthoekige driehoek. De gebruiker kan bijvoorbeeld de lengte van de hypotenusa wijzigen en vervolgens de invloed ervan op de lengte van de rechthoekszijden bekijken.

Deze twee categorieën zijn visueel weergegeven in figuur 4: classificatie van hulpmiddelen in leeractiviteiten. De figuur toont, dat de twee beschreven categorieën elkaar overlappen en een derde categorie van hulpmiddelen vormen, de psycho-technologische tools. Deze worden ook weleens beschreven als cognitieve hulpmiddelen, *mind tools* of *technologies of mind*. De psycho-technologische hulpmiddelen bieden de mogelijkheid voor lerenden om te experimenteren en artefacten te construeren en bovendien binnen deze activiteit tegelijk te leren.

Als we deze indeling van Churchill concreet in een e-learningomgeving toepassen, dan behoren de leerobjecten tot de psychologische hulpmiddelen. Authoring-software, waarmee leerobjecten ontwikkeld kunnen worden, worden onderverdeeld bij de technische tools. Technologie om

bijvoorbeeld mindmappen te creëren, bevindt zich in de overlappende groep van psycho-technologische hulpmiddelen.



**Figuur 4: Classificatie van hulpmiddelen in leeractiviteiten<sup>16</sup>**

### 3.3 Ontwikkeling van een e-learningcursus

Een e-learningcursus is niet gelijk aan simpelweg online geplaatste leerstof, integendeel zelfs: het ontwikkelen van een degelijke e-learningcursus gebeurt aan de hand van een uitgekiend proces.



**Figuur 5: Levenscyclus van een e-learningcursus**

<sup>16</sup> Deze figuur is een eigen bewerking van 'Figure 1. Mediating tools for learning activity' van Churchill (2005, p. 338).

Varlamis en Apostolakis (2006) ontwikkelden een levenscyclus voor de ontwikkeling van een e-learningcursus, zie figuur 5: levenscyclus van een e-learningcursus.

In de voorgestelde cyclus van Varlamis en Apostolakis doorloopt een e-learningcursus vier fasen. In de ontwerpfase moeten de leerdoelstellingen en nodige vereisten voor de leerlingen gespecificeerd worden. De productie-fase omvat de creatie van de inhoudsblokken, het samenvoegen en ten slotte de 'verpakking' ervan. Tijdens de derde fase kan de e-learningcursus in gebruik genomen worden. Het laatste gedeelte van de levenscyclus vormt de beoordeling. Dit slaat enerzijds op de beoordeling van de gebruikers en anderzijds op de beoordeling van het leerproces in zijn geheel.

In de verdere beschrijving van de e-learningcursus zal er voornamelijk gefocust worden op de tweede fase uit de levenscyclus, namelijk de productie-fase. Deze productie-fase kan opgesplitst worden in twee essentiële fasen. Enerzijds de decompositie-fase, waarbij de inhoudselementen gegenereerd worden en anderzijds de creatie van e-learningcursus aan de hand van een scenario-model. Beide fasen komen in de volgende paragrafen achtereenvolgens aan bod.

### **3.3.1.     Decompositie**

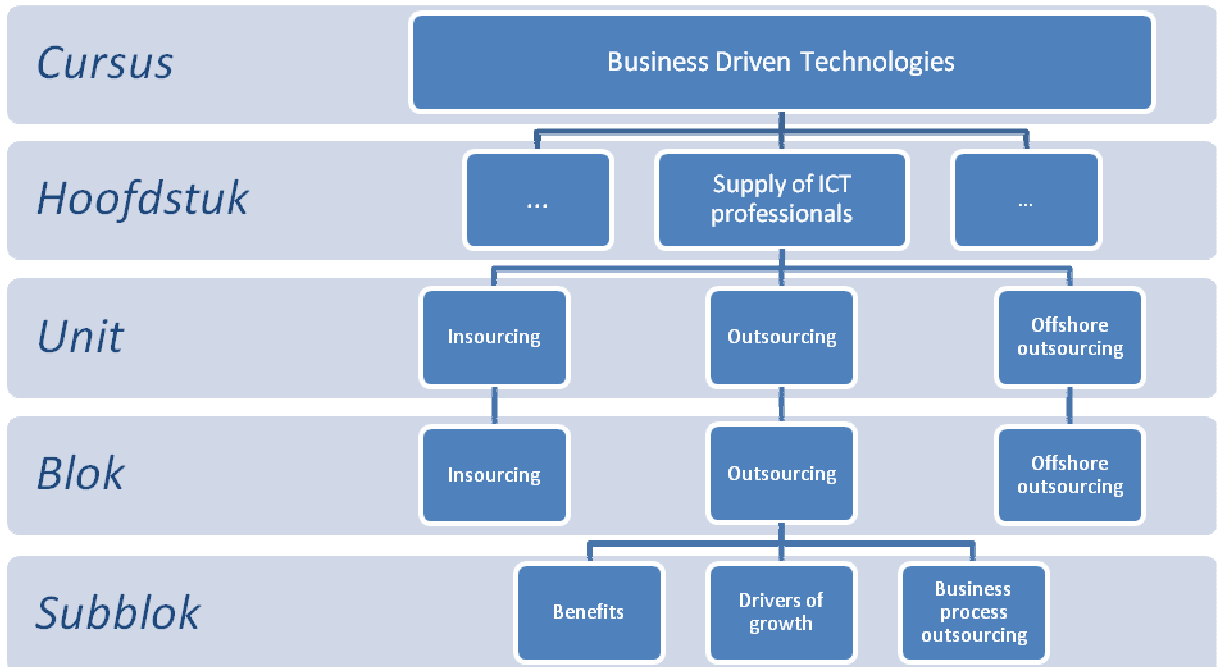
Een eerste belangrijke stap in het ontwikkelingsproces van een e-learningcursus is de decompositie van het brondocument. Reigeluth en Nelson (1997, in Wiley, 2000) noteerden het volgende:

*"When teachers first gain access to instructional materials, they often break the materials down into their constituent parts. Then they reassemble these parts in ways that support their individual instructional goals."*

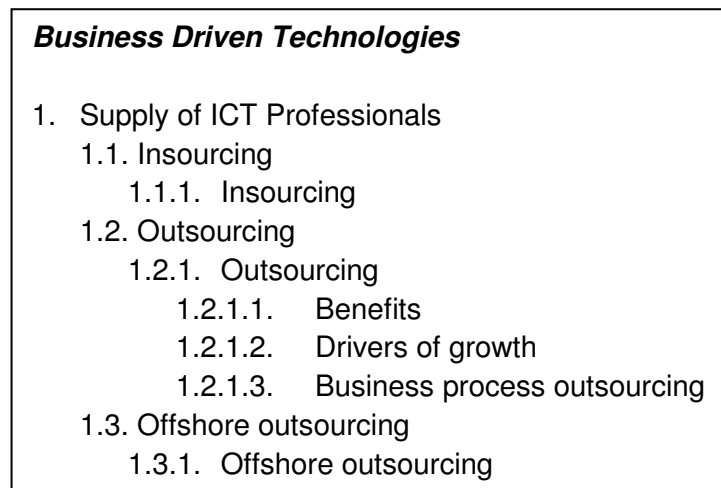
Het opdelen van leerstof om het nadien opnieuw –op een andere manier– samen te voegen is dus eigenlijk niets nieuws. In het ontwikkelen van een e-learningcursus is dit eveneens een eerste en essentiële stap. Hier geldt immers het 'GIGO'-principe, oftewel 'Garbage In, Garbage Out'. De decompositie dient op een logische en verantwoorde manier te gebeuren. Men kan met slordige bouwstenen namelijk geen degelijk eindresultaat verwachten.

Schreurs en Moreau (2006) construeren tijdens deze decompositie eerst een boomstructuur van inhoudsblokken. Zoals in een boek wordt de e-learningcursus eerst opgesplitst in hoofdstukken. Als het hoofdstuk uit meerdere topics bestaat, die afzonderlijk bestudeerd kunnen worden, dan wordt het opgesplitst in twee of meer onafhankelijke units. Deze units worden vervolgens in een bepaalde volgorde geplaatst, zodat ze als één logische inhoudsmodule aangeboden kunnen worden. Een unit zelf is opgebouwd uit één of meerdere blokken, die in een boomstructuur geplaatst worden. Een blok komt overeen met een klein gedeelte van de leerstof. De volledige materie wordt opgesplitst in een aantal tekstdelen en elk tekstdeel wordt gekoppeld aan het overeenkomstige inhoudsblok.

Indien nodig kunnen aan een blok subblokken gelinkt worden. Eén subblok is dan een verdieping van de materie in het bovenstaande blok. Tussen dat blok en de voorkomende subblokken gelden er dan relaties als 'bestaat uit', 'is een voorbeeld van', 'kan verklaard worden door', enz.



**Figuur 7: Decompositie**



**Figuur 6: Boomstructuur**

In figuur 7 is deze opdeling met een concreet voorbeeld uitgewerkt. Uit handboek '*Business Driven Technologies*'<sup>17</sup> werd het hoofdstuk '*Supply of ICT professionals*' gelicht. Aan de hand van de figuur kan men duidelijk vaststellen, dat het hoofdstuk opgedeeld werd in drie units: '*insourcing*',

<sup>17</sup> Haag, S., Baltzan, P., & Philips, A. (2006) *Business Driven Technologies*. New York: McGraw-Hill/Irwin.

'outsourcing' en 'offshore outsourcing'. In dit geval is aan elke unit ook slechts één blok gekoppeld. Het blok 'outsourcing' is hier wel nog verder opgesplitst in een drietal subblokken. Het eindresultaat van deze eerste stap in het decompositieproces is dus een boomstructuur, weergegeven in figuur 6. Tot nu toe is er in de leerstof een duidelijke structuur aangebracht.

In een tweede stap van de decompositiefase kan de leerstof behorende tot één blok of subblok verder uiteengehaald worden. Deze materie bestaat uit het originele gedeelte van de tekst, figuren, tabellen, vragen en antwoorden, tests, links naar webpagina's, ... Al deze elementen worden afgezonderd en vormen also een groep van kleine elementen behorende tot één blok of subblok. Deze 'kleine' presentatie-elementen worden door Schreurs en Moreau (2006) *atomic learning objects* genoemd. Ook het blok als geheel wordt door hun bestempeld als een atomisch leerobject. Daarnaast is het mogelijk om externe atomische leerobjecten toe te voegen, zoals video- en geluidsfragmenten.

Na deze tweede stap is de decompositie voltooid. In eerste instantie werd de totale materie gestructureerd en opgedeeld in inhoudsblokken. Vervolgens werden de blokken nog verder opgesplitst in hun afzonderlijke elementen. Door deze verschillende componenten op een uitgekende wijze terug samen te nemen, wordt er een e-learningcursus ontwikkeld.

### **3.3.2. Scenario-model**

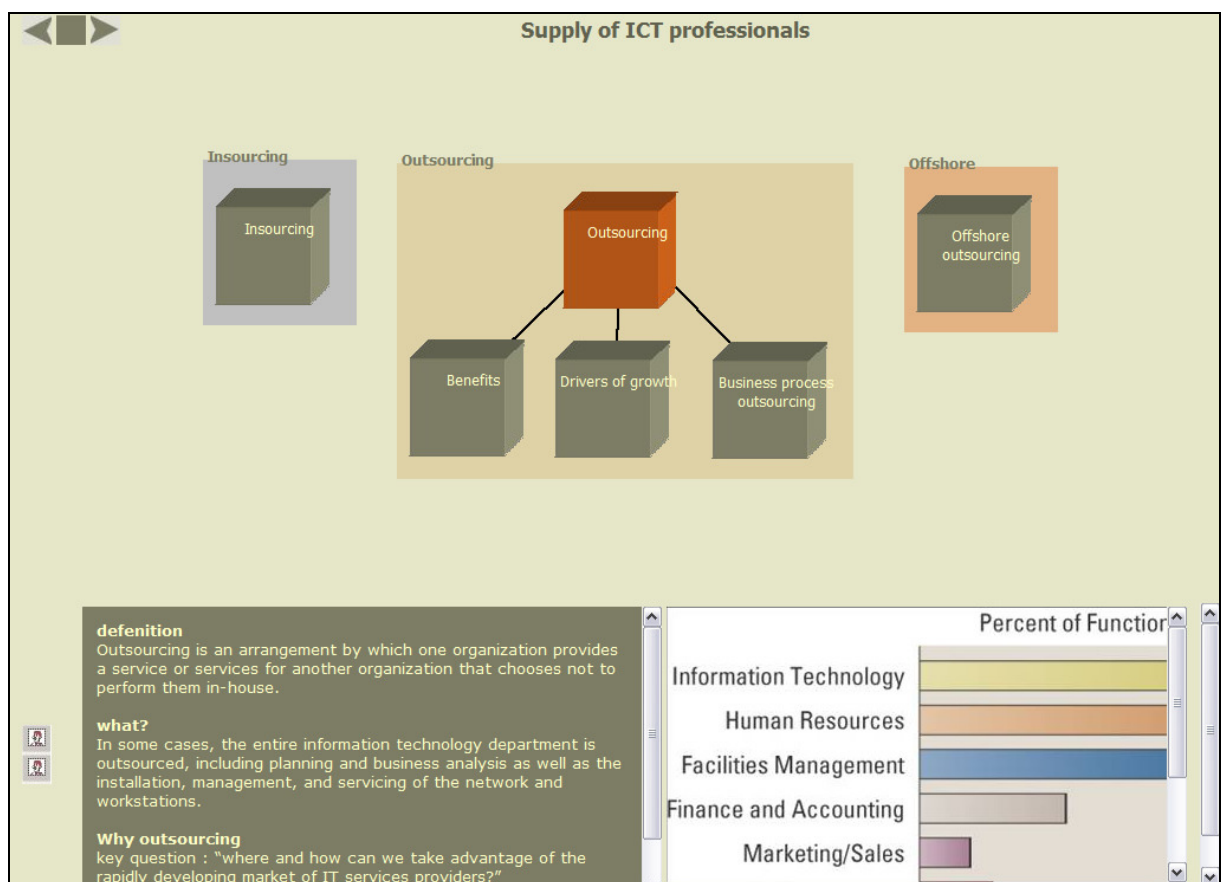
Na de decompositie-fase is de leerstof opgesplitst in een aantal blokken. De volgende fase bestaat er nu uit, om de blokken volgens een gekozen scenario-model te verenigen om een e-learningcursus te bekomen. Voorbeelden van verschillende scenario's zijn *EDUMAP*, *story-telling*, *tree-structured HTML-pages*, ... (Schreurs & Moreau, 2006) Dit betekent dus, dat op basis van dezelfde decompositie, meerdere cursussen volgens verschillende scenario's ontwikkeld kunnen worden.

De basis binnen elk scenario-model zijn de leerobjecten. Afhankelijk van het gekozen scenario heeft een leerobject een specifieke compositie van de blokken onderling en een bepaalde architectuur binnen de blokken. Een blok bestaat enerzijds uit een component met de inhoud, in casu het originele gedeelte van de tekst. Anderzijds wordt dit blok aangevuld met de presentatie-elementen, die deel uitmaken van de presentatie-laag. Deze presentatie-componenten zijn ondersteunende, interactieve en animerende elementen, zoals figuren, video's, animaties, ... (Schreurs & Moreau, 2006) De diverse presentatie-elementen kunnen in één of meerdere scenario's voorkomen. Elk scenario heeft dus een eventueel verschillende compositie van de blokken in een leerobject en maakt elk op zijn manier gebruik van de atomische leerobjecten.

Een leerobject is op zichzelf nog geen cursus, daarvoor worden verscheidene leerobjecten eerst gegroepeerd in hoofdstukken. Een hoofdstuk kan uit één of meerdere leerobjecten bestaan. Enkele hoofdstukken worden vervolgens op hun beurt 'verpakt' in een volwaardige cursus.

Eén van de mogelijke scenario's is *EDUMAP*. In dit scenario wordt de cursus op een grafische manier voorgesteld en in een aantal op elkaar volgende stappen besproken. De informatieblokken waaruit een leerobject opgebouwd is, worden hier voorgesteld door grafische blokken.

Figuur 8: EDUMAP: Supply of ICT professionals is een voorbeeld van een EDUMAP op basis van de decompositie-structuur uit figuur 6. Als we figuur 8 vergelijken met figuur 7 uit de decompositie-fase: het hoofdstuk is opgebouwd door het samenvoegen van de leerobjecten overeenkomstig met de drie units. Deze zijn 'Insourcing', 'Outsourcing' en 'Offshore outsourcing' en zijn duidelijk in de mindmap te onderscheiden door de aparte titels en de verschillende achtergrondkleur. Vervolgens merken we, dat elk blok uit de decompositie-fase ook grafisch in blokvorm getoond wordt. Eventueel worden de blokken nog verder opgesplitst in subblokken zoals bij 'Outsourcing'.



**Figuur 8: EDUMAP: Supply of ICT professionals**



In de mindmap is het mogelijk om de vooropgestelde volgorde van de cursus te doorlopen door de pijltjes rechtsbovenaan de mindmap te gebruiken. Maar de gebruiker kan, indien hij dat wenst, zijn eigen weg gaan door eenvoudigweg op het gewenste grafische blokje te klikken.

Afhankelijk van het geselecteerde blok worden onderaan de bijbehorende atomische leerobjecten getoond. In het voorbeeld is 'Outsourcing' geselecteerd en zien we links de volledige brontekst. Rechts wordt één van de twee beschikbare figuren getoond.

## HOOFDSTUK 4 : Text mining en Natural Language Processing

### 4.1 Inleiding

Geschreven tekst is de meest natuurlijke vorm om informatie uit te wisselen of te bewaren. Karanika en Theodoulidos (2003) halen een rapport<sup>18</sup> aan waarin geargumenteed wordt dat tot 80% van de informatie aanwezig in organisaties in tekstvorm is, zoals rapporten en e-mails. Deze grote hoeveelheden overstijgen de beperkte capaciteit die de mens verwerken kan. Text mining en aanverwante technologieën trachten hiervoor een oplossing te bieden.

Het doel van *text mining*, ook wel *document mining* genoemd, is om door analyse van grote collecties van ongestructureerde documenten nieuwe en interessante patronen, met andere woorden nieuwe kennis, te onthullen. Met text mining is het dus mogelijk om uit deze gigantische berg aan informatie, die met het internet bovendien zeer gemakkelijk toegankelijk is, nieuwe kennis te halen. Gezien de enorme hoeveelheden aan beschikbare tekst, bezit document mining dus een reusachtig potentieel.

Maar er dient wel te worden opgemerkt dat het niet zo eenvoudig is als lijkt. Deze teksten zijn immers bedoeld om door mensen gelezen te worden en bijgevolg niet door computers. Het ontbreken van enige formele structuur en de afwezigheid van metadata maken, dat er niet meteen een standaardmiddel voorhanden is om deze informatieberg te onderzoeken, te ondervragen en te analyseren.

### 4.2 Text mining versus Natural Language Processing

In de literatuur circuleren verschillende alternatieven voor *text mining* (TM), zoals *document mining* en *text data mining*. Daarnaast duiken er voor de term *natural language processing* (NLP) eveneens een aantal varianten op, onder meer: *computational linguistics* en *language processing technology*. Bovendien worden technieken en methoden uit beide domeinen vaker met elkaar gecombineerd. Het gevolg is, dat de termen in bronnen kriskras door elkaar en daarenboven te pas en te onpas gebruikt worden. Alvorens dit hoofdstuk verder uit te diepen, is het dan ook aangewezen eens even stil te staan bij de begrippen *text mining* en *natural language processing*.

---

<sup>18</sup> Merrill Lynch, e-Business Analytics, In-dept Report, 20 November 2000.

#### 4.2.1. Natural Language Processing

Kao en Poteet (2005) stellen dat Natural Language Processing al een geschiedenis van enkele tientallen jaren heeft. De ontwikkelde technieken starten telkens vanuit een taalkundige uitvalsbasis. In een eerste fase wordt de tekst syntactisch ontleed, schrijven Kao en Poteet. Hierbij kan er eventueel gebruik gemaakt worden van informatie uit een woordenboek en de formele grammatica-regels van een taal. Vervolgens wordt het bekomen resultaat semantisch geanalyseerd, zodat het systeem als het ware de tekst begrijpt. Ghahremani (2006) geeft een omschrijving voor NLP in deze zin:

*"A method of extracting meaning from printed words that now allows the software to 'understand' complex phrases about 80 percent of the time."*

De definiëring van Ghahremani is nogal beperkt. Enerzijds wordt de input van NLP gelimiteerd tot geschreven tekst, terwijl er eveneens technologieën zijn waarbij op basis van spraak gewerkt wordt. Anderzijds wijst zijn definitie op éénrichtingsverkeer: enkel van een tekst naar een formele structuur die de software interpreteren kan. Maar voor de omgekeerde weg, van gestructureerde informatie uit bijvoorbeeld een database naar 'normale' tekst (of spraak), worden er binnen het NLP-domein eveneens methoden ontwikkeld. Het artikel over NLP in Wikipedia<sup>19</sup> duidt deze twee richtingen respectievelijk aan met *natural language understanding* en *natural language generation*.

Typerend voor NLP is volgens Kao en Pateet (2005) dat de focus slechts op één document ligt en niet op een collectie van documenten. Daarnaast geven zij voor typische NLP-technieken onder meer de volgende voorbeelden:

- *word stemming*: het herleiden van een woord naar zijn stamvorm, m.a.w. het verwijderen van aanwezige prefixen en suffixen;
- *lemmatization*: een verbogen woord omzetten naar zijn stamvorm;
- *part-of-speech (POS) tagging*: een techniek om de individuele woordgroepen in hun respectievelijke lexicale klassen (zelfstandig naamwoord, werkwoord, voorzetsel, enz.) in te delen;
- *anaphora resolution*: het oplossen van referentiële afhankelijkheid, m.a.w. de betekenis van woorden achterhalen die voor hun interpretatie afhankelijk zijn van een andere referentiële uitdrukking; bijvoorbeeld: de betekenis van persoonlijke voornaamwoorden als 'hij' en 'zij' of woorden als 'de student' in een zin achterhalen (de Swart, z.d.).

In het algemeen kan *natural language processing* dus gedefinieerd worden als het domein, dat de problemen van het automatisch begrijpen en genereren van de menselijke taal bestudeert.

---

<sup>19</sup> *Natural Language Processing*. (2007). Opgevraagd op 26 juni, 2007, via [http://en.wikipedia.org/wiki/Natural\\_language\\_processing](http://en.wikipedia.org/wiki/Natural_language_processing).

#### 4.2.2. Text mining

Kao en Pateet (2005) zien *text mining* als een recente ontwikkeling, in tegenstelling tot NLP. De technieken en methoden die binnen TM worden gehanteerd, werden overgenomen uit het ruime kennisterrein van de statistiek. De omschrijving van Ghahremani (2006) voor TM duidt uitdrukkelijk op de verwantschap met het aan statistiek-gerelateerde vakgebied 'data mining':

*"Text mining, a type of data-mining technology that combs through text and gives it structure so it can be analyzed."*

Text mining-technieken transformeren het originele document naar een specifieke structuur, zodat het mogelijk wordt om technieken uit data mining erop los te laten. Op deze manier tracht men nieuwe of verborgen patronen of informatie te onthullen.

De bedoeling van text mining is dus niet om een brondocument te interpreteren, maar het komt er eerder op aan om patronen in een grote collectie aan documenten te ontdekken. Hiermee komen meteen enkele belangrijke verschilpunten met NLP aan het licht. Ten eerste is het hoofddoel dus, om informatie uit de brondocumenten te halen. De algemene definitie van Spasic, Ananiadou, McNaught en Kumar (2005) voor text mining geeft dit hoofddoel goed weer:

*"Automatic discovery of previously unknown information by extracting information from text."*

Ten tweede focust TM uitsluitend op geschreven tekst, terwijl NLP ook andere tekstvormen als spraak behandelt. Daarenboven zal TM eerder werken met een aantal teksten als input, in tegenstelling tot NLP, waarbij men zich in vele gevallen tot slechts één bron beperkt.

Een aantal technieken die binnen text mining thuishoren zijn:

- *automatic text classification*: een tekst automatisch tot één of meerdere voorgedefinieerde categorieën toewijzen;
- *text clustering*: groeperen van teksten volgens eenzelfde onderwerp;
- *automatic summarization*: met behulp van software genereren van samenvattingen;
- *concept/entity extraction*: herkennen van concepten en entiteiten in teksten.

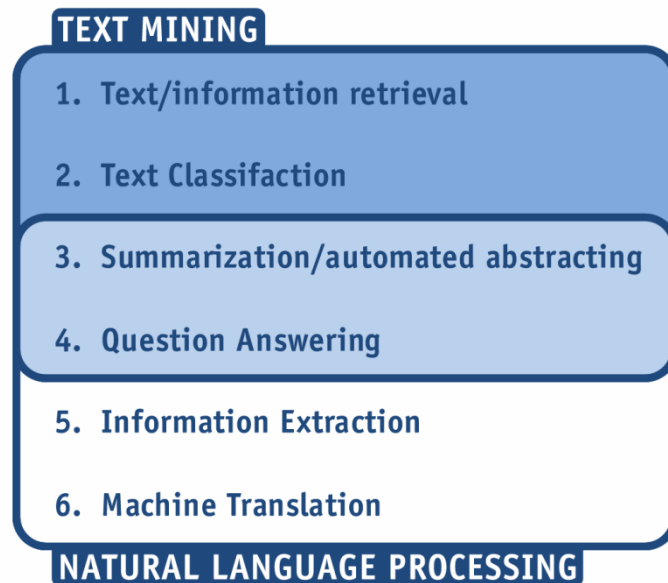
Text mining kan dus gedefinieerd worden als de technologie, om uit grote hoeveelheden geschreven tekst informatie te halen, met in eerste instantie gebruik te maken van aan data mining ontleende technieken.

#### 4.2.3. Overlapping tussen text mining en natural language processing

Uit de vorige paragrafen blijkt reeds, dat de werkgebieden van text mining en natural language processing zeer dicht bij elkaar in de buurt liggen of elkaar zelfs overlappen, maar beide domeinen benaderen dit elk vanuit een andere invalshoek. Het is dan ook niet verwonderlijk, dat men de

laatste tijd vaststelt, dat in concrete toepassingen meer en meer technieken uit TM en NLP gecombineerd worden. Dixon (1997) omschrijft *document mining*, een synoniem voor text mining, eveneens als een combinatie van technieken uit *information extraction*, *information retrieval*, *natural language processing* en *document summarization* met methoden uit *data mining*.

Het overkoepelende domein van de *Language Processing Technology* bevat volgens Lewis (Kao & Poteet, 2004) de toepassingsdomeinen opgesomd in figuur 9.



**Figuur 9: Toepassingsdomeinen Language Processing Technology**

Lewis rekent de eerste twee tot het domein van text mining, de laatste twee tot natural language processing en beide domeinen beweren aanspraak te mogen maken op nummers drie en vier in de figuur hierboven. De correctheid van deze bewering wordt hier niet besproken, maar het toont wel aan, dat er tussen TM en NLP een overlapping bestaat. Indien men er in de toekomst in slaagt om technieken uit beide gebieden met toenemende kans op succes te combineren, dan mag men allicht aannemen, dat de mate waarin TM en NLP elkaar overlappen, alleen nog maar zal toenemen.

Karanikas en Theodoulidis (2003) stellen zelfs uitdrukkelijk in hun definiëring van text mining, dat dit een combinatie is van data mining én NLP-algoritmes. In het verdere verloop van dit werk wordt er dan ook geen onderscheid gemaakt tussen text mining en natural language processing. Het is wel zo, dat het vervolg text mining als uitgangspunt zal nemen en daarbij een aantal concepten uit NLP zal betrekken.

### 4.3 Text mining-kader

Text mining kan volgens Spasic et al. (2005) en Karanika en Theodoulidis (2003) verder opgesplitst worden in drie onderdelen:

1. Information Retrieval
2. Information Extraction'
3. Information Mining

In *Information Retrieval* (IR) komt het erop neer, om relevante documenten te localiseren en te verzamelen. In een tweede onderdeel, namelijk *Information Extraction* (IE), bestaat de kerntaak eruit, informatie uit de (geselecteerde) teksten te destilleren. De laatste stap, *Information Mining* (IM) of het *echte* text mining, tracht in de informatie uit de IE-fase bepaalde patronen te ontdekken. Dixon (1997) en El Wakil (z.d.) voegen hier nog een extra stap aan toe: interpretatie. Het spreekt natuurlijk voor zich, dat aan de gegevens uit IE en aan informatie bekomen door gebruik van IM, een nuttige betekenis moet verbonden worden om verder van enig nut te kunnen zijn.

Aangezien voor dit onderzoek *information retrieval* en voornamelijk *information extraction* de belangrijkste onderdelen zijn, worden deze twee hieronder verder besproken. *Information mining* wordt buiten beschouwing gelaten.

#### 4.3.1. Information retrieval

*Information retrieval*, ook wel *text retrieval* of *document retrieval* genoemd, is de eerste stap in het text mining-proces, waarin relevante documenten worden gelocaliseerd en verzameld. Documenten die irrelevant zijn, dienen uit het systeem geweerd te worden. In deze fase vindt er dus een filtering van de bronteksten plaats en enkel de documenten die ter zake doen voor de specifieke taak, worden bewaard.

Leezenberg, Tonningen en Schoonhoven (2005) schrijven dat *information retrieval* gebruikt kan worden om de ad hoc behoefte aan bijkomende informatie met betrekking tot bepaalde leerstof op te lossen. Met behulp van een IR-systeem kan gezocht worden naar documenten en bestanden aanwezig op het intranet of internet, die het leerproces of een projectuitvoering kunnen ondersteunen.

De internetzoekmachine Google is een zeer bekend voorbeeld van een systeem dat gebruikt wordt voor *information retrieval*. Aan de hand van een of meerdere zoektermen wordt door Google een reeks van relevante documenten en links gefilterd uit het enorme aanbod op het internet. Google is

een voorbeeld van een zoekmachine voor algemeen gebruik, daarnaast bestaan er ook talrijke specifieke information retrieval-toepassingen voor gebruik in bepaalde (sub-)domeinen. Een voorbeeld hiervan is PubMed, een zoekmachine ontworpen om specifiek in databases met biomedische publicaties te zoeken (Spasic, Ananiadou, McNaught, & Kumar, 2005).

#### **4.3.2. Information extraction**

Het tweede onderdeel, *information extraction*, heeft tot doel om informatie uit de (gefilterde) teksten te halen. Dit gebeurt door te gaan zoeken naar feiten die in de tekst voorkomen en vervolgens verbanden tussen de gevonden feiten te leggen (Dixon, 1997). De ongestructureerde bronteksten worden met IE getransformeerd tot ondubbelzinnige gegevens in een formele structuur. De resultaten kunnen vervolgens onmiddellijk getoond worden aan de gebruikers of kunnen opgeslagen worden in een database om eventueel later geanalyseerd te worden met text mining technieken (Cunningham, 2005).

Information extraction is dus wezenlijk verschillend van information extraction. IR zoekt en presenteert relevante (volledige) teksten, terwijl IE teksten analyseert en enkel specifieke informatie die voor de gebruiker van belang is, als output genereert (Cunningham, 2005). Gezien bij IE de teksten zelf geanalyseerd worden, is dit een complex proces waarbij, afhankelijk van de toepassing, algemene of domein-specifieke kennis noodzakelijk is. IE vergt dus veel meer 'rekenwerk' van een computer, maar is bij omvangrijke collecties efficiënter, aangezien het enkel de essentiële informatie uit de teksten presenteert. Dit levert een aanzienlijke besparing in tijd op, in vergelijking met IR waarbij de volledige tekst gelezen moet worden. De gestructureerde output van IE maakt, dat het vertalen ervan relatief eenvoudig kan gebeuren. Dit levert een bijkomend voordeel op ten opzichte van IR, waarbij volledige teksten dienen te worden vertaald, wat veel minder vanzelfsprekend is (Cunningham, 2005).

Information extraction is een zeer belangrijk en omvangrijk aspect van text mining. Het gebeurt niet zelden dat information extraction en text mining als hetzelfde bestempeld worden. Dixon (1997) en El Wakil (z.d.) maken binnen IE de volgende opdeling:

1. Fact extraction
2. Fact integration
3. Kennisvoorstelling

*Fact extraction* bestaat eruit, om individuele feiten binnen één document te zoeken. Hiervoor is domeinspecifieke kennis zeer belangrijk. Enkele technieken die hiervoor gebruikt worden, zijn: patroonherkenning, lexicale analyse en syntactische en semantische structurele analyse. Patroonherkenning is het proces waarbij in de eerste plaats gezocht wordt naar vaak voorkomende

expressies. Lexicale analyse is een techniek waarbij de tekst eerst opgedeeld wordt in tokens<sup>20</sup>. Vervolgens wordt getracht op basis van de tokens, de zinnen te herkennen. Tot slot tracht men bij de lexicale analyse aan de woorden en woordgroepen in hun voorkomende context een juiste betekenis toe te kennen. Een analyse van de syntactische en semantische structuur start met aan de woorden of woordgroepen een syntactische betekenis toe te kennen. Op basis van deze syntactische component wordt er aan de hand van een semantische patroonherkenning getracht, hier een betekenis aan te koppelen.

De individuele feiten samenbrengen en uitzoeken hoe deze in een groter geheel passen, is het onderwerp van *fact integration*. Het oplossen van indirecte verwijzingen in een tekst (bijvoorbeeld naar welke persoon "hij" in een paragraaf refereert) is hier een voorbeeld van.

De laatste stap 'kennisvoorstelling' zorgt voor de presentatie van de geëxtraheerde informatie. Deze presentatie kan op verschillende manieren gebeuren: hetzij heel eenvoudig door een tekstuele weergave waarbij een (voorgedefinieerd) formulier ingevuld wordt; ofwel door een, voor de gebruiker vaak veel interessantere, visuele voorstelling.

De indeling gebruikt door Dixon en El Wakil, wordt hier niet verder aangehouden. In de plaats daarvan komen de volgens Cunnigham (2005) vijf belangrijkste types van information extraction aan bod:

1. Named Entity Recognition
2. Coreference resolution
3. Template Element construction
4. Template Relation construction
5. Scenario Template production

Als we deze categorieën van *information extraction* zouden verdelen volgens de classificatie van Dixon en El Wakil, dan zou *Named Entity Recognition* alvast tot groep *fact extraction* behoren. De overige vier vallen eerder onder het label *fact integration*. In de volgende alinea's worden deze vijf categorieën kort besproken.

#### **4.3.2.1    *Named Entity Recognition***

*Named Entity Recognition* (NER) wordt ook wel *entity identification* of *entity extraction* genoemd. NER classificeert de basiselementen in een aantal voorafgedefinieerde categorieën, zoals namen van personen, organisaties, plaatsen, ... maar ook datums, geldbedragen, percentages, ... NER wordt eveneens gebruikt om termen onder te brengen bij de gerelateerde concepten uit een ontologie of taxonomie (Spasic, Ananiadou, McNaught, & Kumar, 2005).

---

<sup>20</sup> Met tokens worden in dit geval eigenlijk woorden of woordgroepen bedoeld. Het zijn de kleinste bouwblokken van een tekst. Een token zelf, is dan opgebouwd uit een aantal karakters.



Bij NER is een ontologie of een taxonomie een onmisbaar onderdeel, vandaar dat deze begrippen toch wel enige verdere verklaring verdienen. Een ontologie is een formele representatie van een wereldbeeld. Deze representatie bestaat uit een groot aantal concepten, gestructureerd in een strikte hiërarchie. Voor elk concept zijn er bepaalde eigenschappen gedefinieerd. Tussen concepten op verschillende niveau's bestaat er een ontologische relatie. 'Is a kind of' en 'Is a part of' zijn twee voorbeelden van relaties die vaak voorkomen in een ontologie. Dit soort ontologische relaties is bijvoorbeeld ook terug te vinden in mindmappen. Tussen concepten van een bepaald niveau en een hogerliggend niveau is er steeds sprake van overerving van eigenschappen. Eigenschappen van een bepaald concept van een hoger niveau, zijn dus ook van toepassing op alle concepten die zich onder dat bepaalde concept bevinden. Naast ontologieën zijn er ook taxonomieën. Een taxonomie is gelijkaardig aan een ontologie, maar de structuur is minder strikt. Bovendien is hier geen sprake van overerving van eigenschappen.

In domein-specifieke toepassingen is de koppeling tussen de term in de tekst en het concept uit de ontologie of taxonomie door NER, een belangrijke en essentiële stap. Deze combinatie vormt immers de basis voor een verdere analyse. Met deze technologie worden accuraatheidspercentages tot 95% gehaald. Dit is zeer goed, want zelfs indien dit proces door mensen wordt uitgevoerd, worden er geen perfecte scores van 100% bereikt. (Cunningham, 2005)

NER is in kleine mate afhankelijk van het domein, dit wil zeggen dat, wanneer het onderwerp van de geanalyseerde teksten lichtjes wijzigt, dit slechts een kleine aanpassing van het systeem met zich meebrengt. Een drastische wijziging van het onderwerp van de teksten, maakt evenwel aanzienlijke aanpassingen aan het systeem noodzakelijk. (Cunningham, 2005)

#### **4.3.2.2      Coreference resolution**

Het herkennen van identiteitsrelaties tussen de verschillende entiteiten is het onderwerp van *Coreference Resolution* (CO). In andere termen draait het erom te achterhalen welke, door NER ontdekte, entiteiten en verwijzingen zoals voornaamwoorden, naar hetzelfde begrip doelen. CO kan men opsplitsen in twee types van problemen. Een eerste type is het correct identificeren van anaforische verwijzingen<sup>21</sup>. De tweede categorie betreft het op de juiste manier herkennen van eigennamen. (Cunningham, 2005).

Aan de hand van de volgende voorbeeldzin<sup>22</sup> worden beide types geïllustreerd:

*"Ever since Eastman Kodak announced that it was outsourcing its information' systems function in 1988 to IBM, DEC, and Businessland, large organizations have found it acceptable to transfer their IT assets, leases, and staff to outsourcers."*

---

<sup>21</sup> Conform 'anaphora resolution' als NLP-techniek op pagina 43.

<sup>22</sup> Fragment uit Bijlage II.

Een voorbeeld van het eerste type probleem, de anaforische verwijzingen, is 'it' op regel 1. CO dient in dit geval dus te achterhalen dat deze 'it' verwijst naar de entiteit 'Eastman Kodak', net als het woord 'its' op regel 1. 'It' op regel 2 duidt dan weer op het zinsdeel 'to transfer their IT assets, leases, and staff to outsourcers'. Ten tweede is er de moeilijkheid om eigennamen juist te herkennen. Zo dient 'Eastman Kodak' uit het voorbeeld als één entiteit aangeduid te worden en moeten 'IBM' (regel 2) en 'International Business Machines Corporation' als synoniemen behandeld worden. Bovendien moet coreference resolution het verschil tussen 'IBM' en 'IBM Belgium/Luxembourg' kunnen achterhalen.

Uit het voorbeeld blijkt, dat coreference resolution toch niet zo eenvoudig is, als het op het eerste gezicht lijkt. Het is dan ook niet verwonderlijk, dat de resultaten rond 50 en 60% schommelen. Bovendien is deze categorie van information extraction sterk afhankelijk van het kennisdomein. CO dient in de toekomst dus nog zeker te worden verfijnd, zodat de precisie bij de resultaten zal toenemen.

#### **4.3.2.3    *Template Element construction***

De derde categorie van information extraction volgens Cunningham (2005), is de constructie van *Template Elements* (TE). Tijdens de aanmaak van TE wordt er gezocht naar de attributen van de entiteiten die met behulp van NER en CO reeds gevonden werden. Voor elk van deze entiteiten kan een soort van –denkbeeldig– formulier ingevuld worden, in praktijk zal dit een record in een database zijn. Met TE wordt dit formulier aangevuld met attributen uit de tekst. Deze attributen kunnen verschillende vormen aannemen, zoals een bepaalde waarde voor een datum of een synoniemen of alias voor een bepaalde entiteit.

Cunningham (2005) bestempelt TE als een zwak domein-gerelateerde taak en geeft aan dat de huidige resultaten rond 80% liggen. Mensen halen bij het uitvoeren van deze taak een gemiddelde van 95%. Dit duidt erop dat er voor TE nog vooruitgang mogelijk is.

#### **4.3.2.4    *Template Relation construction***

*Template Relation (TR) construction* is een onderdeel van information extraction waarbij, getracht wordt relaties te vinden tussen de verschillende entiteiten uit de constructie van TE. Aan de hand van de brontekst is men uiteraard maar in staat om slechts enkele van de in de realiteit vele mogelijke relaties in te vullen. Een concreet voorbeeld<sup>23</sup>:

*"Michael Palmer, the COO (and former CIO) of Allied Office Products, decided to bring in-house nearly all of the IT functions that had been outsourced."*

---

<sup>23</sup> Fragment uit Bijlage II.

Na het doorlopen van de vorige fases zouden de volgende entiteiten gevonden moeten worden: 'Michael Palmer', 'the COO' en 'Allied Office Products'. TR constructie zou vervolgens de relaties tussen deze entiteiten moeten kunnen vastleggen. In dit geval bezet Micheal Palmer de functie van COO<sup>24</sup> en werkt hij voor het bedrijf Allied Office Products.

De scores bij template relation constructie halen percentages van 75% en zijn eerder zwak domein-afhankelijk. (Cunningham, 2005)

#### **4.3.2.5 Scenario Template production**

De laatste categorie van taken binnen information extraction is *Scenario Template (ST) production*. *Scenario template production* combineert de resultaten van TE and TR in specifieke gebeurtenissen. Een scenario template beschrijft bijvoorbeeld de opvolging van persoon X door persoon Y als CEO van een organisatie. Aan de hand van de resultaten in TE en TR kan een bepaalde gebeurtenis of scenario gedefinieerd worden. Gezien de moeilijkheidsgraad van deze taak halen de betere systemen scores van om en bij de 60%, terwijl mensen die deze taak uitvoeren, minimum 80% scoren. *Scenario template production* is sterk afhankelijk van het domein en is bovendien gebonden aan de scenario's die zijn gedefinieerd. (Cunningham, 2005)

### **4.4 i.Know**

Voor de onderstaande beschrijving van de technologie van i.Know is gebruik gemaakt van een drietal teksten die door i.Know ter beschikking gesteld werden, met name 'Technical White Paper' (i.Know NV, 2005), 'i.Know's functionaliteiten' (i.Know NV, 2005) en 'Bedrijfspresentatie: i.Know, Intelligent Knowledge Management Systems N.V.' (Brands, 2003). Daarnaast is de informatie uit de gebruikte artikels geactualiseerd en aangevuld met de opgedane kennis tijdens de samenwerking met medewerkers van i.Know.

#### **4.4.1. Uitgangspunt**

De technologie van i.Know gaat uit van een relationele semantische benadering om uit ongestructureerde bronnen (voornamelijk tekst-documenten) kennis te winnen. De termen 'relationeel' en 'semantisch' duiden erop dat i.Know zijn technologie op twee verschillende uitgangspunten gebaseerd heeft.

Achter het relationele aspect van de benadering schuilt de statistische methode, vooral toegepast in het pure data mining domein, maar ook bij text mining. Dit 'positivistisch' uitgangspunt gaat er

---

<sup>24</sup> COO is de afkorting voor het Engelse 'Chief Operating Officer' of 'Chief Operations Officer'; in het Nederlands eventueel benoemd als operationeel directeur.

vanuit, dat alles voorspelbaar is, bijgevolg is (nieuwe) kennis te definiëren aan de hand van algoritmes. Hierbij wordt gebruik gemaakt van ontologieën, hetgeen formele voorstellingen zijn van (een deel van) de realiteit<sup>25</sup>.

Het tweede uitgangspunt, aangeduid door het woord 'semantisch', stelt de menselijke capaciteiten centraal. Hierbij gelooft men, dat kennis het resultaat is van het menselijk creatieve vermogen. Dit houdt in, dat kennis dus per definitie niet te voorspellen valt. Dit uitgangspunt wordt voornamelijk gehanteerd in domeinen als natural language processing en in *collaborative workspace*-toepassingen.

i.Know start dus vanuit een relationeel semantische methode waarbij de positivistisch en *human centric* aanpak gecombineerd worden. i.Know is bijgevolg een voorbeeld van een technologie waarbij text mining en natural language processing elkaar overlappen. De belangrijkste elementen van de technologie zijn enerzijds de ontologie en anderzijds het unieke idee van concept-relatie-concept (CRC). De concept-relatie-concept-structuur wordt verder in de tekst besproken.

#### **4.4.2. Technologie**

De basis van de door i.Know ontwikkelde technologie is het int.for<sup>®</sup> platform. Dit is een multidimensioneel platform waarin men drie componenten onderscheiden kan:

1. de multidimensionele kennisvoorstellingsmodule;
2. de component voor de gebruikersfunctionaliteit;
3. de interface voor de kennisuitwisseling.

##### **4.4.2.1 De multidimensionele kennisvoorstellingsmodule**

Deze module is het 'hart' van het volledige systeem en is op zijn beurt een geïntegreerd geheel van drie dimensies:

- perceptuele dimensie;
- logische dimensie;
- inferentiële dimensie.

De perceptuele dimensie bevat een beschrijving van de manier waarop concepten en relaties in de menselijke taal kunnen voorkomen. Deze beschrijvingen kunnen worden weergegeven in formele definities. Een voorbeeld uit de gevalstudie is de volgende taalkundige definitie van het concept outsourcing:

*Outsourcing* = { *outsourcing*, *outsource*, *outsourcer*, *outsourcers*, *IT outsourcing* }

---

<sup>25</sup> Conform de bespreking van ontologie in het kader van NER op pagina 49.

De definitie bestaat dus enerzijds uit een concept of een relatie en anderzijds uit de lexicale representaties (tussen de accolades) van dit concept of deze relatie. In het voorbeeld bevat de perceptuele dimensie voor het concept 'Outsourcing' vijf lexicale representaties, namelijk: 'outsourcing', 'outsource', 'outsourcer', 'outsourcers', en 'IT outsourcing'. Een lexicale representatie is een manier waarop een concept taalkundig voorgesteld wordt. In een geschreven tekst betekent dit, dat een lexicale representatie van een bepaald concept een woord of woordgroep is, dat naar dat specifieke concept doelt. In het voorbeeld duiden zowel het woord 'outsource' als de woordgroep 'IT outsourcing' naar het concept 'Outsourcing' en zijn bijgevolg dus beide een lexicale representatie van het concept 'Outsourcing'.

Aan de hand van deze formele definities is de perceptuele dimensie in staat om de woorden uit de brondocumenten aan reeds gedefinieerde concepten en relaties te koppelen. Als de dimensie in een tekst bijvoorbeeld het woord 'outsourcers' vindt, zal dit door de aanwezigheid van 'outsourcers' als lexicale representatie in de perceptuele definitie van het concept 'Outsourcing' gekoppeld worden aan dit concept.

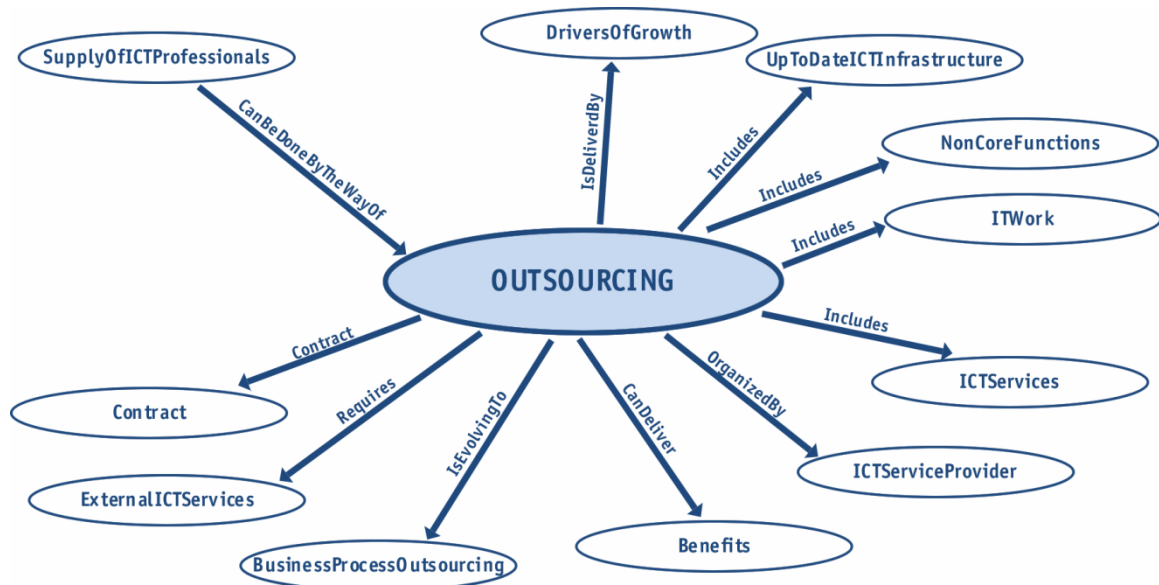
In feite kan de perceptuele dimensie vergeleken worden met de zintuigen van de mens. Hun taak is het registreren van de uitwendige prikkels. Op een vergelijkbare manier vangt de perceptuele dimensie de prikkels, in dit geval de woorden uit de bronteksten, op en koppelt deze aan één of meerdere concepten of relaties. De perceptuele dimensie voert dus het *named entity recognition-proces*<sup>26</sup> uit, volgens Cunningham (2005) één van de vijf belangrijkste types van informatie extractie.

De tweede dimensie wordt de logische dimensie genoemd. Deze dimensie bevat enerzijds definiëringen van specifieke kenmerken van concepten en relaties. Daarnaast beschikt de logische dimensie over informatie van de manier waarop verschillende concepten en relaties met elkaar gecombineerd kunnen worden.

Figuur 10  
Figuur 10 is een voorbeeld van een logische beschrijving van een concept dat voorkomt in het praktijkgeval, namelijk 'Outsourcing'. De figuur toont centraal het concept 'Outsourcing' met errond een aantal door relaties gekoppelde concepten. Zo zijn de concepten 'Outsourcing' en 'ExternalICTServices' gekoppeld door de relatie 'Requires'. De richting van de pijl staat symbool voor de richting van de relatie. In dit geval loopt de relatie van het centrale concept 'Outsourcing' naar 'ExternalICTServices', dus: 'Outsourcing' 'Requires' 'ExternalICTServices'.

---

<sup>26</sup> NER werd beschreven op pagina 48.



**Figuur 10: Logische beschrijving van het concept 'Outsourcing'**

Zo een combinatie van twee concepten, verbonden door een relatie, wordt een concept-relatie-concept (CRC) genoemd. Een CRC vormt een centraal element binnen de technologie van i.Know. De formele notatie van een CRC is als volgt:

NaamConcept\_1; <SNaamRelatie; NaamConcept\_2

Voor het voorbeeld, besproken in de vorige alinea, geeft dit de volgende notatie:

outsourcing; <SRequires; ExternalICTServices;

Als al de logische definities van de beschikbare concepten worden gecombineerd, wordt er een uitgebreid netwerk van CRC's bekomen. Maken we opnieuw de vergelijking met de mens, dan kunnen we de logische dimensie koppelen aan de geheugenfunctie van onze hersenen. Het geheugen is een netwerk van relaties tussen verschillende gebeurtenissen.

Tot slot bevat de multidimensionele kennisvoorstellingsmodule een inferentiële dimensie. Met behulp van deze dimensie kan een waarde toegekend worden aan een combinatie van concepten en relaties in functie van andere CRC's. Zo kunnen aan logische relaties booleaanse waarden (waar en onwaar) en psychologische toestanden (zoals geschikt, gewenst, niet gewenst, ...) gekoppeld worden. Aan logische concepten kunnen logische toestanden (hoog, laag, rijk, arm, ...) toegekend worden. Met inferentiële definities kan de waarde, toegekend aan een eerste CRC, de toestand van een tweede gekoppelde CRC bepalen.

Deze inferentiële definities van concepten en relaties bieden de mogelijkheid, om uit een ongestructureerde bron, processen en redeneringen te definiëren binnen het formeel kennisnetwerk van de logische dimensie. De inferentiële dimensie kan men dus vergelijken met het redeneervermogen van de hersenen van de mens. In het text mining-kader staat de inferentiële

dimensie op dezelfde hoogte als *scenario template production*. In beide gevallen vindt er –in bepaalde mate – een redenering plaats op basis van gegevens die uit een tekst gehaald werden.

De drie dimensies van de multidimensionele kennisvoorstellingsmodule voegen elk op hun beurt aan de gegevens in het tekstdocument een extra dimensie toe. Hierdoor is het mogelijk om uit de ongestructureerde bron de gecodeerde kennis naar boven te halen en deze in een formeel kennisschema te brengen.

#### **4.4.2.2    *De component voor de gebruikersfunctionaliteit***

De component voor de gebruikersfunctionaliteit van het int.for® platform bevindt zich bovenop de multidimensionele kennisvoorstellingsmodule die zojuist besproken werd. Zoals klaarduidelijk blijkt uit de naam, zorgt deze tweede component voor de eigenlijke functionaliteit voor de gebruikers. De component is opgebouwd uit verschillende modules, die elk een specifieke functie vertegenwoordigen. De modulestructuur maakt het mogelijk om twee functies, in feite twee modules, te combineren en alzo de functionaliteit te vergroten. De modulestructuur is gebaseerd op de kennislevenscyclus, die uit de volgende zes fasen bestaat:

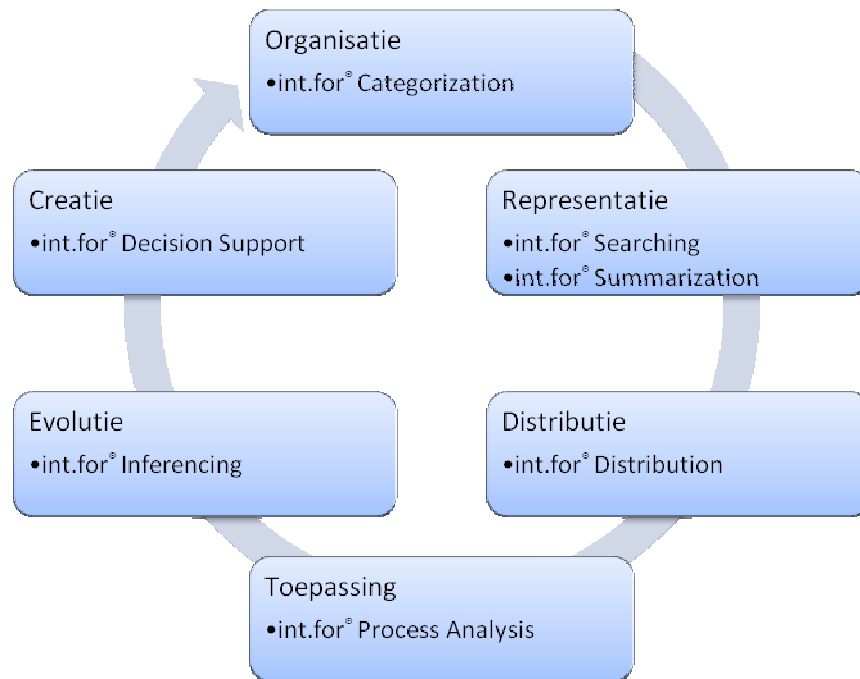
1. Organisatie
2. Representatie
3. Distributie
4. Toepassing
5. Evolutie
6. Creatie

De verschillende functies van de i.Know-technologie kunnen worden gekoppeld aan één fase uit de kennislevenscyclus, zie figuur 11. Elke fase en de bijbehorende functionaliteit worden achtereenvolgens besproken.

Tijdens de eerste fase van kennislevenscyclus, namelijk de organisatie-fase, wordt er een structuur ontwikkeld zodat de kennis op een efficiënte manier ter beschikking gesteld kan worden. Hierbij kan de informatie geclassificeerd worden op basis van een ontologie.

Int.for® Categorization is de functie die deze taak binnen het int.for®-platform uitvoert. De functie zorgt voor de categorisatie van informatie vanuit een ontologie, gedefinieerd op basis van taalonafhankelijke concept-relatie-concept-netwerken. Het feit dat deze ontologie werkt op basis van CRC's en dus niet met synoniemen, maakt het zeer flexibel. Bovendien maakt deze classificatie een contextuele analyse mogelijk. De disambiguering van woorden en woordgroepen gebeurt, net zoals bij het leesproces bij mensen, op basis van de context. In eerste instantie worden de woorden en woordgroepen gelinkt met de gekende concepten. Hiervoor zal de int.for® Categorization intensief gebruik maken van de perceptuele dimensie uit de

multidimensionele kennisvoorstellingsmodule. Indien een bepaalde lexicale representatie aan verschillende concepten gekoppeld is, kan aan de hand van de overige concepten en de relaties tussen deze concepten en de verschillende mogelijke concepten het juiste concept in deze context gekozen worden.



**Figuur 11: Kennislevenscyclus**

Een vereenvoudigd voorbeeld ter illustratie van het disambigueringsproces:

*"Offshore outsourcing is using organizations from developing countries to write code and develop systems."*

'Code' is ondermeer een lexicale representatie van de concepten 'ComputerProgram' en 'Obligation' en aan het woord 'Systems' zijn 'System' en 'Animal' gekoppeld. Door de analyse van de overige concepten en de andere relaties in deze zin, zal de categorisatie-functie voor de lexicale representaties 'code' en 'systems' respectievelijke de concepten 'ComputerProgram' en 'System' als meer relevant beschouwen. Hieruit kan het systeem dan afleiden, dat 'code' en 'systems' in dit voorbeeld eerder in een technologische context dienen geplaatst te worden.

Deze contextuele disambiguering haalt een zeer hoge precisie. De int.for® Categorization kan op dezelfde manier nieuwe en onbekende informatie automatisch herkennen en signaleren. Eventueel kan deze nieuwe informatie vervolgens in de ontologie geïntegreerd worden. Hierdoor wordt het tijdsintensieve definiëeringsproces van ontologieën sterk ingekort.

Nadat de kennis in de eerste fase in een formele structuur ingepast is, kan deze in de aansluitende representatie-fase getoond worden aan de gebruiker. Enerzijds betreft het dan een voorstelling van



de relevante kennis. Anderzijds dienen ook verbanden tussen de verschillende delen beschikbaar gemaakt te worden.

Binnen de i.Know-technologie zijn int.for® Searching en int.for® Summarization de functies die tot deze fase behoren. Int.for® Searching is een zoekmiddel dat via contextuele zoekactie direct de juiste informatie beschikbaar maakt. i.Know zelf noemt deze functie een 'vindmachine' om zich te onderscheiden van de –vaak op het internet gebruikte– 'zoekmachines'. Dankzij de contextuele disambiguering in de eerste fase van de kenniscyclus worden door de int.for® Searching enkel relevante documenten weergegeven en 'vindt' men dus direct de gewenste informatie. Bij zoekmachines daarentegen krijgt men een lijst van alle documenten waarin de opgegeven zoekterm teruggevonden werd, met als gevolg dat er groot aantal irrelevante resultaten gekregen worden. De gebruiker moet dan zelf de resultaten doorlopen en analyseren om zo 'op zoek' te gaan naar de gewenste informatie.

Wordt er bijvoorbeeld met een zoekmachine naar 'Offshore outsourcing in India', dan zullen er waarschijnlijk eerst wel een aantal resultaten gevonden worden die relevant zijn voor de zoekbewerking. Maar er zullen ook resultaten insluipen die het over een andere vorm outsourcing hebben of waar bijvoorbeeld gewoon 'India' in voorkomt. Deze laatste resultaten zijn natuurlijk niet relevant en vandaar ook niet gewenst. Met de 'vindmachine' van i.Know worden alleen documenten gevonden worden waar het daadwerkelijk gaat over 'offshore outsourcing in India.'

De int.for® Summarization-functie maakt het mogelijk om samenvattingen te genereren. Samenvattingen kunnen gemaakt worden van verschillende documenten, die bovendien in meerdere talen geschreven kunnen zijn g. Uiteraard is het ook mogelijk om de lengte van de samenvatting in een percentage ten opzichte van de originele tekst of in aantal zinnen op te geven. Opnieuw wordt er dankzij het contextuele disambigueringsproces een hoge precisie gehaald, waardoor er samenvattingen worden gegenereerd met een hoge consistentie, aangezien de oorspronkelijke interne samenhang behouden blijft.

Distributie is de derde fase in de levenscyclus die kennis doorloopt. Deze stap bestaat dus uit het verspreiden van de informatie alsook het opslaan ervan. De int.for® Distribution-module verzorgt de taken binnen de distributie-fase. Met behulp van deze module kan de gebruiker een individueel profiel beheren, waarin hij of zij kan aanduiden in welke topics hij interesse heeft. Doordat de topics gelinkt zijn met de CRC's, die ook bij de semantische analyse van de informatie gebruikt worden, krijgt de gebruiker enkel de informatie met betrekking tot de door hem opgegeven topics opgestuurd.

Na de distributie-fase is de kenniscyclus halfweg en op dit moment beschikt de gebruiker zelf over de relevante kennis in verband met de door hem gewenste topics. In de volgende fase is het tijd

om de aanwezige kennis te gaan gebruiken in de verschillende werkprocessen. Vandaar dat deze vierde fase ook toepasselijk de naam toepassingsfase meekreeg.

De int.for<sup>®</sup> Process Analysis functie maakt het mogelijk om ook de toepassing van de kennis voor een gedeelte te automatiseren. Deze functie is namelijk in staat om een analyse te maken van grote hoeveelheden van informatie aan de hand van door het systeem zelf ontdekte of door de gebruiker gedefinieerde procesbeschrijvingen. Vervolgens kunnen dan relevante informatiebronnen automatisch met het desbetreffende proces geassocieerd worden. Dit maakt dat de gebruiker de noodzakelijke kennis op de juiste plaats en op het juiste moment ter beschikking heeft en deze zo onmiddellijk tijdens het werkproces kan toepassen.

De vijfde fase in de levenscyclus van kennis heet evolutie. Het toepassen van de bestaande kennis in de vorige fase, leidt tot het ontdekken van nieuwe inzichten en evoluties in deze fase. De int.for<sup>®</sup> Inferencing-functie ondersteunt de evolutie-fase in de technologie van i.Know. Int.for<sup>®</sup> Inferencing zorgt voor een toetsing van de bestaande kennis aan de beschreven processen. Op basis van deze toetsing kunnen dan eventuele suggesties voor aanpassingen voorgesteld worden. Tot slot is het mogelijk om kennisevoluties in een bestaand systeem automatisch te herkennen en te integreren.

Uiteindelijk is met de creatie-fase de kennislevenscyclus rond. Deze fase leidt per definitie naar de creatie van nieuwe kenniselementen. Voor elke van deze nieuwe kenniselementen wordt vervolgens een nieuwe kennislevenscyclus opgestart. i.Know biedt met de int.for<sup>®</sup> Decision Support technologische ondersteuning in deze laatste fase van de kennislevenscyclus. Deze functie voorziet de gebruiker van een visuele weergave van de beschikbare kennis. Op deze manier is de gebruiker in staat om optimale beslissingen te nemen over de gepresenteerde informatie.

#### **4.4.2.3     *De interface voor de kennisuitwisseling***

De derde component van het int.for<sup>®</sup>-platform is de interface voor de kennisuitwisseling. Deze component bevindt zich in feite rond de multidimensionale kennisvoorstellingsmodule en de component voor de gebruikersfunctionaliteit. Deze component maakt het mogelijk om kennis uit te wisselen met aan het systeem externe elementen.

## **HOOFDSTUK 5 : Gevalstudie**

### **5.1 Doel van de gevalstudie**

In de vorige hoofdstukken werd een literatuurstudie gedaan van achtereenvolgens e-learning, leerobjecten en de domeinen van text mining en natural language processing. In de gevalstudie is het doel, om elementen uit deze voorafgaande hoofdstukken samen te brengen. We willen namelijk een poging ondernemen om de ontwikkeling van een e-learningcursus efficiënter te laten verlopen. Om dit te realiseren, wordt er gekeken of het mogelijk is, om de decompositiefase van een e-learningcursus te automatiseren met behulp van technieken uit text mining en natural language processing. Heel concreet, wordt de software van i.Know gebruikt om de decompositiefase van een bestaande e-learningcursus, namelijk 'Strategie en informatiemanagement', gedeeltelijk opnieuw uit te voeren.

Het uitgangspunt is, dat de e-learningcursus voor een bepaald vak opgebouwd wordt op basis van een beschikbare tekst, bijvoorbeeld de behandelde hoofdstukken uit een handboek. Voor het vak 'Strategie en informatiemanagement' werd er gebruik gemaakt van het boek 'Business Driven Technologies'<sup>27</sup>. Om in de tekst een structuur te kunnen aanbrengen, is het noodzakelijk om te weten welke onderwerpen of topics erin besproken worden. Voor elk element uit de structuur dienen vervolgens de 'atomische leerobjecten' samengesteld te worden. Dit zijn onder meer de volledige tekst, een samenvatting, sleutelwoorden, figuren, ... Op basis van de structuur en de atomische leerobjecten kan dan, bijvoorbeeld aan de hand van het EDUMAP-scenario, een mindmap gecreëerd worden. De samenstelling van een e-learningcursus volgens een bepaald scenario valt buiten het bestek van de gevalstudie, aangezien het de bedoeling is, om enkel op de decompositiefase te focussen.

Tijdens het uitvoeren van de gevalstudie is de aanpak gewijzigd, vandaar dat er een onderscheid gemaakt wordt tussen een eerste en een tweede methode. Na een beschrijving van de uitgevoerde stappen binnen beide methoden, worden er voor de volledige gevalstudie enkele conclusies getrokken.

### **5.2 Methode 1**

Deze eerste methode werd toegepast op de hoofdstukken 13 tot en met 20 van het gebruikte boek 'Business Driven Technologies'. De stappen die gezet werden bij een eerste poging om voor een paragraaf topics te herkennen, worden hieronder beschreven.

---

<sup>27</sup> Haag, S., Baltzan, P., & Philips, A. (2006) *Business Driven Technologies*. New York: McGraw-Hill/Irwin.

### 5.2.1. Eerste categorisatie

In een eerste stap werd de volledige brontekst in het geheel door het int.for® platform geanalyseerd. Uit deze eerste analyse werd een overzicht gegenereerd van woorden en woordgroepen die door het systeem niet herkend werden. Dit overzicht bevatte een lijst van 737 niet-herkende monogrammen, dus woorden, samen met de frequentie waarmee ze in de tekst voorkwamen. Daarnaast was er een tweede lijst van 1.123 voor het systeem onbekende bigrammen -woordgroepen bestaande uit twee delen- eveneens met het aantal keer dat ze in de tekst stonden.

Een voorbeeld:

*"Hewlett-Packard, SAP, and Qwest have formed one of the first major alliances for providing ASP services."*

Uit deze zin werd ondermeer 'Hewlett-Packard' niet herkend. In de lijst met niet-herkende monogrammen werd dit weergegeven als 'hewlett-packard 2 hewlett-packard'. De software reduceert de originele vorm al meteen tot een standaardvorm, waarbij bepaalde tekens en hoofdletters verwijderd worden. Het getal in het midden duidt op het aantal keer dat het woord in de volledige tekst voorkomt, in dit geval kwam 'Hewlett-Packard' dus tweemaal voor in de brontekst.

Vervolgens werden de monogrammen aan een eerste ruwe categorisatie onderworpen. Tijdens dit proces werden de woorden in eerste instantie ingedeeld volgens type, namelijk 'Concept', 'Relation' of 'Modifier'.

'Concept'-woorden duiden in het algemeen op een bepaald begrip. De zelfstandige naamwoorden zijn hier een klassiek voorbeeld van. Bij dit type werd er een verdere opsplitsing gemaakt in een aantal vaak voorkomende basiscategorieën uit de ontologie van i.Know, zoals 'Adverb', 'Company', 'Currency', ... Tabel 1 geeft een overzicht van de verschillende categorieën die toegekend werden. Het voorbeeld van 'Hewlett-Packard' uit de vorige paragraaf is een voorbeeld van een 'concept'-woord. Aangezien 'Hewlett-Packard' een bedrijfsnaam is, dient het te worden toegewezen aan de categorie 'Company'. Dit toewijzen werd in deze eerste categorisatie eenvoudigweg gedaan door in het Excel-bestand op de rij van 'Hewlett-Packard' in een volgende kolom 'Company' te noteren.

Het type 'Relation' bevat woorden die informatie geven over de relatie tussen concepten. De werkwoorden vormen onder meer een grote groep binnen dit type. Uit de onderstaande zin stond het woord 'integrate' op de lijst met monogrammen.

*"For instance, newly announced PDAs integrate phones, [...]"<sup>28</sup>*

---

<sup>28</sup> Fragment uit Haag, S., Baltzan, P., & Philips, A. (2006) *Business Driven Technologies*. New York: McGraw-Hill/Irwin.

Het woord 'integrate' vertelt de lezer dus iets over de verhouding tussen 'PDA's' en 'phones', vandaar dat dit als 'Relation' gemarkeerd werd.

**Tabel 1: Eerste categorisatie monogrammen**

| Type     | Categorie           |
|----------|---------------------|
| Concept  | Adverb              |
|          | Company             |
|          | Currency            |
|          | Information measure |
|          | Interval            |
|          | Length measure      |
|          | Name                |
|          | Number              |
|          | Plane angle measure |
|          | Salary              |
|          | Substantive         |
|          | Time interval       |
|          | Volume measure      |
| Relation |                     |
| Modifier |                     |

Het derde type zijn de 'Modifiers'. Deze woorden verschaffen ons extra informatie over concepten of relaties en duiden eventueel op een bepaalde toestand of geven de 'mate' van iets aan.

*"For a business whose IT department is a non-core function, [...]"<sup>29</sup>*

'Non-core', uit de voorbeeldzin hierboven, behoort tot de groep 'modifiers'. In dit geval geeft het aan, over welk type functie het gaat.

Uit Tabel 1 zijn 'number', 'relation', 'substantive' en 'company' de meest toegekende categorieën. Aan ruim de helft van de woorden kon na deze eerste categorisatie nog geen categorie toegekend worden. Voor deze woorden was er in vele gevallen geen geschikt concept aanwezig.

<sup>29</sup> Fragment uit Bijlage II.

### 5.2.2. Ontwikkeling taxonomie en aanpassing ontologie i.Know

Parallel met de eerste ruwe categorisatie werd er een eigen taxonomie voor het ICT-vakgebied ontwikkeld. Deze taxonomie is gebaseerd op 'The ACM Computing Classification System (1998)'<sup>30</sup> en 'MISQ Keyword Classification Scheme'<sup>31</sup> en de expertise van prof. Schreurs. De volledige taxonomie is terug te vinden in Bijlage I.

**Tabel 2: Uittreksel aangepaste ontologie i.Know**

|            |                                               |
|------------|-----------------------------------------------|
| Domain     |                                               |
| Phenomenon |                                               |
|            | GeneralisedStructure                          |
|            | AbstractStructure                             |
|            | Single                                        |
| [...]      |                                               |
|            | TemporalThing                                 |
|            | InformationalThing                            |
|            | Summary                                       |
|            | Webpage                                       |
| [...]      |                                               |
|            | Map                                           |
|            | Software                                      |
|            | ComputerProgram                               |
|            | <i>OperatingSystem</i>                        |
|            | <i>ManagementInformationSystem</i>            |
|            | <i>CommuncationSystem</i>                     |
|            | <i>EmailSystem</i>                            |
|            | <i>InstantMessaging</i>                       |
|            | <i>Webbrowser</i>                             |
|            | <i>CollaborativeWorkSystem</i>                |
|            | <i>ContentManagementSystem</i>                |
|            | <i>WorkflowManagementSystem</i>               |
|            | <i>GroupwareSystem</i>                        |
|            | <i>Peer-to-peerSystem</i>                     |
|            | <i>KnowledgeManagementSystem</i>              |
|            | <i>EnterpriseResourcePlanningSoftware</i>     |
|            | <i>SupplyChainManagementSoftware</i>          |
|            | <i>CustomerRelationshipManagementSoftware</i> |
|            | SoftwareType                                  |
|            | <i>ProprietarySoftware</i>                    |
|            | <i>OpenSoftware</i>                           |
|            | <i>Linux</i>                                  |
|            | Sentence                                      |
|            | Publication                                   |
| [...]      |                                               |

---

<sup>30</sup> Bron: <http://www.acm.org/class/1998/ccs98.html>

<sup>31</sup> Bron: <http://www.misq.org/roadmap/code/level1.html>

Na de eerste categorisatie kon aan vele woorden uit de lijst met monogramman geen concept toegewezen worden, vermits er in de ontologie van i.Know geen geschikt concept voorhanden was. Op basis van de eigen ontwikkelde taxonomie en de lijst van niet-herkende monogrammen werden nieuwe concepten bepaald, die in aanmerking kwamen om aan de ontologie van i.Know toegevoegd te worden. Met behulp van de expertise van i.Know werd bepaald welk van deze voorgestelde categorieën daadwerkelijk opgenomen werden en op welke positie in de hiërarchie deze ingevoegd werden. Een uittreksel uit de aangepaste ontologie is terug te vinden in Tabel 2. De cursief gedrukte woorden uit de tabel duiden op concepten die toegevoegd werden. De toegevoegde concepten werden ingedeeld met behulp van de zelf ontwikkelde taxonomie (zie Bijlage I). Er werden enkel concepten aan de ontologie toegevoegd; er werden dus geen nieuwe elementen van het type 'relation' of 'modifier' toegevoegd.

### 5.2.3. Verfining categorisatie

Vooraleer de eerste categorisatie verder te verfijnen, werden de lijsten met mono- en bigrammen samengevoegd. De woorden of woordgroepen die als relevant werden beschouwd, werden gemarkeerd. 'Relevant' dient in dit geval te worden geïnterpreteerd als noodzakelijk voor een correcte interpretatie van de brontekst. Woorden die in het kader van de brontekst minder of zelfs niet essentieel waren, werden bij de verdere categorisatie buiten beschouwing gelaten. Tabel 3 toont enkele voorbeelden van relevante mono- en bigrammen en het concept waaraan ze gekoppeld werden. De woorden zijn relevant omdat ze binnen het ICT-domein belangrijk zijn.

**Tabel 3: Voorbeelden van relevante lexicale representaties**

| <b>Lexicale representatie</b> | <b>Concept</b> |
|-------------------------------|----------------|
| network administrators        | Administrator  |
| service provider              | WebService     |
| e-business                    | EBusiness      |
| outsourcers                   | Outsourcing    |

'Weather', 'skiers', 'etc', ... zijn voorbeelden van woorden die niet meteen van belang zijn om een tekst uit het ICT-domein te kunnen interpreteren. Deze woorden werden dan ook niet aan een concept gekoppeld. In het geval van de bigrammen bestond deze lijst ook uit betekenisloze combinaties van twee woorden. Dit is te verklaren doordat het systeem bepaalde leestekens, onder meer komma's, automatisch verwijderd. Het gevolg is, dat er in de lijst woordgroepen opgenomen zijn, die in de tekst zelf gescheiden zijn door een komma en dus in feite geen zinvolle woordgroep zijn.

*"In the same way that the computer itself announced the start of the information age, developments in network-centric computing facilitated a new age of knowledge rather than of information."*<sup>32</sup>

Uit de bovenstaande zin maakte bijvoorbeeld 'age developments' deel uit van de lijst met bigrammen. In werkelijkheid zijn deze twee woorden gescheiden door een komma en is het duidelijk dat deze geen betekenisvolle woordgroep vormen. Dit soort bigrammen werd dan ook overgeslagen.

De volledige lijst met relevante woorden en woordgroepen werd vervolgens verder overlopen en elk item kreeg een zo specifiek mogelijk concept toegewezen. Deze toewijzing gebeurde op dezelfde manier als tijdens de eerste categorisatie, door in het Excel-bestand in een volgende kolom het desbetreffende concept, 'modifier' of 'relation' te noteren. De koppeling van concepten met de woorden en woordgroepen in het systeem werd door i.Know uitgevoerd.

#### 5.2.4. Output

Na de categorisatie van de relevante woorden en woordgroepen als specifiek concept, modifier of relatie werd een eerste volledige output van de brontekst gegenereerd. De onderstaande code toont het eerste gedeelte van de output voor de eerste paragraaf.

```
The Internet[the internet] { "Internet" "InformationTechnologicalDevice" }
is[is] { "StoBe" "StoBe" "state" "SKindOf" } a powerful[a powerful] {
(Modifier) } channel[channel] { "Device" "StationaryArtifact"
"StreamWaterArea" "Instrument" } that[that] presents[presents] {
"TimeInterval" "UnilateralGiving" } new[new] { (Modifier) }
opportunities[opportunities] { "CurrencyMeasure" } for[for] {
"SConnectiveRelation" } an[an] { (ModifierPremodifier) }
organization[organization] { "Organization" } to[to] { "Sto" } touch[touch] {
"Communication" } customers,[customers] { "EconomicalSocialRole" "Customer" }
enrich[enrich] { "Drug" } products[products] { "MultiplicationFn" "Products" }
and[and] { "SisCombinedBy" } services[services] { "Service" } with[with] {
"SAssociatedWith" } information,[information] { "InformationalThing"
"InformationMeasure" } and reduce[and reduce] { "SisCombinedBy" }
costs.[costs] { "CurrencyMeasure" "Cost" }
```

Elk woord of elke woordgroep wordt door het systeem eerst herleid naar een standaardvorm, die weergegeven wordt tussen vierkante haakjes. Nemen we de eerste woordgroep 'The Internet' als voorbeeld, dan zien we, dat de hoofdletters verwijderd worden en de standaardvorm 'the internet' is. Vervolgens toont deze output welke verschillende concepten, modifiers of relaties aan deze lexicale representatie verbonden zijn. Deze informatie wordt tussen de accolades weergegeven. Zo

---

<sup>32</sup> Fragment uit Haag, S., Baltzan, P., & Philips, A. (2006) *Business Driven Technologies*. New York: McGraw-Hill/Irwin.



wordt de lexicale representatie 'the internet' gekoppeld aan de concepten 'Internet' en 'InformationTechnologicalDevice'. Relaties kunnen herkend worden aan de hoofdletter 'S' in het begin van de naam. Betreft de lexicale representatie een modifier, dan wordt dit binnen de accolades als volgt aangeduid: '(Modifier)'.

Op basis van dit eerste gedeelte werd er een controle op de categorisatie uitgevoerd. Lexicale representaties die niet, tot een verkeerd concept, of fout als relatie of modifier werden benoemd, werden opgespoord en vervolgens gecorrigeerd.

Het tweede gedeelte van de output betreft een overzicht van de concept-relatie-concept-koppelingen, die de software automatisch in de desbetreffende paragraaf gedetecteerd heeft. De detectie gebeurt door na te gaan welke concepten in een bepaalde paragraaf voorkomen; daarna wordt er nagegaan welke relaties er tussen de concepten kunnen bestaan. De concept-relatie-concepten of CRC's die voor de eerste paragraaf werden gevonden, werden als volgt weergegeven:

```
InformationTechnologicalDevice("The Internet") SConnectiveRelation("for")
Communication("touch")
InformationTechnologicalDevice("The Internet") SisCombinedBy("and")
InformationalThing("information,")
InformationTechnologicalDevice("The Internet") SisCombinedBy("and")
CurrencyMeasure("costs.")
Device("channel") SConnectiveRelation("for") Organization("organization")
Device("channel") SConnectiveRelation("for")
EconomicalSocialRole("customers,")
Device("channel") SConnectiveRelation("for") Drug("enrich")
Device("channel") SConnectiveRelation("for") Service("services")
CurrencyMeasure("opportunities") SConnectiveRelation("for")
MultiplicationFn("products")
```

Daarnaast wordt de volledige output manueel doorzocht en werden zinvolle en liefst voor de tekst relevante CRC's genoteerd. Deze intensieve taak werd verricht om één, de kans dát een topic, en twee, dat ook het juiste topic gevonden wordt, te verbeteren. De achterliggende redenering wordt zometeen verduidelijkt.

```
According to[according to] { "SManifestationOf" } the research[the research] {
"Research" } firm[firm] { "Company" "Corporation" (Modifier) }
Strategy[strategy] { "Plan" } Analytics,[analytics] { "FieldOfStudy" "Company"
} the global[the global] { (Modifier) } m-commerce market[m-commerce market] {
"MCommerce" } will be[will be] { "StoBe" "SIntentionallyCausedProcess" "StoBe"
} worth[worth] { "CurrencyMeasure" (Modifier) } over[over] {
"SAssociativeRelation" "SSpatially-related-to" } $200[$200] {
"UnitedStatesDollar" } billion[billion] { "Billion" } by[by] {
"SSpatialRelation" "SCausedBy" } 2005,[2005] { "Year" } and some[and some] {
"SisCombinedBy" } 350[350] { (Modifier) "PositiveInteger" } million[million] {
"GenValFn" "Million" } customers[customers] { "EconomicalSocialRole"
```

```
"Customer" } will be[will be] { "StoBe" "SIntentionallyCausedProcess" "StoBe"
} generating almost[generating almost] { "SGenerates" } 14 billion[14 billion]
{ "Billion" } transactions[transactions] { "Accounting" "Transaction" }
annually.[annually]
```

Uit de output van bovenstaande paragraaf werd manueel bijvoorbeeld de volgende CRC gehaald:

```
Customer;<SGenerates;Transaction
```

Deze relaties werden aan de hand van bovenstaande notatie in het systeem ingevoerd.

## 5.2.5. Topics

In deze volgende stap werd er een lijst van topics opgesteld. Deze lijst is noodzakelijk, wil het systeem aan een bepaald stuk tekst een topic toewijzen. De lijst met topics werd gebaseerd op de inhoudstabel van de brontekst en de manueel aangemaakte mindmappen over deze tekst. Tabel 4 geeft een overzicht van de topics. Aan elk van de topics werd het meest toepasselijke concept gekoppeld. Elk topic werd dus gekoppeld aan één concept, met uitzondering van de topics 'WirelessOrganization' en 'InformationTechnologicalSecurity', die aan twee verschillende concepten gelinkt werden.

**Tabel 4: Overzicht topics tekst 1<sup>33</sup>**

| Topic                      | Concept 1                 | Concept 2 |
|----------------------------|---------------------------|-----------|
| EBusiness                  | EBusiness                 |           |
| EBusinessModel             | EBusinessModel            |           |
| BusinessToBusiness         | BusinessToBusiness        |           |
| BusinessToBusiness         | BusinessToBusiness        |           |
| ECommerce                  | ECommerce                 |           |
| EMarketplace               | EMarketplace              |           |
| MCommerce                  | MCommerce                 |           |
| Collaboration              | Collaboration             |           |
| CollaborativeWorkSystem    | CollaborativeWorkSystem   |           |
| ContentManagementSystem    | ContentManagementSystem   |           |
| WorkflowManagementSystem   | WorkflowManagementSystem  |           |
| GroupwareSystem            | GroupwareSystem           |           |
| Peer-to-peerSystem         | Peer-to-peerSystem        |           |
| KnowledgeManagement System | KnowledgeManagementSystem |           |
| Partnership                | Partnership               |           |
| Insourcing                 | Insourcing                |           |
| Outsourcing                | Outsourcing               |           |

<sup>33</sup> In Tabel 4 bezitten topic en concept meestal dezelfde naam. Formeel is er echter steeds een onderscheid, aangezien de naam van een topic steeds voorafgegaan wordt door 'Topic\_'. Het topic 'EBusiness' wordt dus formeel genoteerd als 'Topic\_EBusiness'. De notatie voor het bijbehorende concept 'EBusiness' blijft gewoon 'EBusiness'. Als uit de context duidelijk blijkt, dat het om een topic gaat, wordt, om de leesbaarheid te verbeteren, de prefix 'Topic\_' weggelaten. Op plaatsen waar er eventueel toch verwarring zou kunnen ontstaan, wordt bij topics de prefix 'Topic\_' wel toegevoegd.

| Topic                                  | Concept 1                              | Concept 2     |
|----------------------------------------|----------------------------------------|---------------|
| BusinessProcessOutsourcing             | BusinessProcessOutsourcing             |               |
| Offshoring                             | Offshoring                             |               |
| Networks                               | Networks                               |               |
| VirtualPrivateNetwork                  | VirtualPrivateNetwork                  |               |
| ElectronicDataInterchange              | ElectronicDataInterchange              |               |
| ElectronicTradingNetwork               | ElectronicTradingNetwork               |               |
| ServiceProviders                       | WebService                             |               |
| InternetServiceProvider                | InternetServiceProvider                |               |
| ApplicationServiceProvider             | ApplicationServiceProvider             |               |
| OnlineServiceProvider                  | OnlineServiceProvider                  |               |
| Portal                                 | Portal                                 |               |
| EnterpriseInformationPortal            | EnterpriseInformationPortal            |               |
| InternetStandard                       | InternetStandard                       |               |
| OpenSoftware                           | OpenSoftware                           |               |
| Linux                                  | Linux                                  |               |
| WirelessOrganization                   | WirelessMobile                         | Organizationn |
| Wireless                               | Wireless                               |               |
| Mobile                                 | Mobile                                 |               |
| 21stCenturyOrganization                | 21stCenturyOrganization                |               |
| InformationTechnologicalInfrastructure | InformationTechnologicalInfrastructure |               |
| InformationTechnologicalSecurity       | InformationTechnologicalSecurity       | Security      |
| InformationTechnologicalIntegration    | InformationTechnologicalIntegration    |               |

Om een topic te kunnen toewijzen aan een bepaald stuk van de originele tekst, dient de tekst eerst opgesplitst te worden in een aantal aparte delen. Aangezien de software dit niet automatisch kan doen, diende dit manueel te gebeuren. Om de resultaten zo weinig mogelijk op een subjectieve manier te beïnvloeden, werd geopteerd om de tekst op te splitsen volgens de verschillende paragrafen van de originele tekst. Dit maakte dat de tekst opgedeeld werd in 168 paragrafen.

#### 5.2.6. Resultaten topic-herkenning

In de laatste stap werden de 168 paragrafen door de software geanalyseerd om een relevant topic aan elke afzonderlijke paragraaf te kunnen toewijzen. Het bepalen of een topic relevant of niet relevant is voor een specifieke paragraaf, hangt van verscheidene factoren af.

In eerste instantie speelt natuurlijk de aanwezigheid van de concepten gekoppeld aan een bepaald topic, een grote rol voor de relevantie van dat topic voor die bepaalde paragraaf. Hoe meer (lexicale) representaties van een bepaald concept gekoppeld aan het topic in de tekst teruggevonden worden, hoe groter de relevantie van dat topic voor die paragraaf.

Vervolgens speelt ook de aanwezigheid van de voor het topic belangrijke concepten in gevonden en toegevoegde CRC's een rol. Hier wordt er niet enkel rekening gehouden met de frequentie, maar is

ook het type relatie en voorkomen met andere relevante concepten belangrijk. Het hele netwerk van relaties speelt dus een rol. Dit is meteen ook de reden waarom er in een eerdere stap manueel nog relevante CRC's toegevoegd werden, die niet automatisch door de software gevonden werden. Hoe dit algoritme precies werkt, behoort tot de kritische bedrijfskennis van i.Know.

De resultaten van de analyse waren niet goed. Zo werd er voor 102 paragrafen geen relevant topic gevonden. De overige 66 paragrafen kregen allemaal hetzelfde topic 'WirelessOrganization' toegewezen. In een aantal gevallen was dit inderdaad correct, maar voor de overige paragrafen was dit foutief.

Dit enig herkende topic is nota bene juist datgene, waaraan als uitzondering twee concepten toegevoegd werden. Hierin ligt ook al een eerste verklaring. Aan dit topic werden de concepten 'WirelessMobile' en 'Organization' gekoppeld. Dit laatste concept 'Organization' komt zeer frequent voor in de tekst en leidt zo, vaak ten onrechte, tot een verhoogde kans op detectie van het topic 'WirelessOrganization'.

Een andere mogelijke verklaring voor de tegenvallende resultaten is misschien juist het feit, dat aan elk topic maar één concept gekoppeld is. Door meerdere concepten aan één topic te koppelen, stijgt de kans natuurlijk sterk dat dit topic voor een paragraaf als relevant beschouwd wordt. Daarom werd een nieuwe analyse uitgevoerd, waarbij dus het aantal concepten verbonden aan één topic, verhoogd werd. Dit werd enerzijds gedaan door simpelweg bijkomende relevante concepten toe te voegen. Anderzijds werden bepaalde topics samengevoegd tot één breder topic. Maar ook de resultaten van deze nieuwe analyse waren niet significant beter.

Een derde reden die de teleurstellende resultaten kan verantwoorden, is misschien de geringe lengte van de paragrafen. De opsplitsing van de tekst in paragrafen volgens het oorspronkelijke boek heeft geleid tot relatief korte stukken tekst. Hoe korter de paragraaf, hoe minder kans dat een bepaald concept in een paragraaf voorkomt. Bijgevolg neemt ook de kans dat er een topic gevonden wordt, af.

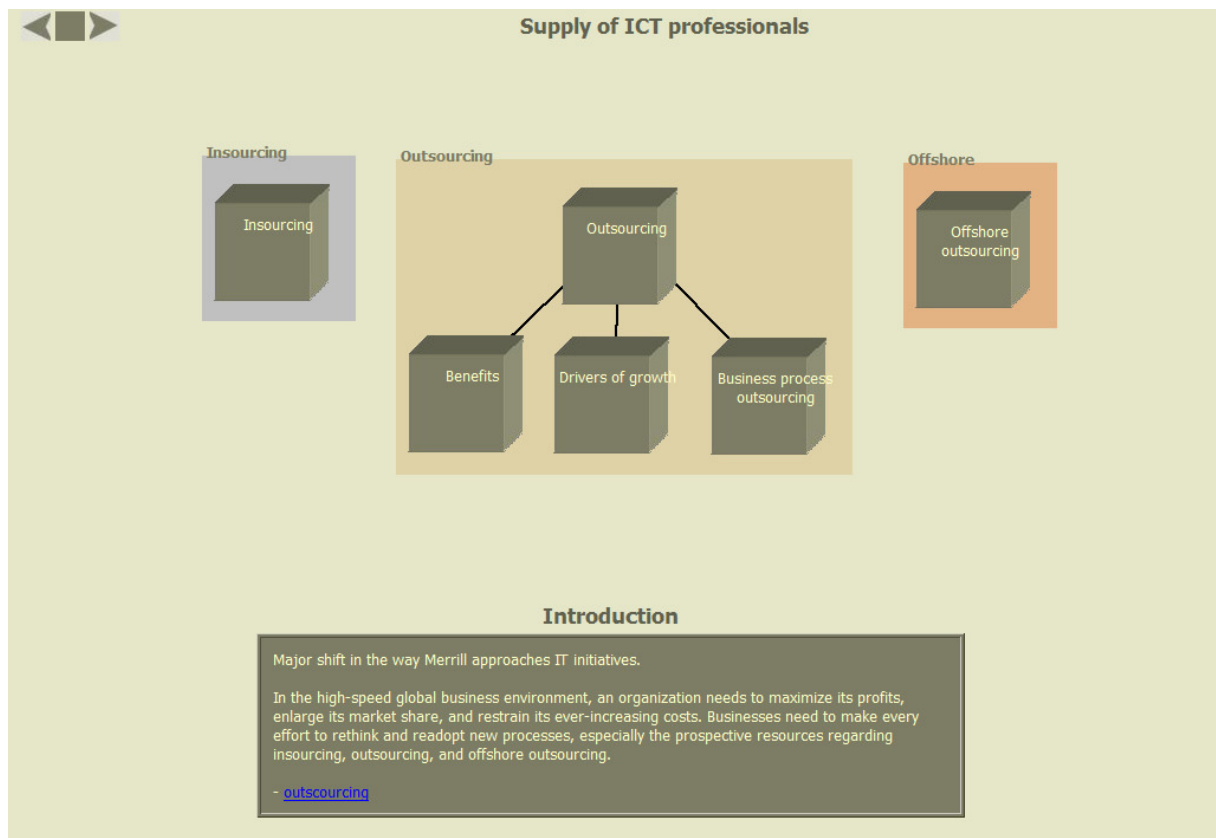
### **5.3 Methode 2**

Vermits de eerste methode zeer povere resultaten opleverde, werd er besloten om de aanpak te veranderen. Dit betekende echter niet, dat al het geleverde werk uit de eerste methode volledig verloren was. De concepten en de koppeling tussen lexicale representaties en concepten, relaties en modifiers uit de vorige werkwijze werden niet uit het systeem verwijderd. Zodoende kon met de gewijzigde aanpak hierop verder gebouwd worden.

### 5.3.1. Uitgangspunt

Het uitgangspunt van deze methode is de volgende: indien we een machine willen laten leren om een mindmap te maken, dienen we te vertrekken van het doorgeven van de kennis die bij het manuele proces gebruikt wordt.

Concreet betekent dit, dat er gestart wordt van een manueel geconstrueerde mindmap. Op basis van deze mindmap werden de topics, concepten en relaties tussen de concepten bepaald en weergegeven in een overzichtelijk schema. Omdat dit proces zeer tijdsintensief is, werd er besloten om te focussen op één hoofdstuk, namelijk 'Chapter 15: Outsourcing Collaborative Partnerships', van de originele tekst 'Business Driven Technologies'. De tekst van dit hoofdstuk is terug te vinden in Bijlage II.



**Figuur 12: Mindmap 'Outsourcing Collaborative Partnerships'**

Figuur 12 toont de originele mindmap en tabel 5 bevat een gedeelte van het ontwikkelde schema. Het volledige schema is weergegeven in Bijlage III. Er kan vastgesteld worden dat de topics overeenkomen met de blokken en subblokken uit de mindmap. Zo vinden we voor de zes blokken uit de mindmap, zes overeenkomstige topics terug in het schema. Op deze manier kan later de hiërarchische structuur van de blokken weerspiegeld worden in de topics-lijst.

**Tabel 5: Schema 'Outsourcing Collaborative Partnerships'**

| Topic      | Concept 1  | Concept 2                | Concept 3 | Rel                       |
|------------|------------|--------------------------|-----------|---------------------------|
| insourcing | insourcing |                          |           |                           |
|            |            | internal ICT work force  | →         | requires                  |
|            |            | supply ICT professionals | ←         | can be done by the way of |
|            |            | IT work                  | →         | includes                  |
|            |            | development of IS        | →         | includes                  |
|            |            | maintenance of IS        | →         | includes                  |

Elk (sub)blok of topic kreeg één hoofdconcept toegewezen en vervolgens verscheidene nevenconcepten. Topic 'Insourcing' bijvoorbeeld werd in de eerste plaats gekoppeld aan het concept 'Insourcing'. Onder dit hoofdconcept bevinden zich drie nevenconcepten, namelijk 'Internal ICT Workforce', 'Supply of ICT professionals' en 'IT Work'. Dit laatste concept kreeg op een tweede subniveau de concepten 'Development of IS' en 'Maintenance of IS' mee.

De relaties werden bepaald door de wijze waarop een concept zich verhoudt tot een concept op een direct lager niveau of in de andere richting. De relatie tussen het concept 'Insourcing' en 'Internal ICT Workforce' werd bestempeld als 'Requires' en loopt, zoals de pijl achter het concept 'Internal ICT Workforce' aangeeft, van het hogerliggende concept 'Insourcing' naar het onderliggende concept 'Internal ICT Workforce'. Concreet impliceert dit dus de volgende relatie 'Insourcing requires Internal ICT Workforce'.

Tot slot bevat de laatste kolom van het schema enkele synoniemen voor een aantal concepten. Deze synoniemen dienen als een eerste aanzet voor het vinden van lexicale representaties van de gebruikte concepten.

### 5.3.2. Concepten

In een volgende stap werd een lijst van de verscheidene concepten uit het schema gehaald. Deze lijst werd vervolgens opgekuist. Zo verwezen oorspronkelijk enkele verschillende conceptnamen naar hetzelfde begrip. Deze verschillende namen werden op elkaar afgestemd en kregen een éénduidige conceptnaam. Een voorbeeld hiervan zijn de concepten 'Cost Reduction' en 'Reduce Costs', die in feite naar hetzelfde begrip verwijzen. Daarnaast kwamen er conceptnamen voor, die sterke gelijkenissen vertoonden met concepten uit de eerste methode. Deze conceptnamen werden

eveneens aan elkaar aangepast, zodat er in het verdere proces gebruik gemaakt kan worden van de reeds aanwezige gegevens in de database voor dit concept.

De nieuwe concepten dienden natuurlijk ook in de ontologie van i.Know ingebracht te worden. In de conceptenstructuur van het schema kwamen een aantal concepten op verschillende plaatsen en op meerdere niveaus terug. Aangezien de ontologie van i.Know vereist, dat elk concept een unieke naam meekrijgt, veroorzaakte dit een probleem. Het was dus niet mogelijk om de originele, uit drie niveaus opgebouwde structuur, te behouden. De oplossing bestond erin, om de hiërarchie van concepten te herleiden naar een lijst met slechts twee niveaus, door niveau één en twee samen te nemen.

Deze aanpassing heeft geen gevolgen voor de werking van het systeem. In principe maakt het immers niet uit, of er met een lijst van concepten of met een boomstructuur gewerkt wordt. De boomstructuur heeft het voordeel, dat ze voor de gebruiker eenvoudiger is om te hanteren. De enige functionaliteit die verloren gaat bij een lijst is, dat het niet mogelijk is om concepten, eigenschappen van hoger geplaatste concepten te laten erven. Maar in het kader van dit werk, was deze functie niet meteen relevant.

Na het definiëren van de concepten, dienen hier eveneens lexicale representaties aan gekoppeld te worden. Hierbij gaan we eigenlijk juist omgekeerd tewerk ten opzichte van methode 1. In de eerste aanpak werd er gestart van niet-herkende woorden, waarvoor vervolgens een geschikt concept gezocht werd. In deze tweede aanpak werd er gestart bij het concept, om vervolgens manueel lexicale representaties te zoeken die naar dit concept verwijzen. Voor het concept 'InternalICTWorkForce' geeft het schema in Bijlage III reeds de volgende synoniemen op: 'in-house ICT services', 'internal ICT experts' en 'internal ICT professionals'. Deze synoniemen werden gebruikt als lexicale representaties voor het concept 'InternalICTWorkForce'. Voor andere concepten werden eventueel nog bijkomende lexicale representaties gezocht. Concreet gebeurt de koppeling tussen de lexicale representaties en het concept binnen het systeem aan de hand van de volgende notatie:

```
in-house ict services;InternalICTWorkForce  
internal ict experts;InternalICTWorkForce  
internal ict professionals;InternalICTWorkForce
```

### **5.3.3. Relaties**

Na de definiëring van de verschillende concepten, kwamen de relaties aan de beurt. Hier werd op dezelfde manier gewerkt als bij de concepten. In een eerste stap werden de verschillende relaties die in het schema gebruikt werden, geïnventariseerd. Vervolgens dienden voor deze relaties ook lexicale representaties gezocht te worden. Het zoeken naar bijpassende woorden en woordgroepen

werd ook ditmaal manueel verricht. Voor de relatie 'Can Deliver' werd bijvoorbeeld de woordgroep 'can give' gevonden, de notatie hiervoor is:

can give;SCanDeliver

Het vinden van goede lexicale representaties voor de relaties uit het schema, was vaak geen eenvoudige taak. Een aantal relaties uit het schema had immers een eerder abstract karakter, waardoor het moeilijk is om er een concrete lexicale representatie voor te vinden.

#### **5.3.4. CRC's**

In een volgende stap werden de in het schema bepaalde concept-relatie-concept-structuren toegevoegd. Al de CRC's werden in het onderstaande formaat genoteerd en vervolgens als input aan de software aangeboden.

Concept;<SRelatie;Concept

Voor de reeds besproken relatie tussen 'Insourcing' en 'InternalICTWorkForce' krijgen we dan de volgende notatie:

Insourcing;<SRequires;InternalICTWorkForce

#### **5.3.5. Topics**

Een laatste stap, alvorens er resultaten gegenereerd kunnen worden, is de definiëring van de topics. De structuur van de topics komt overeen met de topic-hiërarchie uit het opgestelde schema en de mindmap. Voor het behandelde hoofdstuk is er één hoofdtopic, namelijk 'Supply Of ICT Professionals', en daaronder zijn er 3 subtopics ('Insourcing', 'Outsourcing' en 'OffshoreOutsourcing'). Het subtopic 'Outsourcing' kent dan op zijn beurt een opsplitsing in 3 subtopics ('OutsourcingBenefits', 'ITOutsourcingMarket' en 'OffshoreOutsourcing'). Dit geeft de volgende formele notatie van topics:

Topic\_SupplyOfICTProfessionals;Root  
Topic\_Insourcing;Topic\_SupplyOfICTProfessionals  
Topic\_Outsourcing;Topic\_SupplyOfICTProfessionals  
Topic\_OutsourcingBenefits;Topic\_Outsourcing  
Topic\_ITOutsourcingMarket;Topic\_Outsourcing  
Topic\_BusinessProcessOutsourcing;Topic\_Outsourcing  
Topic\_OffshoreOutsourcing;Topic\_SupplyOfICTProfessionals

Elk topic werd vervolgens gekoppeld aan alle concepten die zich in het schema onder dit topic bevinden. De concepten 'Insourcing', 'InternalICTWorkForce', 'SupplyOfICTProfessionals', 'ITWork', 'DevelopmentOfIS' en 'MaintenanceOfIS' werden daarom alle gelinkt aan 'Topic\_Insourcing'.

Alvorens het systeem aan paragrafen één of meerdere topics kan koppelen, is het nodig om de tekst op te splitsen. In de vorige aanpak werd de tekst opgedeeld op basis van de paragrafen zoals



ze in de originele tekst voorkwamen. In de tweede aanpak is ervoor geopteerd, om hier vanaf te stappen en de tekst op te splitsen volgens het behandelde topic. Daarmee komt de opsplitsing van de tekst overeen met de manuele methode voor de mindmap. Voor dit concrete hoofdstuk werd de tekst dus opgesplitst in zeven delen: één deel voor de inleiding en de overige zes komen overeen met de zes gekozen topics.<sup>34</sup>

### 5.3.6. Resultaten topic-herkenning

Bij een eerste analyse van de tekst werd er voor 2 delen geen topic herkend. Vervolgens werden enkele aanpassingen doorgevoerd, om de kans op herkenning van het juiste topic te vergroten. Daartoe werden manueel enkele extra CRC's en enkele bijkomende koppelingen tussen de lexicale representaties en concepten toegevoegd. Enkele onjuiste koppelingen tussen woorden of woordgroepen en bepaalde relaties of modifiers werden verwijderd.

**Tabel 6: Resultaten tekst 1**

| Par. | Origineel topic              | Gevonden topic                    |
|------|------------------------------|-----------------------------------|
| 1    | Inleiding                    | Outsourcing                       |
| 2    | Insourcing                   | <b>Insourcing</b>                 |
| 3    | Outsourcing                  | <b>Outsourcing</b>                |
| 4    | Outsourcing benefits         | <b>OutsourcingBenefits</b>        |
| 5    | IT outsourcing market        | <b>ITOutsourcingMarket</b>        |
| 6    | Business process outsourcing | <b>BusinessProcessOutsourcing</b> |
| 7    | Offshore outsourcing         | <b>OffshoreOutsourcing</b>        |

Tabel 6 bevat de resultaten die na deze aanpassingen bij een tweede analyse gevonden werden. Hieruit blijkt dat voor de paragrafen 2 tot en met 7 het beoogde topic door de software gevonden werd. Het topic 'Outsourcing' werd gekoppeld aan de eerste, inleidende paragraaf. Eventueel kon hier het overkoepelende topic 'SupplyOfICTProfessionals' gevonden worden. De inleiding behandelt een voorbeeld van outsourcing bij Merrill Lynch. Vandaar dat het dus logisch is, dat 'Topic\_Outsourcing' als het meest relevante topic aan het eerste deel toegekend wordt.

---

<sup>34</sup> De nummers aan het begin van een regel in Bijlage II signaleren het begin van een nieuw deel.

### 5.3.7. Tekst 2

Aan de hand van de eerste tekst, hoofdstuk 15 'Outsourcing Collaborative Partnerships' uit het boek 'Business Driven Technologies', werd reeds een deel van de kennis rond outsourcing in het systeem gebracht. De aanwezige kennis werd vervolgens uitgebreid met nieuwe elementen uit de tekst 'Outsourcing Information Systems Functions'<sup>35</sup> (zie Bijlage IV). Voor deze tweede tekst werd het proces, zoals gevolgd bij de eerste tekst, opnieuw doorlopen. In de volgende alinea's wordt dit proces bondig beschreven.

In eerste instantie werd er opnieuw een schema (zie Bijlage V) getekend op basis van de ontwikkelde mindmap voor de tekst. Hieruit werden opnieuw de concepten en relaties gehaald. De concepten en relaties werden indien nodig afgestemd op de concepten uit de eerste tekst. Vervolgens werden voor de nieuwe concepten en relaties enkele lexicale representaties gezocht. Waar nodig werden bij de oude concepten en relaties bijkomende woorden en woordgroepen toegevoegd.

In een volgende stap werden ook de topics uit het schema gehaald en gekoppeld aan de juiste concepten, eveneens op basis van het schema. Verder werd de tekst in vijf delen opgesplitst. De eerste vier delen komen elk overeen met één topic. Het laatste deel is een conclusie van de volledige tekst en werd daarom ook apart gehouden.<sup>36</sup>

Na deze noodzakelijke voorbereiding was het tijd om de nieuwe tekst een eerste maal te analyseren met de software. De resultaten van deze analyse zijn af te lezen in tabel 7: eerste resultaten tekst 2.

**Tabel 7: Eerste resultaten tekst 2**

| Par. | Origineel topic                    | Gevonden topic                    |
|------|------------------------------------|-----------------------------------|
| 1    | Evolution outsourcing IS functions | Geen topic                        |
| 2    | Outsourcing characteristics        | <b>OutsourcingCharacteristics</b> |
| 3    | Special IT outsourcing solutions   | Geen topic                        |
| 4    | Management of outsourcing          | Geen topic                        |
| 5    | Conclusie                          | Geen topic                        |

<sup>35</sup> McNurlin, B.C., Sprague, R.H. Jr. (2006). *Information Systems Management in practice*. Upper Saddle River: Pearson Education, Inc.

<sup>36</sup> De nummers aan het begin van een regel van de tekst in Bijlage IV duiden op het begin van een nieuwe paragraaf.

Hieruit blijkt dat er slechts voor één tekstdeel een topic gevonden kon worden. Het was dus duidelijk, dat er nog enkele aanpassingen dienden te gebeuren. Door i.Know werd er vastgesteld, dat er nauwelijks CRC's gevonden werden met de concepten die als bepalend voor de topics gedefinieerd werden. Daarom was het dus noodzakelijk om meer CRC's toe te voegen.

In de eerste tekst diende men nog manueel te zoeken naar bijkomende CRC's, maar ondertussen was er binnen i.Know een applicatie ontwikkeld, die die deze zoekactie automatisch kon uitvoeren. Deze toepassing genereert een lijst van alle mogelijk CRC's en rangschikt deze volgens relevantie en mate van volledigheid. Een CRC wordt in dit kader als relevant beschouwd, als een concept gekoppeld is aan een topic. De categorieën waarin deze CRC's volgens de toepassing ingedeeld werden, zijn de volgende:

1. Twee bekende concepten, waarvan minstens 1 relevant voor topicbepaling
2. Twee bekende, niet-relevante concepten
3. Relevant eerste concept, onbekend tweede concept
4. Irrelevant eerste concept, onbekend tweede concept
5. Eerste concept onbekend, tweede concept relevant
6. Onbekend eerste concept, irrelevant tweede concept
7. Beide concepten onbekend

De CRC's uit de eerste categorie werden alle toegevoegd. Bij de CRC's uit de overige categorieën werd eerst gekeken of ze zinvol waren. Vervolgens werden alle CRC's die als 'zinvol' beschouwd werden, geanalyseerd op relevantie. De CRC's die uiteindelijk ook als relevant werden gemarkeerd, werden toegevoegd. Voor de CRC's uit de categorieën 3 tot 7, waar het eerste en/of het tweede concept onbekend was, werd de desbetreffende lexicale representatie eerst geannoteerd met het juiste concept.

De vierde categorie, bestaande uit CRC's met een irrelevant eerste concept en een onbekend tweede concept, bevatte ondermeer de CRC in tabel 8.

**Tabel 8: Voorbeeld toegevoegde CRC**

| Lexrep 1 | Concept         | Lexrep 2 | Relatie                     | Lexrep 3     | Concept 2 |
|----------|-----------------|----------|-----------------------------|--------------|-----------|
| software | ComputerProgram | from     | STemporalRelation_SCausedBy | indian firms |           |

Na manuele analyse werd deze CRC als zinvol en relevant bestempeld en diende bijgevolg toegevoegd te worden. Voor de derde lexicale representatie, zie kolom 'Lexrep 3', kon de software geen concept, zie kolom 'Concept 2', terugvinden. Daarom werd de woordgroep 'indian firms' gelinkt aan het concept 'OffshoreOutsourcing', m.a.w. werd de lexicale representatie 'indian firms' geannoteerd met het concept 'OffshoreOutsourcing' en werd de CRC uiteindelijk in de volgende formele notatie toegevoegd:

ComputerProgram;<STemporalRelation\_SCausedBy;OffshoreOutsourcing

Er werden een aantal CRC's toegevoegd die voor het systeem geen relevante concepten bevatten. Dit betekende dus, dat er voor deze CRC's geen van beide concepten gelinkt was aan een topic. Vandaar dat dus ook de definiëring van de topics aangepast werd, zodat deze CRC's wel als relevant beschouwd zouden worden en dus ook een rol zouden spelen bij de bepaling van een topic voor een paragraaf. Met het aangepaste topics-bestand werden de in tabel 9 opgesomde resultaten bekomen.

**Tabel 9: Resultaten tekst 2**

| Par. | Origineel topic                    | Gevonden topic                           |
|------|------------------------------------|------------------------------------------|
| 1    | Evolution outsourcing IS functions | 1 OutsourcingIsFunctions                 |
|      |                                    | 2 <b>EvolutionOutsourcingIsFunctions</b> |
| 2    | Outsourcing characteristics        | 1 OutsourcingIsFunctions                 |
|      |                                    | 2 EvolutionOutsourcingIsFunctions        |
|      |                                    | 3 <b>OutsourcingCharacteristics</b>      |
|      |                                    | 4 SpecialITOutsourcingSolutions          |
| 3    | Special IT outsourcing solutions   | 1 <b>SpecialITOutsourcingSolutions</b>   |
|      |                                    | 2 OutsourcingIsFunctions                 |
|      |                                    | 3 EvolutionOutsourcingIsFunctions        |
|      |                                    | 4 OutsourcingCharacteristics             |
| 4    | Management of outsourcing          | 1 EvolutionOutsourcingIsFunctions        |
|      |                                    | 2 <b>ManagementOfOutsourcing</b>         |
|      |                                    | 3 OutsourcingIsFunctions                 |
| 5    | Conclusie                          | Geen topic                               |

De volgorde waarin de gevonden topics in de tabel weergegeven worden, is een indicatie van de mate van relevantie van dat topic. Ditmaal is er geen *perfect match*, zoals dat het geval was bij de laatste resultaten van tekst 1. Elke paragraaf krijgt, ten laatste in derde instantie, zijn eigen topic toegewezen. Bij de eerste en tweede paragraaf kaapt het overkoepelende topic 'OutsourcingIsFunctions' de eerste plaats weg. Voor de laatste paragraaf met de conclusies wordt er geen topic gevonden. De geringe lengte van dat deel is hier allicht een verklaring voor.

### 5.3.8. Controleteksten topic-herkenning

Om de software nu te testen, werden twee extra teksten over outsourcing gezocht. De twee teksten die werden geselecteerd zijn: 'Outsourcing as an Economic Strategy'<sup>37</sup> en 'Strategic Dimensions of IT Outsourcing'<sup>38</sup>. Beide teksten zijn integraal opgenomen in respectievelijk Bijlage VI en Bijlage VII. Vooraleer de teksten te analyseren, werden de twee lijsten van topics van de eerste twee teksten gecombineerd tot één algemene lijst met definiëringen van topics (zie tabel 10). Vervolgens werden de nieuwe teksten ingedeeld in kleinere delen.<sup>39</sup>

**Tabel 10: Overzicht topics**

| Nr | Topic                           |
|----|---------------------------------|
| 1  | SupplyOfICTProfessionals        |
| 2  | Insourcing                      |
| 3  | Outsourcing                     |
| 4  | OutsourcingBenefits             |
| 5  | ITOutsourcingMarket             |
| 6  | BusinessProcessOutsourcing      |
| 7  | OffshoreOutsourcing             |
| 8  | OutsourcingIsFunctions          |
| 9  | EvolutionOutsourcingIsFunctions |
| 10 | OutsourcingCharacteristics      |
| 11 | SpecialITOutsourcingSolutions   |
| 12 | ManagementOfOutsourcing         |

Om de resultaten op een correcte manier te kunnen bestuderen, werden teksten 1 en 2 ook opnieuw geanalyseerd op basis van de samengevoegde topics-lijst. Tabel 11 en tabel 12 bevatten de resultaten van de nieuwe analyse voor respectievelijk tekst 1 en tekst 2.

---

<sup>37</sup> Turban, E., McLean, E., & Wetherbe, J. (2002). *Information Technology for Management*. New York: John Wiley & Sons, Inc.

<sup>38</sup> Marchand, D.A., & Davenport, T.H. (z.d.) *Mastering Information Management*. London: Prentice Hall.

<sup>39</sup> De indeling is ook nu in Bijlage VI en Bijlage VII weergegeven aan de hand van een nummer in het begin van de regel.

**Tabel 11: Definitieve resultaten tekst 1**

| Par. | Origineel topic              | Gevonden topic                      | Rel. |
|------|------------------------------|-------------------------------------|------|
| 1    | Inleiding                    | 1 Outsourcing                       | 0.69 |
|      |                              | 2 OutsourcingBenefits               | 0.29 |
|      |                              | 3 ITOutsourcingMarket               | 0.29 |
|      |                              | 4 BusinessProcessOutsourcing        | 0.29 |
| 2    | Insourcing                   | 1 <b>Insourcing</b>                 | 3.32 |
|      |                              | 2 SupplyOfICTProfessionals          | 1.07 |
|      |                              | 3 Outsourcing                       | 1.07 |
|      |                              | 4 OffshoreOutsourcing               | 1.07 |
| 3    | Outsourcing                  | 1 <b>Outsourcing</b>                | 0.96 |
| 4    | Outsourcing benefits         | 1 <b>OutsourcingBenefits</b>        | 0.74 |
|      |                              | 2 Outsourcing                       | 0.42 |
|      |                              | 3 ITOutsourcingMarket               | 0.42 |
|      |                              | 4 BusinessProcessOutsourcing        | 0.42 |
| 5    | IT outsourcing market        | Geen topic gevonden                 |      |
| 6    | Business process outsourcing | 1 <b>BusinessProcessOutsourcing</b> | 2.50 |
|      |                              | 2 OutsourcingCharacteristics        | 1.50 |
|      |                              | 3 Outsourcing                       | 0.45 |
| 7    | Offshore outsourcing         | 1 <b>OffshoreOutsourcing</b>        | 6.16 |
|      |                              | 2 SpecialITOutsourcingSolutions     | 2.17 |
|      |                              | 3 ITSourcingMarket                  | 1.30 |

**Tabel 12: Definitieve resultaten tekst 2**

| Par. | Origineel topic                    | Gevonden topic                           | Rel. |
|------|------------------------------------|------------------------------------------|------|
| 1    | Evolution outsourcing IS functions | 1 Outsourcing                            | 1.67 |
|      |                                    | 2 ITOutsourcingMarket                    | 0.25 |
|      |                                    | 3 OutsourcingBenefits                    | 0.21 |
|      |                                    | 4 BusinessProcessOutsourcing             | 0.21 |
|      |                                    | 5 <b>EvolutionOutsourcingIsFunctions</b> | 0.10 |
| 2    | Outsourcing characteristics        | 1 Outsourcing                            | 1.46 |
|      |                                    | 2 BusinessProcessOutsourcing             | 0.95 |
|      |                                    | 3 OutsourcingBenefits                    | 0.89 |
|      |                                    | 4 ITOutsourcingMarket                    | 0.89 |
|      |                                    | 5 <b>OutsourcingCharacteristics</b>      | 0.69 |
|      |                                    | 6 EvolutionOutsourcingIsFunctions        | 0.05 |
|      |                                    | 7 SpecialITOutsourcingSolutions          | 0.04 |
|      |                                    | 8 OffshoreOutsourcing                    | 0.04 |
| 3    | Special IT outsourcing solutions   | 1 Outsourcing                            | 0.56 |
|      |                                    | 2 <b>SpecialITOutsourcingSolutions</b>   | 0.56 |
|      |                                    | 3 OffshoreOutsourcing                    | 0.51 |
|      |                                    | 4 BusinessProcessOutsourcing             | 0.31 |
|      |                                    | 5 OutsourcingBenefits                    | 0.30 |
|      |                                    | 6 ITOutsourcingMarket                    | 0.30 |
|      |                                    | 7 OutsourcingCharacteristics             | 0.01 |
| 4    | Management of outsourcing          | 1 Outsourcing                            | 0.86 |
|      |                                    | 2 <b>ManagementofOutsourcing</b>         | 0.17 |
|      |                                    | 3 OutsourcingBenefits                    | 0.03 |
|      |                                    | 4 ITSourcingMarket                       | 0.03 |
|      |                                    | 5 BusinessProcessOutsourcing             | 0.03 |
| 5    | Conclusie                          | Geen topic                               |      |

Deze definitieve resultaten voor tekst 1 en tekst 2 wijken lichtjes af ten opzichte van de resultaten weergegeven in tabel 6 en tabel 9. De verklaring hiervoor is, dat er ondertussen ook gesleuteld is aan de kennisbank. Er zijn onder andere lexicale representaties van concepten en relaties toegevoegd en verwijderd. De lijst van CRC's is bovendien ook aangepast. Maar de grootste verandering is natuurlijk, het samenvoegen van de topic-lijsten van beide teksten. Vandaar dus het optreden van enkele verschillen ten opzichte van de eerste resultaten, maar in het algemeen zijn de bevindingen gelijk aan de eerste bevindingen.

De resultaten voor de teksten 3 en 4 zijn terug te vinden in tabel 13 en tabel 14. De bepaling van het originele topic voor de laatste twee teksten is niet gebeurd op basis van een mindmap, zoals bij de twee vorige. De twee teksten werden doorlopen op basis van de topics uit tabel 10 en aan elke paragraaf werden de meest toepasselijke topics toegekend. Dit toekennen van topics is uiteraard tot op een zekere hoogte afhankelijk van persoon tot persoon. Maar het geeft zeker en vast een eerste indicatie of de gevonden topics de goede richting uit gaan.

**Tabel 13: Definitieve resultaten tekst 3**

| Par. | Origineel topic                   | Gevonden topic               | Rel. |
|------|-----------------------------------|------------------------------|------|
| 1    | 1 Outsourcing                     | 1 <b>Outsourcing</b>         | 1.04 |
|      | 2 EvolutionOutsourcingISFunctions | 2 OutsourcingBenefits        | 0.77 |
|      |                                   | 3 OutsourcingMarket          | 0.77 |
|      |                                   | 4 BusinessProcessOutsourcing | 0.77 |
| 2    | 1 SpecialITOutsourcingSolutions   | Geen topic gevonden          |      |
|      | 2 OffshoreOutsourcing             |                              |      |
| 3    | 1 ManagementOfOutsourcing         | 1 Outsourcing                | 1.00 |
|      |                                   | 2 OutsourcingBenefits        | 0.53 |
|      |                                   | 3 ITOutsourcingMarket        | 0.53 |
|      |                                   | 4 BusinessProcessOutsourcing | 0.53 |
| 4    | 1 Outsourcing                     | 1 <b>Outsourcing</b>         | 7.37 |

Tekst 3 bekijkt het gebruik van outsourcing als een economische strategie. In verband met de automatische topic-herkenning voor deze tekst kunnen we concluderen, dat de resultaten voor 50% correct zijn. Paragrafen 1 en 4 kregen beide bij de manuele analyse het topic 'Outsourcing' mee en dit topic kwam ook bij de automatische analyse op de eerste plaats. In paragraaf 2 gaat het zeer bondig over enkele outsourcing-oplossingen, met ondermeer 'offshore outsourcing'. Buiten 'offshore outsourcing' en een verwijzing naar 'business process outsourcing' worden er andere



oplossingen voorgesteld, dan deze die behandeld zijn in teksten 1 en 2. Deze twee elementen kunnen een verklaring zijn voor het feit, dat de software geen topic kan toewijzen aan de paragraaf. Voor paragraaf 3 werden aan de hand van het systeem een aantal topics gevonden, maar niet het voorgestelde topic 'ManagementOfOutsourcing'. Deze paragraaf handelt over de risicobeheersing bij outsourcing. Hierbij zijn er enkele overeenkomsten met het topic 'ManagementOfOutsourcing' uit tekst 2, maar deze paragraaf gaat eigenlijk specifiek over de risico's en deze komen in de vorige teksten niet zo duidelijk aan bod.

**Tabel 14: Definitieve resultaten tekst 4**

| Par. | Origineel topic              | Gevonden topic                    | Rel. |
|------|------------------------------|-----------------------------------|------|
| 1    | 1 Outsourcing                | 1 <b>ITOutsourcingMarket</b>      | 1.08 |
|      | 2 ITOutsourcingMarket        | 2 EvolutionOutsourcingIsFunctions | 1.08 |
|      |                              | 3 <b>Outsourcing</b>              | 0.91 |
|      |                              | 4 OutsourcingBenefits             | 0.65 |
|      |                              | 5 BusinessProcessOutsourcing      | 0.65 |
| 2    | 1 Outsourcing                | 1 <b>Outsourcing</b>              | 1.08 |
|      | 2 OutsourcingBenefits        | 2 <b>OutsourcingBenefits</b>      | 0.10 |
|      | 3 OutsourcingCharacteristics | 3 ITOutsourcingMarket             | 0.10 |
|      |                              | 4 BusinessProcessOutsourcing      | 0.10 |
| 3    | 1 Outsourcing                | 1 <b>Outsourcing</b>              | 1.03 |
|      | 2 OutsourcingBenefits        | 2 <b>OutsourcingBenefits</b>      | 1.03 |
|      | 3 OutsourcingCharacteristics | 3 ITOutsourcingMarket             | 1.03 |
|      |                              | 4 BusinessProcessOutsourcing      | 1.03 |
| 4    | 1 Outsourcing                | 1 <b>Outsourcing</b>              | 0.28 |
|      | 2 OutsourcingBenefits        | 2 <b>OutsourcingBenefits</b>      | 0.28 |
|      |                              | 3 ITOutsourcingMarket             | 0.28 |
|      |                              | 4 BusinessProcessOutsourcing      | 0.28 |
| 5    | 1 Outsourcing                | Geen topic                        |      |
|      | 2 OutsourcingBenefits        |                                   |      |
| 6    | 1 Outsourcing                | 1 <b>Outsourcing</b>              | 0.29 |
|      | 2 OutsourcingBenefits        |                                   |      |

| Par. | Origineel topic           |   | Gevonden topic             | Rel. |
|------|---------------------------|---|----------------------------|------|
| 7    | 1 Outsourcing             | 1 | <b>Outsourcing</b>         | 1.23 |
|      | 2 OutsourcingBenefits     | 2 | <b>OutsourcingBenefits</b> | 0.41 |
|      |                           | 3 | ITOutsourcingMarket        | 0.41 |
|      |                           | 4 | BusinessProcessOutsourcing | 0.41 |
| 8    | 1 Outsourcing             | 1 | <b>Outsourcing</b>         | 0.62 |
|      | 2 OutsourcingBenefits     |   |                            |      |
| 9    | 1 Outsourcing             | 1 | <b>Outsourcing</b>         | 1.65 |
|      | 2 OutsourcingBenefits     |   |                            |      |
| 10   | 1 Outsourcing             | 1 | <b>Outsourcing</b>         | 1.82 |
|      | 2 ManagementOfOutsourcing | 2 | OutsourcingBenefits        | 0.66 |
|      |                           | 3 | ITOutsourcingMarket        | 0.66 |
|      |                           | 4 | BusinessProcessOutsourcing | 0.66 |
| 11   | 1 Outsourcing             |   | Geen topic                 |      |
|      | 2 OutsourcingBenefits     |   |                            |      |

Net als tekst 3, behandelt tekst 4 outsourcing op een strategisch niveau. In de tekst worden een aantal strategische voordelen beschouwd, die in tekst 1 en 2 nauwelijks of niet aan bod komen. Hierdoor zijn de verschillen tussen de verschillende paragrafen van deze laatste tekst ook niet weerspiegeld in de topics-lijst. Dit uit zich zowel bij de manuele als de geautomatiseerde analyse in het feit dat er relatief weinig variatie zit in de topics die als relevant beschouwd worden. De vergelijking van de beide analyses leert, dat de digitale analyse toch een goede benadering is van de manuele analyse.

Het valt op dat een cluster van topics meermaals terugkeert, zowel in tekst 3 als in tekst 4. Het gaat namelijk over de volgende topics: 'Outsourcing', 'OutsourcingBenefits', 'ITOutsourcingMarket' en 'BusinessProcessOutsourcing'. De hoofdreden hiervoor is te zoeken in de koppeling tussen de topics en de concepten. Voor elk van deze vier topics is 'Outsourcing' als een relevant concept gedefinieerd. Aangezien de gevalstudie toegespitst is op teksten met als onderwerp 'outsourcing', is het logisch dat dit concept zeer vaak voorkomt. Topics waarvoor 'Outsourcing' als een belangrijk concept aangeduid is, hebben bijgevolg een grotere kans om als relevant voor een bepaalde paragraaf beschouwd te worden.

### **5.3.9. Samenvattingen**

Naast het herkennen van topics werd er ook een samenvatting gemaakt van de verschillende teksten. Voor elke tekst is er een samenvatting van de volledige tekst, alsook een aparte samenvatting van elke paragraaf. Bijlagen VIII tot en met XI bevatten de samenvattingen van de vier gebruikte teksten in de gevalstudie. De parameter voor de lengte van de samenvattingen werd, voor zowel de samenvatting van de volledige tekst als deze per paragraaf, ingesteld op 25%.

De generatie van de samenvattingen is gebaseerd op de extractie van zinnen. Dit houdt in, dat de meest relevante zinnen uit de originele tekst volledig opgenomen worden in de samenvatting. Enerzijds wordt er rekening gehouden met de frequentie waarmee concepten voorkomen. Zinnen die lexicale representaties bevatten van concepten die vaak voorkomen, hebben op deze manier een streepje voor op de andere. Anderzijds speelt ook hier de relatie tot andere concepten een zeer belangrijke rol. Net als bij topic-herkenning heeft het achterliggende netwerk van CRC's een grote invloed op de relevantie. Vandaar is het ook zo, dat de zinnen die voor een paragraaf als relevant aanzien worden, niet noodzakelijk opgenomen zijn in de samenvatting van de volledige tekst en vice versa.

Hierbij aansluitend kan er ook geen verband gelegd worden tussen het feit of er voor een bepaalde paragraaf een topic gevonden wordt en het voorkomen van zinnen uit die paragraaf in de volledige samenvatting. Van paragraaf 5 uit tekst 1, paragraaf 2 uit tekst 3 en paragraaf 11 uit tekst 4 komen er geen zinnen voor in de samenvatting en wordt er ook geen topic gevonden. Maar de paragrafen 5 uit tekst 2 en paragraaf 5 uit tekst 4 kregen ook geen topic toegewezen, maar hiervan werden wél zinnen in de volledige samenvatting gebruikt.

De kwaliteit van de samenvattingen zelf is middelmatig. De overgangen zijn niet zo vloeiend, daar er vaak van het ene onderwerp onmiddellijk naar het andere gesprongen wordt. Maar dit is nu eenmaal onvermijdelijk, indien men de extractie-methode gebruikt. Het belangrijkste bij een samenvatting is natuurlijk, dat ze de kern van de tekst weerspiegelt. Wat juist tot de essentie behoort, is in beperkte mate ook een subjectief gegeven. Daarbij moet dan wel vastgesteld worden, dat het belangrijkste vaak niet volledig weergegeven wordt. Een bepaald argument wordt soms aangehaald, maar het gebeurt niet helemaal volledig. Allicht is dit ook deels het gevolg van de extractie-methode. Als de auteur een idee of argument over verschillende zinnen verspreid heeft, dan zal, als niet alle zinnen omtrent dat idee als relevant beschouwd worden, dit slechts gedeeltelijk in de samenvatting teruggevonden worden.

Daarnaast valt het op, dat er voor een aantal begrippen eerder een voorbeeld in de samenvatting staat, dan een definiëring of verklaring ervan. Wat de voorkeur geniet, is misschien eveneens van

subjectieve aard. Om dit op te lossen, kan bijvoorbeeld aan eerder definiërende relaties een groter gewicht toegekend worden. Op die manier neemt de kans toe, dat een zin met een concept uit een CRC met een definiërende relatie, in de uiteindelijke samenvatting terecht komt.

### 5.3.10. Sleutelwoorden

De generatie van sleutelwoorden is een laatste stap in de gevalstudie. 'Smart indexing', de techniek die hiervoor gebruikt werd, is pas recent door i.Know ontwikkeld. Traditionele methoden voor het maken van een lijst van sleutelwoorden, splitsen een zin eerst op in afzonderlijke woorden. Vervolgens worden stopwoorden verwijderd en de woorden die overblijven, worden als zogenaamde sleutelwoorden weergegeven. De 'Smart indexing'-technologie daarentegen is in staat, om zonder enige domein-specifieke kennis, uit de zin eerst concepten en relaties te herkennen. Dit betekent, dat samenhangende woordgroepen ook samen als één concept herkend worden, terwijl deze met traditionele methoden in aparte sleutelwoorden opgesplitst worden. Vervolgens worden de woordgroepen die als concept herkend werden, ontdaan van de overbodige stopwoorden. De woorden die uiteindelijk overblijven, zijn dan de sleutelwoorden.

Nemen we een voorbeeld uit de hieronder afgedrukte derde paragraaf van tekst 1.

"**OUTSOURCING**

**Outsourcing** is an arrangement by which one organization provides a service or services for another organization that chooses not to perform them in-house. In some cases, the entire information technology department is outsourced, including planning and business analysis as well as the installation, management, and servicing of the network and workstations. Outsourcing can range from a large contract under which an organization like IBM manages IT services for a company such as Xerox, to the practice of hiring contractors and temporary office workers on an individual basis. Refer to Figure 4.16 for a comparison of the functions companies have outsourced."

De 'Smart indexing'-technologie haalde hier de volgende sleutelwoorden uit:

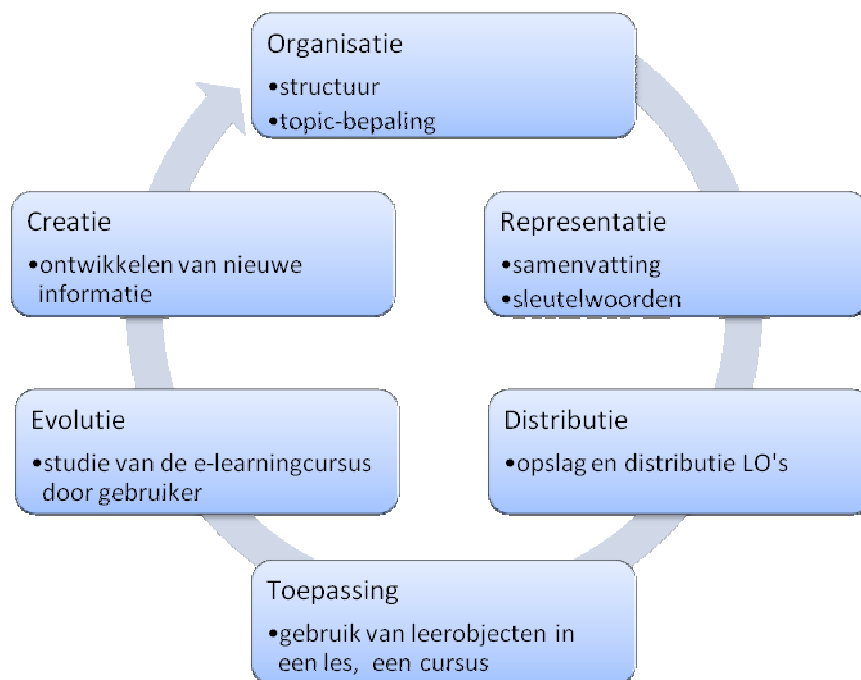
|                |                |                    |
|----------------|----------------|--------------------|
| arrangement;1  | range;1        | hiring             |
| cases;1        | refer;1        | contractors;1      |
| company;1      | service;1      | individual basis;1 |
| comparison;1   | services;2     | large contract;1   |
| ibm;1          | servicing;1    | one organization;1 |
| in-house;1     | workstations;1 | outsourcing        |
| installation;1 | xerox;1        | outsourcing;1      |
| management;1   | business       | temporary office   |
| network;1      | analysis;1     | workers;1          |
| organization;2 | figure 416;1   | entire information |
| outsourcing;1  | functions      | technology         |
| planning;1     | companies;1    | department;1       |
| practice;1     |                |                    |

Het voorlaatste sleutelwoord, of misschien beter gezegd 'sleutelwoordgroep', 'temporary office workers' toont duidelijk aan, dat er rekening gehouden wordt met woordgroepen die samenhangen. De traditionele methoden zouden deze woordgroep opsplitsen in de aparte woorden: 'temporary', 'office' en 'workers'. Hetgeen niet de context weerspiegelt en dus uiteindelijk een veel grotere ruimte voor interpretatie laat.

## 5.4 Conclusie gevalstudie

Het opzet van deze gevalstudie was, om de decompositiefase van de ontwikkeling van een e-learningcursus te automatiseren. Vooraleer specifieke conclusies te trekken voor de decompositiefase wordt de ontwikkeling van een e-learningcursus in het licht van de kennislevenscyclus bekeken.

De kennislevenscyclus werd in het vorige hoofdstuk reeds gebruikt om de functionaliteit van de technologie van i.Know te bespreken. De ontwikkeling van een e-learningcursus is ook een proces waarbij er kennis gegenereerd wordt, bijgevolg is het mogelijk om dit ook aan de hand van deze levenscyclus te bekijken (zie figuur 13).



**Figuur 13: De ontwikkeling van een e-learningcursus aan de hand van de kennislevenscyclus**

In de eerste fase, organisatie, komt het erop aan om in beschikbare gegevens een structuur aan te brengen. In dit geval betekent dat dus, om de tekst op te splitsen in coherente delen en voor elk deel het juiste topic te bepalen. De functie uit de technologie van i.Know die hiervoor gebruikt

wordt, is int.for<sup>®</sup> Categorization. Vervolgens dient in de lijst van de gevonden topics een logische structuur aangebracht te worden.

De representatie-fase zorgt voor een voorstelling van de structuur uit de eerste fase. De samenvattingen en sleutelwoorden zijn hiervan twee voorbeelden uit de gevalstudie. De samenvattingen kunnen aangemaakt worden met behulp van de int.for<sup>®</sup> Summarization en de sleutelwoorden op basis van de recent ontwikkelde 'Smart Indexing'-technologie.

De gegevens uit de representatie-fase kunnen in de distributie-fase opgeslagen worden en verspreid als leerobjecten. In de vierde fase, toepassing, is het mogelijk om de verscheidene LO's te combineren tot een les of cursus rond één bepaald topic of een collectie van verwante topics. Evolutie is de voorlaatste stap in de kennislevenscyclus. De fase komt overeen met de studie of het raadplegen van de e-learningcursus, bijvoorbeeld een mindmap, door de gebruiker. De lerende kan bijvoorbeeld elementen uit de e-learningcursus combineren met andere, externe bronnen en dit kan leiden tot de laatste fase, waarin nieuwe informatie aangemaakt wordt. Met deze nieuwe informatie kan eventueel een nieuwe kenniscyclus opgestart worden.

Het voorgaande toont aan, dat de introductie van text mining-technieken in het ontwikkelingsproces van een e-learningcursus een zeker theoretisch potentieel heeft. Realiteit is wel, dat daarvoor nog een lange weg af te leggen is, maar de vooruitzichten althans zijn boeiend.

Deze gevalstudie behandelt weliswaar enkel de decompositie-fase. Deze eerste stap in de ontwikkeling van een e-learningcursus komt overeen met de eerste twee fases uit de juist besproken kennislevenscyclus, namelijk 'Organisatie' en 'Representatie'.

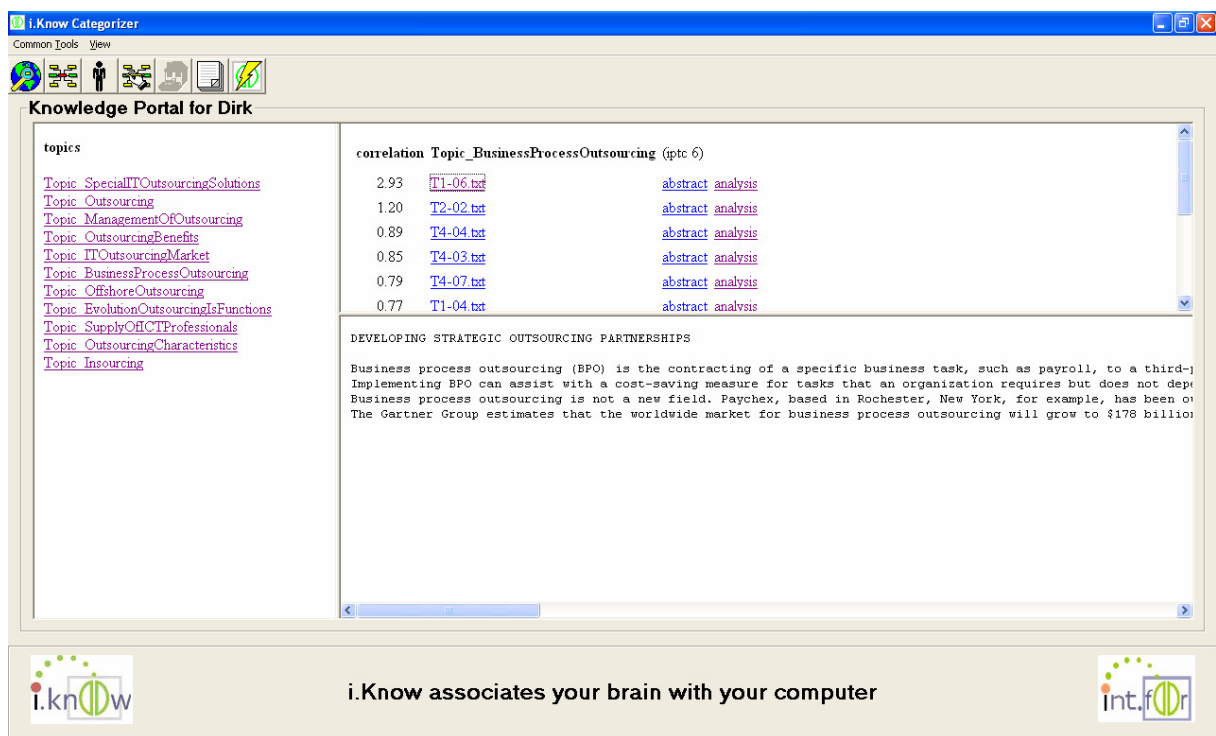
De 'Organisatie'-fase is zonder twijfel de meest belangrijke stap in het decompositie-proces. Een goede basis is immers onontbeerlijk om voor een degelijke kwaliteit in de volgende stappen en fasen te zorgen. In de gevalstudie werd aangetoond dat, mits er een redelijke hoeveelheid kennis over een onderwerp, in dit geval outsourcing, aan het platform van i.Know toegevoegd wordt, de software in staat is, om aan een paragraaf een relevant topic te koppelen. Het onderscheid tussen verscheidene vormen van outsourcing, zoals 'insourcing', 'business process outsourcing' en 'offshore outsourcing' kan goed gemaakt worden met de inbreng van relatief weinig kennis. De reden hiervoor is, dat deze vormen gebruik maken van duidelijk te onderscheiden specifieke termen. Het verschil tussen de topics als 'Outsourcing benefits' en 'Outsourcing characteristics' is bijgevolg een stuk moeilijker, aangezien de nuance ertussen kleiner is.

In de resultaten was er sprake van 'overgeneratie' van topics. Om nu de resultaten inzake topic-bepaling te verbeteren, dient het hele proces verder te worden verfijnd. Deze 'tuning' kan gebeuren, door de koppeling tussen de lexicale relaties en concepten en relaties opnieuw te

bekijken. Foute of overtollige lexicale representaties dienen dan te worden verwijderd en ontbrekende woorden of woordgroepen kunnen toegevoegd worden. Een tweede methode om de resultaten te verbeteren, is het aanpassen van de CRC's. Hier geldt eigenlijk hetzelfde: verwijderen van foute of overtollige CRC's en het toevoegen van relevante concept-relatie-concept-combinaties.

Een belangrijke opmerking voor de organisatie-fase is het feit, dat het nog niet mogelijk is om een tekst automatisch op te splitsen volgens zijn aparte topics. Hiervoor zijn wel technologieën in ontwikkeling, maar deze werden in deze gevalstudie niet behandeld.

Bovendien moet de structuur ook nog voorgedefinieerd worden. In de tweede methode is de structuur van de eerste twee mindmappen weerspiegeld in de hiërarchie van de topics. Om de software rekening te laten houden met de structuur, moet er op voorhand veel domein-specifieke kennis aanwezig zijn. Daarentegen kan men ook de door de software aan de tekstdelen toegekende topics gebruiken om manueel een structuur van de topics te ontwerpen; hiervoor is de beschikbaarheid van kennis en de inbreng ervan veel minder noodzakelijk.



**Figuur 14: Schermafdruk relevante paragrafen per topic**

De huidige resultaten van de herkenning van topics, kunnen aangewend worden om uit verscheidene teksten, delen omtrent hetzelfde onderwerp te selecteren. De schermafdruk in figuur 14 toont bijvoorbeeld de paragrafen uit de verschillende teksten die te maken hebben met het topic 'BusinessProcessOutsourcing'. Op deze manier kan snel en eenvoudig de beschikbare kennis rond een topic opgevraagd worden.

In de representatie-fase draait het in de praktijkstudie voornamelijk rond de generatie van automatische samenvattingen en van lijsten met sleutelwoorden. De momenteel gehaalde resultaten zijn nog van onvoldoende kwaliteit om bijvoorbeeld de samenvattingen dadelijk in de praktijk te gebruiken. Hiervoor dient het systeem eerst te worden verfijnd. Het is wel reeds perfect mogelijk om deze syntheses als hulpmiddel te gebruiken bij het decomposeren. Ook de sleutelwoorden kunnen bij het decompositieproces een ondersteunende rol aannemen. Ze kunnen helpen bij de bepaling van de topics en de ontwikkeling van een decompositie-structuur. Daarnaast kunnen ze ook gebruikt worden als een atomisch leerobject met 'key terms'. Hiervoor dient op basis van de huidige resultaten eerst nog een bijkomende selectie gemaakt te worden.

Uit de gevalstudie kunnen we besluiten dat, op basis van de huidige resultaten, de integratie van text mining-technieken in het decompositieproces enkel een ondersteunde functie kan hebben. Het is echter wel mogelijk om de automatisatie te verhogen, indien het systeem verder verfijnd wordt.



## **HOOFDSTUK 6 : Conclusies**

### **6.1 Algemene conclusies**

Vandaag de dag is er een enorme hoeveelheid aan informatie beschikbaar. Het grote probleem bestaat erin, om deze informatie te analyseren en er kennis uit te halen. Bedrijven kunnen de beschikbare informatie aanwenden om eventueel een strategisch voordeel te verwerven. Onderwijsinstellingen dienen deze berg aan informatie op een aantal manieren te verwerken. Eén ervan is het creëren van een e-learningcursus om deze informatie te kunnen doorgeven. De ontwikkeling van een e-learningcursus is een tijdsintensief proces. Het opzet van deze eindverhandeling bestond er dan ook uit, om na te gaan op welke wijze text mining-technieken geïntegreerd konden worden in de ontwikkeling van een dergelijke e-learningcursus.

Om op deze vraag een antwoord te kunnen geven, werd in de eerste plaats een literatuurstudie verricht. Hierbij werd eerst de evolutie naar nieuwe onderwijsconcepten in e-learning onder de loep genomen. De niet-traditionele onderwijsvormen, met bovenaan het afstandslernen, hebben in de geschiedenis steeds snel nieuwe technologieën geïntroduceerd, met e-learning was dit niet anders. E-learning kende eerst een pijlsnelle groei, waardoor er hoge verwachtingen gecreëerd werden. Toen bleek dat deze verwachtingen niet konden waargemaakt worden, ging het minstens even snel bergaf en verzeilde e-learningsector in een diepe crisis. Vervolgens brak er een periode van gestage groei aan, die nu nog steeds voortduurt. Tijdens deze evolutie werd de 'lerende' zelf steeds meer op het voorplan geplaatst. Recent tekent zich met e-learning 2.0 een nieuwe trend af, sommigen durven zelfs spreken van een opkomende revolutie. Microcontent is hét nieuwe woord bij deze Web 2.0-toepassingen. Microcontent duidt op kleine elementen die door de gebruikers van toepassingen samen gecreëerd worden. De eenvoudigheid waarmee gebruikers zelf microcontent kunnen aanmaken, herbruiken, citeren, bewerken, linken, enz. versterkt de koerswijziging van de oude situatie, met de kennisoverdracht als centraal punt, naar de nieuwe, waarbij de lerende voorop staat.

Leerobjecten vormden het onderwerp van het volgend hoofdstuk van de literatuurstudie. Daarin werd vastgesteld, dat de leerobjecten omwille van hun ideale mix tussen herbruikbaarheid en gespecificeerde inhoud, de centrale elementen zijn tijdens de ontwikkeling van een e-learningcursus. Dit ontwikkelingsproces valt in te delen in twee grote onderdelen. Het eerste onderdeel is het decompositieproces, waarbij leerstof tot zijn atomische objecten ontleed wordt. De tweede fase bouwt een e-learningcursus op aan de hand van een uitgekozen scenario.

Dit eerste luik van de literatuurstudie liet toe, om een goed beeld te krijgen van e-learning zelf en op welke manier een e-learningcursus tot stand komt. Vervolgens werd er een verkennend

literatuuronderzoek uitgevoerd naar text mining en aanverwante domeinen. Daaruit kan worden geconcludeerd, dat text mining en natural language processing steeds meer naar elkaar toegroeien en op bepaalde punten elkaar reeds overlappen. Uit het text mining-kader blijkt, dat information extraction het belangrijkste onderdeel is van text mining. De technologie van i.Know start vanuit een relationeel semantische benadering. De kern en tevens kracht van hun software zijn de concept-relatie-concept-structuren of CRC's. Kenmerkend hiervoor is, dat de concepten niet als afzonderlijk beschouwd worden, maar dat deze altijd in relatie tot andere concepten geanalyseerd worden. De benadering met CRC's laat toe, om aan de hand van de voorkomende relaties van een bepaald woord of woordgroep, een betere afweging te maken, welke interpretatie er relevant is en bijgevolg welk concept van toepassing is. Met de deze technologie is een uitgebreide functionaliteit mogelijk.

Het doel van de gevalstudie bestond eruit na te gaan welke technologieën uit text mining en natural language processing in de decompositiefase van de ontwikkeling van een e-learningcursus kunnen worden geïntegreerd. Meer bepaald omvatte het praktijkprobleem de integratie van de text mining-technieken van i.Know in de decompositie-fase van de ontwikkeling van een e-learningcursus. Concreet werd er geprobeerd om, voor paragrafen uit verscheidene teksten, de relevante topics te bepalen, samenvattingen te genereren en een lijst van sleutelwoorden samen te stellen.

De topic-herkenning is een eerste aanzet om automatisch een structuur van de tekst te maken. Het herkennen van topics werkt behoorlijk, maar dient zeker nog verfijnd te worden. Om tot een structuur te komen zal er ook een methode gevonden moeten worden waarbij de software zelfstandig het begin en het einde van een paragraaf over één topic kan bepalen.

Daarnaast werden er ook samenvattingen en sleutelwoorden gegenereerd. De kwaliteit van de samenvattingen is nog niet optimaal, maar ook hier kunnen verbeteringen worden aangebracht. De lijst van de sleutelwoorden is nog te lang om in het kader van e-learning meteen als *key terms* te worden gebruikt. Maar hieruit kunnen wel de dadelijk de meest belangrijke sleutelwoorden worden geselecteerd.

De resultaten van de gevalstudie zijn dus momenteel nog van een middelmatige kwaliteit en dienen zeker nog verfijnd te worden. Op dit moment beschikt men over een drietal middelen dat het manuele decompositieproces zeker kan vooruit helpen. De topic-herkenning, de samenvattingen en de sleutelwoorden kunnen alledrie het ontwikkelen van een structuur voor een tekst versnellen.

Voorlopig moet men het besluit trekken, dat de technologie van i.Know slechts een ondersteunende functie inneemt tijdens de decompositie-fase van de creatie van een

e-learningcursus innemen kan. Maar met de gesuggereerde aanpassingen moet de automatisatiegraad van het decompositieproces zeker verhoogd kunnen worden.

## **6.2 Suggesties voor verder onderzoek**

In de conclusies van de gevalstudie wordt het gebruik van text mining in e-learning voornamelijk een helpende rol toebedeeld. Er zijn echter wel vooruitzichten dat een verdere integratie van text mining, deze rol aanzienlijk zou kunnen uitbreiden. Enkele mogelijke, kleine aanpassingen werden reeds voorgesteld in deze conclusies, maar er zijn nog een aantal pistes die ook onderzocht dienen te worden.

Een eerste mogelijkheid waarnaar verder onderzoek dient te gebeuren, is de stelling, dat de indien lijsten met sleutelwoorden als basis gebruikt worden, de resultaten verbeteren. Aan de hand van deze lijst kan men enerzijds het algemene topic van een paragraaf bepalen en anderzijds de gebruikte concepten definiëren. Bovendien kunnen deze sleutelwoorden een eerste suggestie zijn voor de verschillende lexicale representaties. Een definiëring van concepten, relaties en topics zou de resultaten van de topischerkenning en de kwaliteit van de samenvattingen kunnen verbeteren. De achterliggende reden is het feit, dat het systeem deze woorden en woordgroepen, zonder enige voorkennis, reeds als relevant beschouwt. Zo vertrekt men reeds van een sterke basis en is het mogelijk dat dit tot betere resultaten leidt.

Verder is het aan te raden om na te gaan, of de huidige 'text segmentation'-technieken, waarbij een verandering van topic in de tekst opgespoord wordt, reeds geïntegreerd kunnen worden. Vervolgens dient te worden onderzocht, of de opsplitsing van de tekst op basis van text segmentation ook daadwerkelijk verbeterde resultaten als uitkomst geeft. In het huidige praktijkgeval zijn de twee controleteksten grotendeels ingedeeld op basis van de door de auteur aangebrachte indeling. Dit kan in het kader van dit onderzoek misschien niet de meest optimale indeling zijn, aangezien één paragraaf niet steeds beperkt is tot één topic. De paragrafen gebaseerd op 'text segmentation' zouden meer coherent moeten zijn en vandaar dat men betere resultaten zou kunnen verwachten bij de herkenning van topics.

In de tweede stap van methode 1 is er een taxonomie voor het ICT-domein ontworpen. Bij het gebruik ervan om de concepten aan de ontologie van i.Know toe te voegen, is reeds gebleken dat deze taxonomie op bepaalde gebieden niet volledig of niet gedetailleerd genoeg is. Maar het strekt misschien tot aanbeveling om deze taxonomie in de eerste plaats te vervolledigen. Vervolgens dient men passende lexicale representaties voor de gebruikte begrippen op te sporen en dan de taxonomie in zijn geheel aan de ontologie van i.Know toe te voegen. Verder onderzoek moet dan uitwijzen of deze methode dan toch geen meerwaarde leveren kan.

In het algemeen dienen de nieuwe trends in e-learning steeds te worden opgevolgd. Ook de verdere evoluties in text mining en gerelateerde domeinen worden best in het oog gehouden. Bovendien is het aangewezen om dit eerste verkennend onderzoek nog verder uit te breiden en zo mogelijk bijkomende technologieën op te sporen en vervolgens na te gaan op hun bruikbaarheid in deze gevalstudie.

### **6.3 Persoonlijke conclusies**

Deze eindverhandeling zou ik graag afsluiten met een korte persoonlijke noot. Het literatuuronderzoek gaf de indruk dat een text mining-toepassing zeer eenvoudig verlopen kan. Het is een vaak gebruikt cliché, maar de praktijk toonde het tegendeel aan. De tegenvallende resultaten van de eerste methode uit de gevalstudie waren dan ook teleurstellend. Stilaan werd duidelijk dat een goede precisie enkel haalbaar is binnen een beperkt domein, waarbij vooraf veel elementen worden gedefinieerd. Hoe uitgebreider het domein waarbinnen de toepassing moet functioneren, hoe lager de precisie.

Een positief punt is wel, dat het onderzoek en ontwikkeling in deze sector toch relatief vlug evolueren. Dit heb ik namelijk zelf kunnen vaststellen in de praktijk met de applicatie voor het zoeken naar mogelijke CRC's en de 'Smart Indexing' die tijdens het uitvoeren van de gevalstudie bij i.Know in gebruik genomen werden.

Tot slot ben ik zeer benieuwd naar het wanneer en hoe deze gevalstudie ook in de praktijk zal toegepast kunnen worden. Theoretisch lijkt het haalbaar, mits een aantal serieuze inspanningen. De praktijk is, zoals juist reeds gesteld, vaak heel anders. Ik hoop toch dat, de bevindingen uit dit beperkte onderzoek ooit in één of andere vorm in de realiteit omgezet worden.

## LIJST VAN DE GERAADPLEEGDE WERKEN

- Alexander, B. (2006, March/April). Web 2.0: A New Wave of Innovation for Teaching and Learning. *EDUCAUSE Review* , 41 (2), pp. 32-44.
- Allan, J. (2002). Detection As Multi-Topic Tracking. *Information Retrieval* , 5 (2-3), 139-157.
- Bloehdorn, S., Cimiano, P., Hotho, A., & Staab, S. (2005). An Ontology-based Framework for Text Mining. *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology* , 20, 87-112.
- Boehle, S. (2006, januari 17). The State of the E-Learning Market. *Training Magazine* .
- Brands, M. (2003). *Bedrijfspresentatie: i.Know, Intelligent Knowledge Management System N.V.* Diepenbeek: i.Know NV.
- Churchill, D. (2005, december). Learning objects: an interactive representation and a mediating tool in a learning activity. *Educational Media International* , 42 (4), pp. 333-349.
- Clarke, A. (2007, maart). The future of e-learning. *Adults Learning* , pp. 14-15.
- Cramer, S. R. (2007, Januari/Februari). Update Your Classroom with Learning Objects and Twenty-First Century Skills. *The Clearing House* , 80 (3), pp. 126-132.
- Cross, J., & Hamilton, I. (2002). *The DNA of eLearning*. Opgeroepen op juli 13, 2006, van Internet Time Group: <http://www.internettime.com/learning/articles/DNA.pdf>
- Cunningham, H. (2005). Information Extraction, Automatic. In K. Brown (Red.), *Encyclopedia of Language and Linguistics* (2nd Edition ed., pp. 665-677). Elsevier.
- de Swart, H. (z.d.). *Dynamische Semantiek*. Opgeroepen op augustus 18, 2007, van Inleiding Taalkunde 2007 (Universiteit Utrecht): <http://www.phil.uu.nl/~mariekes/ITK07/07ithoordyn.ppt>
- Dixon, M. (1997, oktober 4). *An Overview of Document Mining Technology*. Opgeroepen op mei 16, 2006, van -: <http://www.geocities.com/ResearchTriangle/Thinktank/1997/mark/writings/main.html>
- Droste, J., & Rikhof-van Eijck, M. (2002). *E-learning en onderwijsvernieuwing*. 's-Hertogenbosch: CINOP (Centrum voor innovatie van opleidingen).
- El Wakil, M. M. (z.d.). Introducing Text Mining.
- Fallon, C., & Brown, S. (2003). *E-learning standards: a guide to purchasing, developing, and deploying standards-conformant e-learning*. Boca Raton: St. Lucie Press.
- Gahremani, Y. (2006, januari 1). The Joy of Text. *CFO Magazine* , pp. 79-81.
- Hahn, U., & Mani, I. (2000). The Challenges of Automatic Summarization. *Computer* , 33 (11), 29-36.

- Harting, K., & Erthal, M. J. (2005). History of Distance Learning. *Information Technology, Learning, and Performance Journal* , 23 (1), pp. 35-44.
- Huisman, E. (2001). *Corporate e-learning, gids voor managers van leren*. Amsterdam: Uitgeverij Boom.
- i.Know NV. (2005). *i.Know's functionaliteiten*. Diepenbeek: i.Know NV.
- i.Know NV. (2005). *Technical White Paper*. Diepenbeek: i.Know NV.
- Kao, A., & Poteet, S. (2004). Report on KDD Conference 2004 Panel Discussion: Can Natural Language Processing Help Text Mining? (S. Sarawagi, Red.) *ACM Special Interest Group on Knowledge Discovery and Data Mining* , 6 (2), 132-133.
- Kao, A., & Poteet, S. (2005). Text Mining and Natural Language Processing – Introduction for the Special Issue. (S. Sarawagi, Red.) *ACM Special Interest Group on Knowledge Discovery and Data Mining* , 7 (1), 1-2.
- Karanikas, H., & Theodoulidis, B. (2003, september 30). *Knowledge Discovery in Text and Text Mining Software*. Opgeroepen op december 14, 2006, van Parmenides: [http://www.crim.co.umist.ac.uk/parmenides/internal/docs/Karanikas\\_NLDB2002%20.pdf](http://www.crim.co.umist.ac.uk/parmenides/internal/docs/Karanikas_NLDB2002%20.pdf)
- Kruse, K. (2002, september). *The State of e-Learning: Looking at History with the Technology Hype Cycle*. Opgeroepen op juli 13, 2006, van E-Learning Guru: [http://www.e-learningguru.com/articles/hype1\\_1.htm](http://www.e-learningguru.com/articles/hype1_1.htm)
- Leezenberg, P. B., Tonningen, B. R., & Schoonhoven, E. (2005, februari 1). *Using Knowledge Management and E-learning for Strategic Change*. Opgeroepen op december 13, 2005, van ManagementSite: <http://www.managementsite.nl/content/articles/296/296.asp>
- Lehman, R. (2007). Learning Objects Repositories. *New Directions for Adult and Continuing Education* (113), pp. 57-66.
- Masie. (2002, maart 8). *Making sense of learning specifications and standards: A decision maker's guide to their adoption*. Opgeroepen op mei 28, 2007, van The MASIE Center: [http://www.masie.com/standards/S3\\_Guide.pdf](http://www.masie.com/standards/S3_Guide.pdf)
- Remels, K. (2002). *Blended learning: een nieuwe stap in de evolutie van e-learning*. Diepenbeek: Limburgs Universitair Centrum.
- Rosenberg, M. J. (2001). *E-Learning: Strategies for Delivering Knowledge in the Digital Age*. New York: McGraw-Hill.
- Rubens, W. (2003, september). *De prille geschiedenis van e-learning: Omzien in verwondering*. Opgeroepen op juli 5, 2006, van Wilfred Rubens: technology enhanced learning: <http://www.te-learning.nl/omzieninverwondering.pdf>

- Rubens, W., & Admiraal, W. (2003, mei). *Samenwerkend leren met behulp van ICT binnen het Europese onderwijs: ervaringen met Synergeia in vier landen*. Opgeroepen op juni 5, 2007, van Wilfred Rubens: technology enhanced learning: <http://www.te-learning.nl/ord2003cscl.pdf>
- Schreurs, J., & Moreau, R. (2006). Converting digital learning content into Learning Objects (LO). *Proceedings of the International Conference Knowledge Acquisition and Management KAM2006*, (p. 227). Wroclaw-Poland.
- Schreurs, J., & Moreau, R. (2006). Converting Learning Content to Learning Objects (LO) and Atomic LO's. *Proceedings of the Virtual Conference: IADIS Virtual Multi Conference on Computer Science and Information Systems*.
- Schreurs, J., & Moreau, R. (2006). Learning objects (LO) aligning different learning styles. *Proceedings of the International Conference on e-learning (ICEL2006)*, (pp. 415-422). University of Quebec Montréal.
- Schreurs, J., & Moreau, R. (2005). Learning Styles in an E-Mindmap. *Proceedings of EDEN 2005 conference*. Helsinki.
- Schreurs, J., & Moreau, R. (2006). Modelling of E-Learning Course Scenario's.
- Shepherd, C. (2000). *Objects of interest*. Opgeroepen op mei 28, 2007, van Fastrak Consulting Ltd.: <http://www.fastrak-consulting.co.uk/tactix/features/objects/objects.htm>
- Spasic, I., Ananiadou, S., McNaught, J., & Kumar, A. (2005). Text mining and ontologies in biomedicine: Making sense of raw text. *Briefings in Bioinformatics*, 6 (3), 239-251.
- The New Media Consortium and the EDUCAUSE Learning Initiative. (2006). *The Horizon Report 2006*. Opgeroepen op juli 13, 2006, van The New Media Consortium: [http://www.nmc.org/pdf/2006\\_Horizon\\_Report.pdf](http://www.nmc.org/pdf/2006_Horizon_Report.pdf)
- Vandewijer, H. (2004). *Learning Service Providing, een ASP-oplossing voor e-learning. De ontwikkeling van een LSP-portaal*. Diepenbeek: Limburgs Universitair Centrum.
- Varlamis, I., & Apostolakis, I. (2006). The Present and Future of Standards for E-Learning Technologies. (A. Koohang, Red.) *Interdisciplinary Journal of Knowledge and Learning Objects*, 2, 59-76.
- Wiley, D. A. (2000). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley, *The Instructional Use of Learning Objects*. Bloomington: Association for Instructional Technology en Association for Educational Communications and Technology.
- Williams, J. B., & Goldberg, M. (2005). The evolution of e-learning. *ASCILITE 2005 Conference Proceedings* (pp. 725-728). Brisbane: Sociaty for Computers in Learning in Tertiary Education.

## BIJLAGEN

### Bijlage I      Taxonomie

| Categorieën                     |                  | Sleutelwoorden                   | Synoniemen |                           |
|---------------------------------|------------------|----------------------------------|------------|---------------------------|
| <b>C Information Technology</b> |                  |                                  | C_s1       | Computer science          |
|                                 |                  |                                  | C_s2       | IS Technology             |
|                                 |                  |                                  | C_s3       | Technology                |
|                                 |                  |                                  | C_s4       | Computer technology       |
|                                 |                  |                                  | C_s5       | Technological environment |
| CA                              | Computer Systems |                                  | CA_s1      | Computers                 |
|                                 |                  |                                  | CA_s2      | Hardware                  |
|                                 | CA01             | Data communications              |            |                           |
|                                 | CA02             | Input/output devices             |            |                           |
|                                 |                  | CA02_1 Terminal                  |            |                           |
|                                 |                  | CA02_2 Optical disk              |            |                           |
|                                 |                  | CA02_3 CD-ROM                    |            |                           |
|                                 |                  | CA02_4 ATM                       |            |                           |
|                                 |                  | CA02_5 Pen                       |            |                           |
|                                 | CA03             | Memory                           |            |                           |
|                                 | CA04             | Processors                       |            |                           |
|                                 | CA05             | Storage devices                  |            |                           |
|                                 | CA06             | Work stations                    |            |                           |
|                                 | CA07             | Computer type                    |            |                           |
|                                 |                  | CA07_1 Supercomputer             |            |                           |
|                                 |                  | CA07_2 Minicomputer              |            |                           |
|                                 |                  | CA07_3 Microcomputer             | CA07_1_s1  | Personal computer         |
|                                 |                  |                                  | CA07_1_s2  | Desktop computer          |
|                                 | CA08             | Computer architecture            |            |                           |
|                                 | CA09             | Distributed systems              |            |                           |
|                                 | CA10             | Networks                         |            |                           |
|                                 |                  | CA10_1 LAN                       |            |                           |
|                                 |                  | CA10_2 Wide area networks        |            |                           |
|                                 |                  | CA10_3 ISDN                      |            |                           |
|                                 | CA11             | Telecommunications               |            |                           |
|                                 | CA12             | Front-end computers              |            |                           |
|                                 | CA13             | Back-end computers               |            |                           |
|                                 | CA14             | Multimedia                       |            |                           |
|                                 | CA15             | Client-server                    |            |                           |
| CB                              | Software         |                                  |            |                           |
|                                 | CB02             | Machine Language                 |            |                           |
|                                 | CB04             | Operating Systems                |            |                           |
|                                 |                  | CB04_1 Virtual memory            |            |                           |
|                                 |                  | CB04_2 Unix                      |            |                           |
|                                 | CB05             | Compilers                        |            |                           |
|                                 | CB06             | DBMS                             |            |                           |
|                                 |                  | CB06_1 Relational model          |            |                           |
|                                 |                  | CB06_2 Network model             |            |                           |
|                                 |                  | CB06_3 Hierarchical model        |            |                           |
|                                 |                  | CB06_4 Semantic data model       |            |                           |
|                                 |                  | CB06_5 Entity-relationship model |            |                           |
|                                 |                  | CB06_6 DDL                       |            |                           |



|          |                      |                                    |       |                                 |
|----------|----------------------|------------------------------------|-------|---------------------------------|
|          | CB06_7               | Query languages                    |       |                                 |
|          | CB06_8               | SQL                                |       |                                 |
|          | CB06_9               | Object-oriented DBMS               |       |                                 |
| CB07     |                      | Programming languages              |       |                                 |
|          | CB07_1               | COBOL                              |       |                                 |
|          | CB07_2               | FORTRAN                            |       |                                 |
|          | CB07_3               | COBOL                              |       |                                 |
|          | CB07_4               | BASIC                              |       |                                 |
|          | CB07_5               | Pascal                             |       |                                 |
| CB08     |                      | 4th Generation languages           |       |                                 |
| CB09     |                      | Software Packages                  |       |                                 |
|          | CB09_1               | Modeling languages                 |       |                                 |
|          | CB09_2               | Program generators                 |       |                                 |
|          | CB09_3               | Report generators                  |       |                                 |
|          | CB09_4               | Statistical packages               |       |                                 |
|          | CB09_5               | Electronic spreadsheets            |       |                                 |
|          | CB09_6               | Word processing software           |       |                                 |
|          | CB09_7               | Hypertext                          |       |                                 |
|          | CB09_8               | Knowledge-based software           |       |                                 |
|          | CB09_9               | Communications software            |       |                                 |
| CB10     |                      | Natural languages                  |       |                                 |
| <b>E</b> | <b>IS Management</b> |                                    | E_s1  | MIS Management                  |
|          |                      |                                    | E_s2  | DP Management                   |
|          |                      |                                    | E_s3  | Information Management          |
|          |                      |                                    | E_s4  | Information Resource Management |
|          |                      |                                    | EA_s1 | Database administration         |
| EA       |                      | Database management                |       |                                 |
| EC       |                      | Hardware resource management       |       |                                 |
|          | EC01                 | Hardware selection                 |       |                                 |
|          | EC03                 | Hardware maintenance               |       |                                 |
|          | EC04                 | Hardware contracts                 |       |                                 |
| ED       |                      | Software resource management       |       |                                 |
|          | ED01                 | Software selection                 |       |                                 |
|          | ED03                 | Software contracts                 |       |                                 |
| EE       |                      | IS Project management              |       |                                 |
| EF       |                      | IS Planning                        |       |                                 |
|          | EF01                 | IS Planning methods                |       |                                 |
|          |                      | EF01_1 CSF                         |       |                                 |
|          |                      | EF01_2 BSP                         |       |                                 |
|          |                      | EF01_3 Enterprise modeling         |       |                                 |
|          |                      | EF01_4 Information engineering     |       |                                 |
|          | EF04                 | IS Strategic planning              |       |                                 |
|          | EF05                 | IS Operational planning            |       |                                 |
|          |                      | EF05_1 Capacity planning           |       |                                 |
|          | EF09                 | IS Architecture                    |       |                                 |
| EG       |                      | Organizing IS                      |       |                                 |
|          | EG01                 | IS centralization/decentralization |       |                                 |
| EH       |                      | IS Staf                            |       |                                 |
|          | EH01                 | IS Job categories                  |       |                                 |
|          |                      | EH01_1 CIO                         |       |                                 |
|          |                      | EH01_2 Information analyst         |       |                                 |
|          |                      | EH01_3 Programmer                  |       |                                 |
|          |                      | EH01_4 Computer operator           |       |                                 |
|          |                      | EH01_5 EDP auditor                 |       |                                 |

|                                        |                                              |                              |                             |         |                             |
|----------------------------------------|----------------------------------------------|------------------------------|-----------------------------|---------|-----------------------------|
|                                        |                                              | EH01_6                       | Knowledge engineer          |         |                             |
| EI                                     | IS Evaluation                                |                              |                             |         |                             |
| EJ                                     | IS Control                                   |                              |                             |         |                             |
|                                        | EJ01                                         | IS Control methods and tools |                             |         |                             |
|                                        |                                              | EJ01_1                       | EDP audit                   |         |                             |
| EK                                     | IS Security                                  |                              |                             |         |                             |
|                                        | EK01                                         | Data security                |                             |         |                             |
|                                        | EK02                                         | Data encryption              |                             |         |                             |
|                                        |                                              | EK02_1                       | Public key encryption       |         |                             |
|                                        | EK03                                         | Access control               |                             |         |                             |
|                                        | EK07                                         | Passwords                    |                             |         |                             |
|                                        | EK08                                         | Disaster plans               |                             |         |                             |
|                                        | EK09                                         | Computer crime               |                             |         |                             |
|                                        | EK10                                         | Computer viruses             |                             |         |                             |
| EL                                     | IS Management                                |                              |                             |         |                             |
|                                        | EL06                                         | IS Intergration              |                             |         |                             |
|                                        | EL07                                         | Outsourcing of IS            |                             |         |                             |
|                                        | EL08                                         | IS Risk management           |                             |         |                             |
|                                        | EL10                                         | IS Downsizing                |                             |         |                             |
| <b>F IS Development and Operations</b> |                                              |                              |                             | F_s1    | Applications development    |
|                                        |                                              |                              |                             | F_s2    | Software development        |
|                                        |                                              |                              |                             | F_s3    | Software engineering        |
|                                        |                                              |                              |                             | F_s4    | Systems analysis and design |
| FA                                     | IS Development strategies                    |                              |                             |         |                             |
|                                        | FA02                                         | Prototyping                  |                             |         |                             |
|                                        | FA11                                         | Object-oriented approach     |                             |         |                             |
|                                        | FA12                                         | Reverse engineering          |                             |         |                             |
| FB                                     | Systems development life cycle               |                              |                             |         |                             |
|                                        | FB03                                         | Systems analysis             |                             |         |                             |
|                                        | FB04                                         | IS Design                    |                             | FB04_s1 | Systemes design             |
|                                        |                                              | FB04_1                       | Database design             |         |                             |
|                                        |                                              | FB04_2                       | Normalization               |         |                             |
|                                        | FB05                                         | Programming                  |                             |         |                             |
|                                        |                                              | FB05_1                       | Modular programming         |         |                             |
|                                        |                                              | FB05_2                       | Interactive programming     |         |                             |
|                                        |                                              | FB05_3                       | Structured programming      |         |                             |
|                                        |                                              | FB05_4                       | Top-down programming        |         |                             |
|                                        |                                              | FB05_5                       | Object-oriented programming |         |                             |
|                                        | FB06                                         | Testing                      |                             |         |                             |
|                                        | FB07                                         | Conversion                   |                             |         |                             |
|                                        | FB08                                         | Program maintenance          |                             |         |                             |
| FC                                     | Software development methodologies and tools |                              |                             |         |                             |
|                                        | FC04                                         | CASE                         |                             |         |                             |
|                                        | FC05                                         | Data dictionary              |                             |         |                             |
|                                        | FC06                                         | DFD                          |                             |         |                             |
|                                        | FC08                                         | Decision tables              |                             |         |                             |
|                                        | FC09                                         | Decision trees               |                             |         |                             |
|                                        | FC10                                         | Entity diagrams              |                             |         |                             |
|                                        | FC12                                         | Flowcharts                   |                             |         |                             |
|                                        | FC13                                         | Functional decomposition     |                             |         |                             |
|                                        | FC19                                         | Structured modeling          |                             |         |                             |
|                                        | FC21                                         | Pseudo-code                  |                             |         |                             |
|                                        | FC28                                         | Object-oriented analysis     |                             |         |                             |

|    |                   |                               |
|----|-------------------|-------------------------------|
|    | FC29              | Object-oriented design        |
| FD | IS Implementation |                               |
| FE | IS Operations     |                               |
|    | FE01              | Data processing               |
|    | FE02              | Computer systems installation |

#### **G IS Usage**

|    |                          |                    |
|----|--------------------------|--------------------|
| GA | Organizational use of IS |                    |
|    | GA01                     | Strategic IS       |
|    | GA02                     | Office automation  |
|    | GA02_1                   | Word processing    |
|    | GA03                     | Personal computing |
|    | GA05                     | Telecommuting      |
| GB | Users                    |                    |
|    | GB05                     | User requirements  |
| GC | Type of IS Support       |                    |
|    | GC03                     | Personal IS        |
| GD | IS Access                |                    |
| GE | Type of processing       |                    |
|    | GE01                     | Online IS          |
|    | GE02                     | Batch IS           |
|    | GE03                     | Real-time IS       |
|    | GE04                     | Interactive IS     |

#### **H Information systems**

|    |                              |                                           |                                 |
|----|------------------------------|-------------------------------------------|---------------------------------|
| HA | Types of information systems |                                           |                                 |
|    | HA01                         | Transaction processing systems            |                                 |
|    | HA02                         | MIS                                       |                                 |
|    | HA03                         | DSS                                       |                                 |
|    | HA03_1                       | Group DSS                                 |                                 |
|    | HA04                         | Expert systems                            | HA04_s1 Knowledge-based systems |
|    | HA05                         | Executive support systems                 |                                 |
|    | HA06                         | Office systems                            |                                 |
|    | HA06_1                       | EFT                                       |                                 |
|    | HA06_2                       | Electronic markets                        |                                 |
|    | HA08                         | Communication systems                     |                                 |
|    | HA08_1                       | Electronic mail                           |                                 |
|    | HA08_2                       | Computer conferencing                     |                                 |
|    | HA08_3                       | Facsimile transmission                    |                                 |
|    | HA08_4                       | Videotext                                 |                                 |
|    | HA08_5                       | Teletext                                  |                                 |
|    | HA08_6                       | Voice/data sharing                        |                                 |
|    | HA08_7                       | Message systems                           |                                 |
|    | HA08_8                       | Audio conference                          |                                 |
|    | HA08_9                       | Voice messaging                           |                                 |
|    | HA09                         | Information storage and retrieval systems |                                 |
|    | HA09_1                       | Information search and retrieval          |                                 |
|    | HA10                         | Document management systems               |                                 |
|    | HA11                         | Electronic meeting systems                |                                 |
|    | HA12                         | Collaborative work systems                |                                 |
| HB | IS Application               |                                           | HB_s1 Applications software     |
|    |                              |                                           | HB_s2 Business applications     |
|    |                              |                                           | HB_s3 Information systems       |
| HC | Components of IS             |                                           |                                 |
|    | HC01                         | User interface                            |                                 |

|    |                    |                            |
|----|--------------------|----------------------------|
|    | HC02               | Database                   |
|    | HC03               | Programs                   |
|    | HC04               | Input                      |
|    | HC05               | Output                     |
| HD | IS Characteristics |                            |
|    | HD01               | Interface characteristics  |
|    | HD01_1             | Computer graphics          |
|    | HD01_2             | Graphical user interface   |
|    | HD01_3             | Natural language interface |
|    | HD01_4             | Interactive interface      |

## **Biilage II      Tekst 1: 'Outsourcing Collaborative Partnerships'**

### 1 Chapter Fifteen:                      Outsourcing Collaborative Partnerships

More than 400 people from Merrill Lynch, Thomson Financial (a large market data vendor), and a number of other vendors have been working feverishly on Merrill Lynch's biggest outsourcing initiative ever. This highly complex \$1 billion makeover of its wealth management system is designed to improve the efficiency of Merrill's financial advisers. With the new system, Merrill Lynch's financial advisers can obtain more of the assets of their high-net-worth customers.

The new system also represents a major shift in the way Merrill approaches IT initiatives. In the 1990s, Merrill developed its previous system, Trusted Global Advisor (TGA), as it did any other major system - it developed it in-house. The thought of outsourcing a critical business system to a vendor was viewed as highly unfavorable by most financial services organizations. Last year, Merrill Lynch signed a contract that outsourced much of the responsibility for its new platform to Thomson Financial.

In the high-speed global business environment, an organization needs to maximize its profits, enlarge its market share, and restrain its ever-increasing costs. Businesses need to make every effort to rethink and readopt new processes, especially the prospective resources regarding insourcing, outsourcing, and offshore outsourcing.

### 2 INSOURCING

Insourcing is a common approach using the professional expertise within an organization to develop and maintain the organization's information technology systems. Insourcing has been instrumental in creating a viable supply of IT professionals and in fact in creating a better quality workforce combining both technical and business skills.

Michael Palmer, the COO (and former CIO) of Allied Office Products, decided to bring in-house nearly all of the IT functions that had been outsourced. Palmer calculated that he could cut the \$24,000-a-month hardware monitoring and maintenance costs by two-thirds and save several hundred ff on development by insourcing. Since Palmer turned to insourcing 85 percent of the IT work, he is now saving nearly \$500,000 a year.

### 3 OUTSOURCING

Outsourcing is an arrangement by which one organization provides a service or services for another organization that chooses not to perform them in-house. In some cases, the entire information technology department is outsourced, including planning and business analysis as well as the installation, management, and servicing of the network and workstations. Outsourcing can range from a large contract under which an organization like IBM manages IT services for a company such as Xerox, to the practice of hiring contractors and temporary office workers on an individual basis. Refer to Figure 4.16 for a comparison of the functions companies have outsourced.

4 Ever since Eastman Kodak announced that it was outsourcing its information' systems function in 1988 to IBM, DEC, and Businessland, large organizations have found it acceptable to transfer their IT assets, leases, and staff to outsourcers. In view of the changes in sourcing, the key question now is not "should we outsource IT?" but rather "where and how can we take advantage of the rapidly developing market of IT services providers?" Organizations should consider outsourcing in order to achieve the following benefits:

Financial savings

Increase technical abilities

Market agility

#### Financial Savings

Cost competitiveness has driven the Big Three - Ford, DaimlerChrysler, and General Motors - to collaborate in a unique way by building an IT service. Automotive Network eXchange (ANX) is a network that joins suppliers and buyers across the automotive supply chain and allows the sharing of CAD / CAM drawings, as well as exchanging trade e-mail messages and shipping communications. Ironing out such inefficiencies across the supply chain alone could decrease the cost of producing a car as much as \$1,200.

#### Increase Technical Abilities

Rapid technological change has become an issue for many businesses throughout all industries, resulting in more expense to upgrade systems, more time to install, and increased complexity to master. For a business whose IT department is a non-core function, maintaining a "best-of-breed" status under these conditions is next to impossible, especially for small and medium-sized enterprises where cost is a critical factor.

Nike's footwear business generates \$6 billion in revenues per year. It requires an achievement in global communication to keep its global Web of Nike designers, product developers, marketing teams, sales staff, distributors, and dealers in line with product developments and marketing plans. Nike uses a system called GPIN (Global Product Information Network), an innovative IT service that allows Nike's employees to collaborate throughout their processes and keeps every partner up-to-date with developments.

#### Market Agility

Market agility includes the ability to expand core businesses more rapidly depending on the outsourcer's capabilities to provide efficient transitions to new systems, better information management for decision making, and expansion to new geographical markets.

Kodak is over 100 years old. So is chemical-based photography. Therefore, Kodak moved its focus from being in the film business to being in the picture-processing business, and from this developed Kodak's exciting e-service, Photonet, a password-protected application that allows users to post photographs on the Internet, order copies, send picture postcards, and customize postcard greetings.

IT outsourcing is a fast-growing industry since it provides access to state-of-the-art technologies with expert guidance, thus curtailing the need to open up expensive in-house departments. According to a report by IDC, global spending on IT services will rise to \$700 billion by 2005, an increase from \$440 billion in 2002. Many factors have converged to prompt firms to outsource. Figure 4.17 displays the reason why many organizations outsource key departments.

5 Focusing a company's resources on core business functions allows outsourcing noncore functions. Outsourcing can give the right combination of people, processes, and technology to operate efficiently and effectively in the global marketplace without burdening time and budget. Some of the influential drivers affecting the growth of the outsourcing market include:

Core competencies- Many companies have recently begun to consider outsourcing as a means to fuel revenue growth rather than just a cost-cutting measure. Outsourcing enables an organization to maintain an up-to-date technology infrastructure while freeing it to focus on revenue growth goals by reinvesting cash and human capital in areas offering the greatest return on investment.

Rapid growth- A company's sustainability depends on both speed to market and ability to react quickly to changes in market conditions. By taking advantage of outsourcing, an organization is able to acquire best practices process expertise. This facilitates the design, building, training, and deployment of business processes or functions.

Industry changes- High levels of reorganization across industries have increased demand for outsourcing to better focus on core competencies. The significant increase in merger and acquisition activity created a sudden need to integrate multiple core and noncore business functions into one business, while the deregulation of the utilities and telecom industries created a need to ensure compliance with government rules and regulations. Companies in either situation turned to outsourcing so they could better focus on industry changes at hand.

The Internet- The pervasive nature of the Internet as an effective sales channel has allowed clients to become more comfortable with the outsourcing process.

#### 6 DEVELOPING STRATEGIC OUTSOURCING PARTNERSHIPS

Business process outsourcing (BPO) is the contracting of a specific business task, such as payroll, to a third-party service provider. Business process outsourcing is increasingly becoming the strategic choice of companies looking to achieve cost reductions while improving their service quality, increasing shareholder value, and focusing on their core business capabilities. Many organizations are looking beyond traditional IT outsourcing to business process outsourcing as the next logical step.

Implementing BPO can assist with a cost-saving measure for tasks that an organization requires but does not depend upon to maintain its position in the marketplace. BPO is divided into two categories: (1) back-office outsourcing that includes internal business functions such as billing or purchasing, and (2) front-office outsourcing which includes customer-related services such as marketing or technical support.

Business process outsourcing is not a new field. Paychex, based in Rochester, New York, for example, has been outsourcing payroll processing for small businesses since 1971. However, the market is heating up these days, as companies' need for strategic cost cutting, desire to improve business methods, and comfort with outsourcing arrangements grow.

The Gartner Group estimates that the worldwide market for business process outsourcing will grow to \$178 billion in 2005. Management consulting and technology services company Accenture is a BPO player. Accenture's first BPO deal came in the early 1990s, when it outsourced finance and accounting functions for British Petroleum. Accenture now handles a variety of outsourcing tasks, such as airline ticket processing and call center staffing for AT&T.

#### 7 SOURCING'S NEW SURGE - OFFSHORING

Offshore outsourcing is using organizations from developing countries to write code and develop systems. Numerous countries have substantially well-trained IT professionals and clerical staff who have lower salary expectations compared to their U.S. counterparts. Offshore outsourcing has become a small but rapidly growing sector in the overall outsourcing market. Nearly half of all businesses use offshore providers, and two-thirds plan to send work overseas in the near future, according to Forrester Research. India receives most of the outsourcing functions from all over the world. However, as more American companies seek to source globally, more countries are emerging to benefit from that demand - from Canada to Malaysia - each with its own particular strengths and weaknesses. Organizations have much to gain from offshore outsourcing.

**Bijlage III**    **Schema tekst 1: 'Outsourcing Collaborative Partnerships'**

| Topic       | Concept 1   | Concept 2                     | Concept 3                      |   | Relatie                   | Synoniemen                                                                  |
|-------------|-------------|-------------------------------|--------------------------------|---|---------------------------|-----------------------------------------------------------------------------|
| insourcing  | insourcing  |                               |                                |   |                           |                                                                             |
|             |             | internal ICT work force       |                                | → | requires                  | in-house ICT services<br>internal ICT experts<br>internal ICT professionals |
|             |             | supply of ICT professionals   |                                | ← | can be done by the way of |                                                                             |
|             |             | IT work                       | development of IS              | → | includes                  | development<br>ICT application<br>development                               |
|             |             |                               | maintenance of IS              | → | includes                  | maintaining                                                                 |
| outsourcing | outsourcing |                               |                                |   |                           |                                                                             |
|             |             | external ICT services         |                                | → | requires                  | not in-house                                                                |
|             |             | supply of ICT professionals   |                                | ← | can be done by the way of |                                                                             |
|             |             | IT work                       | entire IT department planning  | → | includes                  | ICT planning<br>ICT strategy<br>system analysis                             |
|             |             |                               | business analysis              | → | includes                  | networks                                                                    |
|             |             |                               | installation of infrastructure | → | includes                  | workstations                                                                |
|             |             | up-to-date ICT infrastructure |                                | → | includes                  |                                                                             |
|             |             | non-core functions            |                                | → | includes                  |                                                                             |
|             |             | contract                      |                                | → | requires                  |                                                                             |
|             |             | ICT services                  |                                | → | includes                  |                                                                             |
|             |             | ICT service provider          |                                | → | organized by              |                                                                             |



| outsourcing benefits  | benefits                      |   |                      | advantage     |
|-----------------------|-------------------------------|---|----------------------|---------------|
|                       | outsourcing                   | ← | can deliver          |               |
|                       | financial savings             | → | includes             |               |
|                       | cost competitiveness          | → | can be achieved by   | cost decrease |
|                       | increase technical abilities  | → | includes             |               |
|                       | systems upgrade               | → | do more efficient    |               |
|                       | installation                  | → | do more efficient    |               |
|                       | systems                       | → | do more efficient    |               |
|                       | master complexity             | → | solves the           |               |
|                       | rapid technological change    | → | problem limited with |               |
|                       | market agility                | → | includes             |               |
|                       | expansion of core business    | → | speeds up            |               |
|                       | new systems                   | → | efficient            |               |
|                       | new geographical markets      | → | transition to        |               |
|                       | expansion to                  | → |                      |               |
|                       | reasons of outsourcing        |   |                      |               |
|                       | state of the art technology   | → | includes             |               |
|                       | core business focus           | → | includes             |               |
|                       | outside sources of expertise  | → | includes             |               |
|                       | reduce costs                  | → | includes             |               |
|                       | minimal investment in ICT     | → | includes             |               |
|                       | revenue growth                | → | focus on             |               |
| IT outsourcing market | drivers of growth             |   |                      |               |
|                       | outsourcing market            | ← | is influenced by     |               |
|                       | outsourcing                   | ← | is delivered by      |               |
|                       | core competencies             | → | includes             |               |
|                       | rapid growth                  | → | includes             |               |
|                       | speed to market               | → | results in           |               |
|                       | reaction to changes in market | → | improves             |               |
|                       | industry changes focus        | → | includes             |               |

|                              |                              |                                                         |                                                |                                |
|------------------------------|------------------------------|---------------------------------------------------------|------------------------------------------------|--------------------------------|
|                              |                              | internet as sales channel                               | → includes                                     |                                |
| Business process outsourcing | business process outsourcing |                                                         |                                                | BPO                            |
|                              |                              | outsourcing                                             | ← is evolving to/includes                      |                                |
|                              |                              | business process                                        | → is applicable                                |                                |
|                              |                              | strategic decision                                      | → is a                                         |                                |
|                              |                              | cost reduction<br>service quality<br>stakeholders value | → to achieve<br>→ to increase<br>→ to increase |                                |
|                              |                              | back-office outsourcing                                 | → can be                                       |                                |
|                              |                              | internal business functions                             | → includes                                     | billing<br>purchasing          |
|                              |                              | front-office outsourcing                                | → can be                                       |                                |
|                              |                              | customer-related services                               | → includes                                     | marketing<br>technical support |
| offshore outsourcing         | offshore outsourcing         |                                                         |                                                |                                |
|                              |                              | ICT services from developing countries                  | → requires                                     |                                |
|                              |                              | well trained IT professionals<br>lower salary           | → is characterized by<br>→ is characterized by |                                |
|                              |                              | supply of ICT professionals                             | ← can be done by the way of                    |                                |
|                              |                              | program code                                            | → orders for                                   |                                |
|                              |                              | system development                                      | → orders for                                   |                                |
|                              |                              | IT work                                                 | → orders for                                   |                                |
|                              |                              | global sourcing                                         | → includes                                     |                                |

## **Biilage IV      Tekst 2: 'Outsourcing Information Systems Functions'**

### **1 Outsourcing Information Systems Functions**

In the IT world, outsourcing means turning over a firm's computer operations, network operations, or other IT function to a provider for a specified time; generally, at least a few years. In 1989, outsourcing became a legitimate management strategy by CIOs. Until that time, the only companies that outsourced their IS departments were those that were poorly run. However, in 1989, Kodak outsourced its well-run IS operations to become a more competitive company. That surprising move caused top executives around the world to consider the use of outsourcers. Today, CIOs are expected to investigate outsourcing sufficiently to satisfy executive management that their IS operations are as efficient and effective in-house as they would be if they were outsourced; otherwise they should outsource what they do not do well.

#### **The Driving Forces Behind Outsourcing**

At a meeting of the Chicago Chapter of the Society for Information Management, Mel Bergstein of DiamondCluster Int'l. gave an overview of outsourcing. His main message was that outsourcing is another step in the evolution of the IT field.

Outsourcing descended on IS departments as a follow-on to the merger and acquisition activities in the 1980s, says Bergstein. In the 1960s, only 10 percent of the U.S. economy had global competition. In the 1970s, that rose to 70 percent. In response, companies had to focus on core businesses in the 1980s, which led to the huge amount of merger and acquisition activity. This activity was also driven by a new market for corporate control. High-yield bonds allowed a few people to buy a company and leverage it with debt. Companies were "priced" based on their shareholder value; that is, their discounted cash flow.

These two drivers – focus and value – are still leading companies to restructure and focus on core businesses by asking themselves, "Where do we really add value?" As examples, some apparel companies no longer cut, sew, manufacture, or distribute goods because they see their core businesses as design and marketing. Likewise, some publishers no longer manufacture books. They manage and finance projects, and outsource everything else.

Thus, outsourcing is part of the drive for focus and value, and it is not solely an IT issue, says Bergstein; it is a business issue. Because top management must stress value, they must consider outsourcing in all their functions.

#### **Changing Customer-Vendor Relationships**

IS outsourcers perform the same activities for a company that an IS organization performs in-house. Over time, the amount of work done by outsiders has increased, says Bergstein, as the following expansion in customer-vendor relationships illustrates.

Traditionally, IS organizations bought professional services, such as planning (or consulting), building or maintaining applications, building or maintaining networks, and training. They also bought products, which may or may not have included training. They also bought transactions, such as payroll check processing from a service bureau or credit reports from a credit rating service. Purchasing transactions allows buyers to shift fixed costs to variable costs and gives sellers higher margins because they take on the risks.

With the increasing use of packages and the need to integrate them to create integrated systems, companies have contracted with systems integrators. They generally handle the entire life cycle – planning, development, maintenance, and training – for major systems projects. Finally, the most bundled approach to contracting for IT services is outsourcing, where the outsourcer contracts to handle all or most of certain IT activities. The main difference between the latter two options is that systems integration is project-based, whereas outsourcing is time-based.

This five-option continuum, shown in Figure 8-2, demonstrates how the IT field has moved, says Bergstein. As you move from the more traditional professional services category (on the left) to outsourcing (on the right), four changes occur in the vendor-customer relationship:

IS management loses an increasing amount of control because more of the activities are turned over to outsiders.

Providers take more risks as they offer options on the right.

Providers' margins improve as they offer services on the right.

The importance of choosing the right provider becomes more important to the right, because more is at risk in using an outside source.

#### Outsourcing's History

In 1989, essentially only one kind of outsourcing involving IT was available. Since then, the field has expanded significantly. Here is a glimpse of its history; based largely on attendance at the semiannual conferences of the Sourcing Interests Group, founded and led by Barry Wiegler since 1991.

##### IT Outsourcing.

IT outsourcing essentially began with "big bang" deals, or mega-deals, which consisted of outsourcing all a company's data center operations for up to 10 years. These deals involved selling existing equipment to the outsourcer, transferring all software licenses, moving significant numbers of in-house IS personnel to the outsourcer's payroll, negotiating how the outsourcer would help in the transition and which party would carry which costs, establishing desired service levels and ways to measure performance, and specifying every single service to be provided – because if it wasn't in the contract, it would be an added cost.

In those early days, the goal of these large data center contracts was purely financial. Companies wanted to remove the huge IT infrastructure investments from their books and shift those fixed costs to variable costs; and they wanted to save money, generally about 15 percent. The deals were front loaded, with the outsourcers losing money or breaking even the first year or two, but then becoming profitable after that, as the cost of technology dropped, as they leveraged licenses across clients, as they shared expertise across clients, and as they invested in productivity tools that made them more efficient.

Several problems occurred, though. An "us versus them" mindset often set in because neither the clients nor the outsourcers handled the transition well. A lot of finger pointing took place as outsourcers tried to charge for services clients thought were included in the contract. In addition, service levels did not always live up to expectations, or interpretations of the contract language differed.

Furthermore, cultures clashed. Former employees might have kept their same desk, but once they became an employee of the outsourcer, they became a provider and were treated differently. Users had higher expectations of outsourcers than of their IS organizations. In short, companies learned that managing the relationship was really the tough job. Formerly, they had thought that negotiating the deal was the difficult part, so they had not carefully defined governance structures, that is, how the relationship would be managed.

2 Today, the IT outsourcing industry has matured. Providers have learned that heavy-handed treatment of clients can backfire. They are much more careful in transition planning. Clients' attorneys have learned what is important in a contract and where the pitfalls lie. Those early contracts have been renegotiated, and although the clients may not have changed providers, they have generally become more adept at renegotiating because they now know what they really need.

Of course, not all outsourcing deals were mega-deals. But even the small deals felt like a big bang to the employees who moved to the outsourcer.

##### Transitional Outsourcing.

In the early 1990s, a new type of computing arose: client-server computing, as noted in Chapter 5. IT outsourcing had been around for a few years, so CIOs with their hands full supporting legacy systems looked into using outsourcing to transition to client-server computing. They chose one of two routes. Either they outsourced maintenance of their legacy systems so their staff could concentrate on building new client-server systems or they outsourced client-server development to specialists and kept maintenance in-house. In either case, once the new systems were brought in, the legacy systems they replaced were shut down.

Then, in the late 1990s, when the immense size of Y2K compliance surfaced – to retrofit old applications so they would work after the Year 2000 – most companies outsourced as much of their Y2K work as they could. Because of the enormous volume of work, offshore outsourcing to India, Ireland, and other countries grew significantly. Unlike traditional IT outsourcing, however, contracts were generally shorter and did not include operations. This project-based outsourcing has been called transitional outsourcing.

##### Best-of-Breed Outsourcing.

All through the 1990s, IS departments outsourced different pieces of their work – mainly infrastructure support, as noted in Chapter 2 and IS Lite. However, CIOs learned that although selecting one outsourcer with broad capabilities might be easiest to manage, no single company was best in class in all areas. Thus, selective

outsourcing began, where one outsourcer handled desktop operations, another data center operations, and a third network management. Even though the concept was good for getting the best-of-breed providers, coordination among multiple providers became a nightmare.

A more recent trend has been collaborative outsourcing, where one company becomes the prime contractor for numerous facets of IS operations but some of the work is provided by other ESPs. Often an operations partner, a development partner, and a telecom partner collaborate to bid on the work, but one is the prime partner. Thus, teams of large ESPs now bid against other teams for contracts. In some cases, these contracts take on quite a bit more than simply operations; the work includes development of new systems as well. Best-of-breed outsourcing has perpetuated the tradition of long and complex outsourcing contracts.

### 3 Shared Services.

When IT outsourcing began to gain credibility, executives wondered, "Can we get the same economies of scale by pulling disparate non-core functions together into one shared services group?" In many cases, they felt they could. So they "insourced" to themselves, creating a shared services organization to handle such functions as IT, legal, facilities management, real estate, mail room, finance, and on and on. The goal was to improve efficiencies and save money. Generally, companies created a center of expertise in each area, with all the centers reporting to one shared services vice president.

IT was not always included, but, as in the case of MeadWestvaco Corporation (see Chapter 1), it was. Some executives believe having IT in shared services gives them the ability to leverage the IT underpinnings of the other services. Shared services also centralizes the management of outsourced functions because, in many cases, the functions are centralized and then outsourced. Shared services groups have become adept at negotiating and managing contracts and supplier relationships because these tasks are a main part of their job.

### 2 Business Process Outsourcing.

As the IT outsourcing field matured, data center outsourcing, desktop outsourcing, and other standard IT outsourcing areas became so well understood that they became like commodity services, hence profit margins dropped as the number of competitors rose. To move into higher-margin services, ESPs began specializing in specific functional areas, offering to handle specific business processes as well as their IT underpinnings. This business process outsourcing (BPO) is defined as outsourcing all or most of a reengineered process that has a large IT component.

Improving a non-core process by tapping the expertise and investments of a provider that focuses solely on that process (rather than cut costs) has been the main goal in BPO. Companies are outsourcing logistics, customer service, and many other essential, yet peripheral, functions to the experts.

Balboa Travel, a travel agency in San Diego, California, handed over its ticket accounting to Unisys. Each week, travel agencies must report the tickets they have sold to the Airline Reporting Corporation. The process is important, yet burdensome, and the president of Balboa Travel did not want to hire a programmer to maintain such a reporting system, which is what he would have had to do if he had not outsourced the work. Unisys provides him a more sophisticated service than he could afford in-house. It lets him offer his clients – corporate travel departments – reports about their employees' travel habits via an extranet. Balboa is also able to do data mining on its ticket sales to uncover trends that will help it offer new and better travel services.

As is obvious, BPO moves IT-based outsourcing out beyond the IS organization; it involves business units as well. BPO outsourcing is often quite a bit more complex than IT outsourcing because it requires clients to change their business processes to fit with the processes of the service provider. Furthermore, some clients want to retain parts of a process, so complex coordination may be necessary between the two firms as well.

BPO has brought a mindset change to the field. Whereas IT outsourcing moved suppliers and customers closer to one another in terms of working together, the two parties still do not have the same goals. Clients want to save more money, whereas outsourcers want to make more money.

In BPO, though, when providers take over an entire process, they can be measured and paid based on outcomes rather than transactions. Outcome-based outsourcing gives the two parties common goals. The client focuses on "what" needs to be done; the provider focuses on "how" to do it.

Rita Terdiman noted this phenomenon in a speech to the Sourcing Interests Group. As shown in Figure 8-3, BPO deals aim to be more like joint ventures and alliances, moving them to the right side of the chart. Therefore, trust, joint financial investments, and partnering are part of the deal.

As an example of BPO, consider ANZ Banking Group and its outsourcing of its procurement function, as described in a Sourcing Interests Group research report.

### E-business Outsourcing.

With the arrival of business use of the Internet, outsourcing has enabled companies to quickly get Web sites up and handling business. In large companies, e-business outsourcing started with marketing departments outsourcing the development of their corporate Web site. Once developed, IS took over operations. However, in dot-coms and Internet-based operations, outsourcing all or most of the IS function has been the preferred mode of operation for several reasons. Even with the dot-com crash, this is still a legitimate way to mobilize for e-business.

One, outsourcing allows a company to move fast. When a firm cannot spend a year developing a system; outsourcing the e-business infrastructure can help it get a site up and running within months, perhaps weeks. Two, companies can remain flexible, which can mean staying small and focusing only on a few key functions. Generally, IT has not been seen as a core differentiating area when tailorable off-the-shelf products and services have been available. Three, outsourcing does not tie up a firm's funds in computer and networking equipment, which could become obsolete fairly soon. A company can rent rather than buy. It can draw on best-of-breed as well as change course quickly if need be, swapping out one ESP and swapping in another, to keep pace with the market.

Unlike traditional IT outsourcing, with e-business outsourcing, machines do not need to be purchased from the client, personnel do not have to be moved, and software licenses do not have to be transferred. The outsourcing starts from scratch. Some large companies have followed this route as well to get into e-business, for the same reasons.

Utility computing.

E-business outsourcing and IT outsourcing are combining into a new form of managed services outsourcing that is being referred to by various names: utility computing, on-demand computing, virtual data centers, autonomic (self-healing) systems, and grid computing (a take-off on the term electricity grid). This idea is that computing power can be treated like electricity: You plug in and pay only for what you use.

Numerous vendors, especially IBM, HP, and Sun, are promoting access rather than ownership. By supplying computer processing, network bandwidth, and applications on demand, they are selling the idea of turning clients' fixed IT costs into variable costs. Clients can then scale their IT infrastructure or handle spikes in computing demand without having to buy more IT assets.

To educate themselves and their business on the utility computing concept, CIOs can test it by using an IT outsourcer in a high-need area, such as e-mail, security, or redundancy, notes Thickins. It is important to understand the amount of consulting this form of outsourcing requires, the chargeback mechanism (for only paying for use), the contract terms, where it could provide benefits, and so on. Mobil Travel Guide illustrates the convergence of e-business and IT outsourcing into utility computing in one of its high-need areas: its Web site.

Thus, in 15 years' time, IT outsourcing has expanded significantly, from outsourcing data center operations to outsourcing business processes and from domestic outsourcing to offshoring, which is discussed shortly.

#### 4 Managing Outsourcing

Numerous aspects to managing outsourcing need to be handled well to create a successful working relationship. Here are just four – organizational structure, governance, day-to-day working, and supplier development. All are based on research reports published by the Sourcing Interests Group.

Organizational Structure.

Managing outsourcing is different from managing internal staff because, for one thing, it is a joint effort between parties that may not have the same goals, as noted earlier. Therefore, during contract negotiations, the two parties need to figure out and negotiate how they are going to jointly manage the contract they sign. In fact, governance needs to be explicitly addressed in the contract.

Typically, parties establish layers of joint teams. A top-level team of a couple executives from both companies has the final word on conflict resolution. An operational team, with members from both companies, oversees day-to-day functioning. They hold a formal meeting, say, once a week to once a month; but they are generally in daily contact. Also, some joint special-purpose teams may be created from time to time to deal with pressing issues. Some companies have ongoing standing committees, such as a pricing committee or a change management committee, to oversee the use of formal change management procedures.

Although joint committees are a common management structure, each side needs a single executive in charge of their side of the relationship. On the client side, this executive is the relationship manager. This job position has not been prevalent in IS departments, but we believe it is going to become the norm as companies move toward IS Lite. Needless to say, the skills of a relationship manager are far different from those of a data center

manager. A relationship manager needs to be good at negotiating, cajoling, and being an effective liaison between end users and service providers.

To illustrate how one company has managed its outsourcing, we look at Eastman Kodak Company because it created a thoughtful and effective governance structure. The following description comes from the Sourcing Interests Group; it focuses on the outsourcing between Eastman Kodak and IBM Global Services.

#### Governance.

The foundations of governing an outsourcing relationship are laid out in the contract, which can be hundreds of pages long (with appendixes). A major governance item in the contract is the service level agreements (SLAs) because they are used to gauge supplier performance. For every contracted service, its SLA spells out responsibilities, performance requirements, penalties, bonuses, and so on. Completeness is an important attribute of good SLAs; generally everything should be detailed, perhaps even with times of deliveries, who will deliver what to whom, and so on.

Another important component of SLAs is metrics. An SLA needs to be measurable to be of use. Establishing metrics can be tricky because, in many cases, IS departments have not kept good measures of their own performance. In BPO, the situation is even worse; companies do not know how many people are in the process, departmental costs do not reflect overhead or IT costs, and so on. Measures are needed to establish benchmarks, against which vendors want to demonstrate improvements. Clients also need metrics to negotiate better deals. Clients who do not know their own performance levels negotiate from weakness; they know less than the vendor because they have not kept track of details, and vendors are not apt to correct mistaken impressions. Furthermore, they are likely to overlook important details, which will later cost them money.

In addition to SLAs, parties establish simple governance rules to be used when either party is making a decision, so that both are "singing from the same hymnal." Most parties in strong relationships say they put the contract in the drawer after it has been signed and work from trust and agreed-upon rules. It is only when trust in one another breaks down that they turn to the contract. Figure 8-4 lists some governance rules from a number of different enterprises.

#### Day-To-Day Working.

The Sourcing Interests Group reports provide advice from outsourcing executives on how to manage the day-to-day interactions of the two parties. Here are a few of those recommendations.

##### Manage expectations, not staff.

The outsourcer's staff is no longer under the purview of the client, so command-and-control is not a wise option – it only results in an acrimonious relationship. Facilitation becomes the mode of working. Rather than say "do this," the approach becomes "how can we solve this together?" Furthermore, the relationship managers have the important role of influencing the users' expectations so that delivery meets business objectives.

Realize that informal ways of working may disappear.

More formality is inevitable as outcomes are measured and more tightly controlled, especially if the relationship is handled strictly by the book, which happens in some cases. This change can be a real shock to people who are used to, say, getting a small job done by calling their friend "Joe" in the IS department. Once Joe works for the supplier, he may no longer be able to provide that service; he must follow the work authorization process defined in the contract. This can cause unhappiness as users see providers as "them", making them the scapegoat. The two parties need to find ways to reduce this tendency.

Loss of informal ways of working can add rigor.

Rigor frequently improves work quality. Users may think twice before requesting changes and prepare better definitions of what they want. Furthermore, better processes can streamline work, improve effectiveness, and potentially reduce unnecessary work. Service providers do introduce new discipline; the client should prepare employees for this change and assist them in adapting because it is generally best to use the provider's processes. This is one reason why transition planning is so important: to help the client personnel move to new procedures with the least disruption and disgruntlement.

Integration of the two staffs requires explicit actions.

Integration does not happen naturally. Explicit policies are likely to be needed. Some examples are to (1) grant outsourcing staff access to appropriate work areas, not unduly restrict them; (2) hold joint celebrations and social events; (3) invite each other to meetings; and (4) perhaps even have a client executive move to the provider for 2 years to learn how they work internally. However, integration generally can only go so far; the client still needs to remain in control and guide the relationship. Furthermore, the more side-by-side the parties work, the more likely they will experience scope creep in which the provider takes on more work.

The best way to manage day-to-day is to communicate frequently.

One executive said he carried around a top-10 list in his shirt pocket, which he revised every week. They were the most important items he had to handle. They kept him on focus and turned out to be his best informal management technique.

#### Supplier Development.

A topic that is receiving increased attention in the production sourcing arena – that is, buying parts and services that go into one's own products and services – is supplier development. It means assisting one's suppliers to improve their products and services, generally by improving their processes. Although supplier development has not been prevalent in IT sourcing, we think it will be. It will likely be prevalent in BPO.

Here is an example of supplier development from manufacturing, not IT, from a Sourcing Interests Group research report.

### 3 Offshoring

To round out our discussion of outsourcing, we turn to the topic receiving much attention today: sending work offshore. In the late 1990s, when labor markets were especially tight and IS organizations needed to retrofit their systems to make them Y2K compliant, the use of offshore outsourcers to maintain IT applications grew dramatically. Offshore, of course, is relative. For U.S. companies, near-shore means outsourcing to Canadian and Mexican companies, whereas offshore means Ireland, India, the Philippines, and other countries. For European companies, near-shore is Ireland, Poland, and other Eastern European countries.

Companies turn to offshoring to tap lower labor costs and an ample supply of qualified people. During the recent economic downturn, offshoring gave companies a way to cut costs. The trickle of IT jobs in the late 1990s has turned into a steady stream of white-collar work going offshore. Application maintenance and development, call centers, customer service, back-office processing, BPO, claims processing, and other functions can all be moved offshore.

Now that the U.S. economy is improving, offshoring has become a political issue, because companies are not expanding their domestic labor force as rapidly as some had expected. There's an outcry that offshoring is taking jobs from away domestic workers. Politicians are trying, at the very least, to "shame" companies into not moving domestic work abroad. But as with manufacturing jobs in the 1980s, offshoring is unstoppable because the economics are so strong. Once one company in an industry lowers its costs by moving work to lower-wage countries, its competitors need to reduce their costs as well. That may mean that they, too, need to offshore.

For all the "this is terrible" talk, offshoring might actually be good for developed countries, because increasing the living standards in other countries increases their citizens' demands for consumer products that can, in turn, be supplied by highly efficient companies in developed countries.

Furthermore, white-collar offshoring has been inevitable, argues author Daniel Altman, because service sector productivity in the United States has not kept up with manufacturing sector productivity. In the early 1950s, each service sector employee produced about \$39,000 in output (in 2000 dollars), he notes; in manufacturing, the output was \$48,000. Now, 50 years later, service productivity has increased to \$54,000 (a 47 percent increase), whereas manufacturing productivity is at \$207,000 (a 330 percent increase)!

Manufacturers have faced international competition, whereas service firms have not. Manufacturers have been forced to increase their productivity to stay in business – mainly by increasing the quality of their products, notes Altman. It should come as no surprise that service companies are tapping cheaper sources of labor now that they can because of globally available telecommunications technology, he notes. This global competition in services will force American companies to increase the productivity of their workforce in the same way and increase the quality of their services, Altman believes.

Offshore outsourcing differs in some unique ways from domestic outsourcing. In two reports, the Sourcing Interests Group (SIG) explores IT offshoring and the areas CIOs and their staff need to consider. Here are four points from those reports.

#### Offshoring Options Are Broadening.

India has become the premier IT and BPO offshoring country because its huge, highly educated workforce speaks English and is hardworking. College graduates in India apply for jobs that only high school graduates will apply for in the United States, such as working in a customer service center. Hence, the quality of work is often higher, while the labor costs are lower. Furthermore, in IS, all of the major IT outsourcers in India have achieved the highest level (Level 5) in the Capability Maturity Matrix, which means their processes produce high-quality software. Most IS organizations in the United States and in other developed countries are not even



up to Level 3. So the quality of software from these Indian firms is very high-as long as the specifications for the work are clear and truly what the customer wants.

After having good initial IT experiences in India, client companies are offshoring higher-value white-collar work beyond short-term software maintenance projects to ongoing software maintenance work, new development projects, and even BPO. The types of firms offering offshore options are also broadening. U.S. outsourcers are building their own offshore capabilities or acquiring existing offshore providers, often in India, sometimes in China, to lower their own costs. Client companies not wanting to deal directly with unfamiliar overseas providers can now tap offshore resources through onshore outsourcers. In fact, as noted in the SIG reports, all large IT outsourcing deals will likely have an offshore component. Global sourcing is becoming the norm.

#### Both Parties Need Cultural Training to Bridge Cultural Differences.

Offshoring does bring cultural differences into sharp relief, which can hinder success unless properly handled. Outsourcers and offshore advisory firms (who advise clients on the offshore marketplace) now realize that both parties need cultural training to overcome management and communication gaps.

For clients, offshorers now routinely put client employees who will deal with the offshore employees through a cultural integration program. Here is a brief description of the program at Syntel, a large U.S.-based IT provider that uses an offshore (India) model, as recounted in one of the SIG reports:

We initially hold a 1.5-day workshop, and then a follow-up orientation a few months later. We find that clients often restructure their operations during the first few months, sometimes bringing in new people. These new people need the orientation, also, which is why we repeat the orientation.

In the workshop we talk about how lifestyles and cultures differ between the US. and India. In fact, we tailor each workshop to the client's location because Midwest employees differ from New Yorkers, for example. We illustrate, for instance, how the same words have different meanings in the two countries. We point out differences in dress, foods, even eating times. We show a video of these differences, we have people wear the native dress, and we demonstrate how Indians prepare a lunch, which we then serve. We have learned that U.S. employees often have a third-world view of India that is not accurate. Our goal in the orientation is to familiarize them with today's modern India.

On the provider side, offshorers often put their own employees through an accent neutralization program if they have an accent the client country will find difficult to understand. This practice is especially prevalent for Indian providers with American clients. Providers also teach their employees about the client's company culture (going so far as to hang the same posters on the walls as the client has on its walls) and the client country's culture (going so far as to familiarize employees with holidays and keep them up-to-date on sports teams and major sporting events).

#### Communication Issues Need to Be Addressed from the Outset.

Different cultures have different communication norms, even when they speak the same language. These differences show up immediately when people from two cultures try to negotiate an offshore contract. The differences need to be addressed, beginning with the negotiations.

Here is just one example. The word "yes" in Asia typically means, "I hear what you are saying," whereas in the United States it means "I can do what you ask" or "I agree with you." When conversing, an Indian might nod his head "yes," which an American might misinterpret as an agreement. The American is then surprised at the next meeting when there has been no agreement after all.

There are substantially more risks in negotiating offshore than onshore contracts because of such communication misunderstandings, notes the global law firm of Mayer, Brown, Rowe & Maw LLP in a SIG report. The firm uses many mechanisms to mitigate these risks. Here are just five:

- Avoid colloquialisms, such as sports analogies ("we've got to punt on this"), because these statements do not bridge cultures.
- Simplify communications by using short, concise sentences with common words, rather than the typical legal practice of lengthy, convoluted sentences and paragraphs.
- Have the offshore provider write a "statement of work," to gauge their understanding of what they think they are being asked to do.
- Get all commitments in writing.
- Include a person on the client negotiating team who knows the offshore country culture so that he or she can explain the country's norms to the client team as issues arise and prevent the offshorer's team from holding side discussions in their own language at the bargaining table. A "country-wise" member can significantly change negotiating dynamics.

Communication issues continue throughout offshore relationships. This is why migrating the work, which is often fairly quickly defined in onshore contracts, can require lengthy discussions in offshore deals. Common understandings do not exist, so every step in moving work from one country to another, and who pays for what, must be agreed upon. Furthermore, country laws and norms must be taken into account.

#### Country Laws Need to Be Followed.

Offshoring contracts must bridge a myriad of differences between cultures. To preclude dealing with some legal issues, offshore vendors now typically sign their contract in the client's country. For example, as noted earlier, Syntel is a US. company. Its contracts with US. clients are signed in the United States, and are governed by US. law.

Even so, clients need to be aware of data privacy laws, which may not allow customer data to leave its home country or enter certain other countries. The European Union (EU), for instance, does not allow personal data to be moved outside of EU countries. Likewise, clients need to understand the enforceability of intellectual property laws in the outsourced country. Lack of enforceability may preclude moving some IT work to a country, such as China, which does not have enforceable intellectual property laws. Likewise, taxation, employment, immigration, and other laws can significantly affect an offshoring arrangement. All of these must be taken into account.

The list of issues in offshoring can be truly mind-numbing. However, as the trend continues to grow, common practices are arising that make the negotiating task less daunting and the resulting arrangement less risky and more likely to succeed. As an example from the SIG report of a company using offshoring for its IT maintenance work, consider Exult, which is itself an example of a BPO company. It manages human resources for other firms.

#### Use Offshoring to Advantage.

A main criticism of offshoring is that it decreases the skills and know-how of the client's IS organization. This need not be so. Kate Kaiser of Marquette University and Stephen Hawk of the University of Wisconsin-Parkside<sup>8</sup> describe an 8-year arrangement between an unnamed financial services firm in the United States and an unnamed IT outsourcer in India. In their recent article in MIS Quarterly Executive, the two authors note that the U.S. firm wanted to reduce its system development costs but also increase its in-house IS staff's knowledge—a fairly unusual dual goal. To do so, the two firms have evolved their relationship to "IT cosourcing," which Kaiser and Hawk define as "when the vendor and client collaborate so closely that the vendor can replace or augment the client's IT competencies." In essence, resources from both firms are used to meet the client's needs. Project teams are mixed, and team leadership can come from either firm - both of which require the vendor to have a large onsite presence. Two mechanisms, in particular, have ensured that the U.S. firm's IS staff gain competencies rather than lose them: formalized knowledge transfer between the two firms and a dual projectmanagement hierarchy.

To formalize knowledge transfer from the Indian staff to the US. staff, US. staff members are sometimes formally assigned to projects to learn from an Indian mentor. The two firms initially made such assignments informally, but that approach did not improve skills. So the two firms formalized learning-oriented mechanisms, such as mentoring. At the US. firm, each IS employee has a development plan, which includes career goals and steps to achieve the needed skills. For the Indian firm, job assignments with the US. client include such tasks as mentoring a specific US. team member in specific areas. The cost of mentoring is included in the cost of the project.

To create the dual project management hierarchy, the hierarchies of the two firms now mirror each other to improve communication across them. Furthermore, the tiers of leadership in projects can come from either firm, depending on the circumstances. The Indian firm may lead to provide mentoring. The US. client may lead for business reasons. Again, the roles of each job are formally agreed upon, and mentoring or other knowledge-transfer skills are part of the cost of the project.

Both mechanisms increase the cost of outsourcing, note Kaiser and Hawk, because, for one thing, dual leadership requires more Indian staff to be located in the United States. But the additional cost improves the in-house staff's technical and application skills, which the firm believes is well worth the added expense, because they highly value their employees' career development.

#### Redefine Services Using Offshoring, Automation, and Self-Service.

Uday Karmarkar,<sup>9</sup> research director of UCLA's Center for Management in the Information Economy, like author Daniel Altman<sup>7</sup> believes that outsourcing of services is inevitable and that the real concern of service firms should be their loss of competitiveness, not the loss of jobs in their own country. He believes that the service

economy is in the midst of restructuring itself, which is terribly painful when it happens in any industry. Offshoring, automation, and self-service are all combining to cause "the industrialization of services," he believes. Like manufacturing firms before them, service firms therefore need to find new ways to add value. He suggests looking in five places for determining new strategies for surviving and using offshoring, automation, and self-service to execute these strategies.

Understand customers.

Companies that understand niches of customers, and serve them well, will themselves do well, believes Karmarkar, especially as they move their services to the Web. Edmunds, a company that has published books for car buyers for many years, now focuses on its car-buyer Web site. The site has been so well designed that it has won numerous awards. Edmunds understands its customers and caters to them, states Karmarkar.

Understand demographics.

Look for underserved niches, like Wells Fargo Bank has done in offering specific services to the Hispanic population in the western United States. The bank now opens 22,000 Hispanic accounts a month by understanding and catering to this growing group, notes Karmarkar.

Stay in touch with customers.

Do not outsource customer service, Karmarkar recommends. Many have done it and found that they lose touch with their customers. Some have realized the effect and brought the function back in-house. A far better strategy is to deliver responsive and error-free service, he believes.

Offer end-to-end service.

Customers are most interested in buying services from firms that offer end-to-end services. For example, Virgin Atlantic Airways provides its businessclass passengers with limousine service to the airport and drive-through check-in. People are willing to pay for such pampering to relieve themselves of hassles. This is an important strategy some service firms are delving into by using data mining to understand preferences of customer clusters and then catering to those preferences, notes Karmarkar.

Dominate the screen.

As information moves online, companies are vying to control "the screen," that is, where the information ends up. Service companies have a shot at dominating the screen, because the design of the service and the interface (rather than the technology or the appliance) will determine success, states Karmarkar. For example, NIT DoCoMo, the telecommunications company in Japan, knows how to sell content on cell phones and appliances - from service providers to consumers. Due to this skill, NIT has claimed a disproportionate degree of control over the information chain, notes Karmarkar; it dominates the screen. In short, service companies need to understand what kinds of information various clusters of customer want and cater to them. For many service firms, that catering will be via the Web and mobile devices, using offshoring, self-service, and automation where each can contribute the most value.

5 In conclusion, outsourcing has become a strategic alternative. With the rapid pace of business change, the best hope of many enterprises is to tap the expertise of companies that are keeping pace rather than trying to do everything themselves. That is why so much inter-company collaboration is taking place. However, outsourcing does not mean relinquishing responsibility. In fact, taken to its extreme, it can mean devoting resources to assist suppliers improve their processes. We believe it is a coming focus in the world of IS Lite.

**Biilage V**      **Schema tekst 2: 'Outsourcing Information Systems Functions'**

| Topic                                 | Concept 1                              | Concept 2                                 | Concept 3                         | Concept 4                   |              | Relatie                   | Synoniemen            |                          |                     |
|---------------------------------------|----------------------------------------|-------------------------------------------|-----------------------------------|-----------------------------|--------------|---------------------------|-----------------------|--------------------------|---------------------|
| evolution to outsourcing IS functions | evolution to outsourcing IS functions  | driving forces                            |                                   |                             | →            | caused by                 | drivers               |                          |                     |
|                                       |                                        |                                           |                                   |                             |              |                           | driving forces        |                          |                     |
|                                       |                                        |                                           | definition                        |                             | →            | is defined as             | means                 |                          |                     |
|                                       |                                        | drive for focus and value                 |                                   |                             | →            | includes                  | drivers               |                          |                     |
|                                       |                                        |                                           |                                   |                             |              |                           | focus                 |                          |                     |
|                                       |                                        |                                           |                                   |                             |              |                           | value                 |                          |                     |
|                                       |                                        | evolution of outsourcing services         |                                   |                             | →            | caused by                 |                       |                          |                     |
|                                       |                                        |                                           | traditional professional services |                             | →            | includes                  | professional services |                          |                     |
|                                       |                                        |                                           |                                   |                             |              |                           | traditionally         |                          |                     |
|                                       |                                        | contracting for outsourcing IT activities |                                   |                             | →            | includes                  | integrated services   |                          |                     |
|                                       |                                        |                                           | IT outsourcing characteristics    | outsourcing characteristics |              |                           |                       |                          | systems integrators |
|                                       |                                        |                                           |                                   |                             |              |                           |                       |                          | contracting         |
|                                       |                                        |                                           |                                   |                             |              |                           |                       |                          | contracted          |
| customer-vendor relationships         |                                        |                                           |                                   | →                           | changes      | vendor-customer           |                       |                          |                     |
|                                       |                                        |                                           |                                   |                             |              | relationship              |                       |                          |                     |
|                                       |                                        |                                           |                                   |                             |              | provider                  |                       |                          |                     |
| beginning of IT outsourcing           |                                        |                                           |                                   | →                           | results in   | it outsourcing            |                       |                          |                     |
|                                       | problems of first outsourcing projects |                                           |                                   |                             | →            | includes                  | problems              |                          |                     |
| IT outsourcing characteristics        | outsourcing characteristics            | mature outsourcing                        |                                   |                             |              | →                         | includes              | matured                  |                     |
|                                       |                                        |                                           |                                   | transitional outsourcing    |              | →                         | started with          | transitional outsourcing |                     |
|                                       |                                        | best-of-breed outsourcing                 |                                   | →                           | started with | best-of-breed outsourcing |                       |                          |                     |
|                                       |                                        |                                           |                                   |                             |              |                           |                       |                          |                     |

|                                  |                                  |                              |                              |                             |                                   |
|----------------------------------|----------------------------------|------------------------------|------------------------------|-----------------------------|-----------------------------------|
|                                  |                                  |                              | selective outsourcing        | → includes                  | selective outsourcing             |
|                                  |                                  |                              | collaborative outsourcing    | → includes                  | collaborative outsourcing         |
|                                  |                                  |                              |                              |                             | prime contractor                  |
|                                  |                                  |                              |                              |                             | ESP                               |
|                                  |                                  | business process outsourcing |                              | → includes                  | business process outsourcing      |
|                                  |                                  |                              |                              |                             | bpo                               |
|                                  |                                  |                              | definition                   | → is defined as             | defined                           |
|                                  |                                  |                              | fitting processes outsourcer | → requires                  | fit                               |
|                                  |                                  |                              |                              |                             | processes of the service provider |
|                                  |                                  |                              | outcome-based outsourcing    | → can be done by the way of | outcome-based                     |
|                                  |                                  |                              |                              |                             | outcomes                          |
|                                  |                                  | e-business outsourcing       |                              | → includes                  | e-business outsourcing            |
|                                  |                                  |                              |                              |                             | e-business                        |
|                                  |                                  |                              | reasons to outsource         | → caused by                 | reasons                           |
|                                  |                                  |                              | characteristics              | → is characterized by       | characteristics                   |
| special IT outsourcing solutions | special IT outsourcing solutions |                              |                              |                             |                                   |
|                                  |                                  | shared services              |                              | → includes                  | shared services                   |
|                                  |                                  | offshoring                   |                              | → includes                  | offshore                          |
|                                  |                                  |                              |                              |                             | offshoring                        |
|                                  |                                  |                              |                              |                             | near-shore                        |
|                                  |                                  |                              | offshoring options           | → includes                  | offshoring options                |
|                                  |                                  |                              | special issues               | → includes                  |                                   |
|                                  |                                  |                              | bridge cultural differences  | → includes                  | cultural differences              |
|                                  |                                  |                              |                              |                             | culture                           |
|                                  |                                  |                              |                              |                             | differ                            |
|                                  |                                  |                              | communication norms          | → includes                  | communication norms               |
|                                  |                                  |                              |                              |                             | communication                     |

|                           |                           |                          |              |                             |                                                                |
|---------------------------|---------------------------|--------------------------|--------------|-----------------------------|----------------------------------------------------------------|
|                           |                           |                          | country laws | → includes                  | norms<br>country laws<br>laws                                  |
|                           |                           | advantages               |              | → includes                  | advantages<br>advantage                                        |
|                           |                           | redefine services        |              | → can be done by the way of | redefine services<br>new strategies                            |
| management of outsourcing | management of outsourcing |                          |              |                             | managing outsourcing                                           |
|                           |                           | organizational structure |              | → requires                  | organizational structure<br>management structure               |
|                           |                           | governance               |              | → requires                  | governance<br>governing<br>service level agreement             |
|                           |                           | day-to-day working       |              | → requires                  | SLA<br>day-to-day working<br>day-to-day<br>daily<br>every week |

## **Biilage VI      Tekst 3: 'Outsourcing as an Economic Strategy'**

### **1 Outsourcing as an Economic Strategy**

If organizations can effectively define and measure information technology costs and benefits, as we discussed in the preceding sections, they are in a better position to effectively manage their IT. On the other hand, they still may not be able to manage it as well as firms that specialize in managing IT. For such organizations, the most effective strategy for obtaining the economic benefits of IT and controlling its costs may be outsourcing, which is obtaining services from vendors rather than from within the organization.

Information technology is now a vital part of almost every organization and plays an important supporting role in most functions. However, IT is not the primary business of many organizations. Their core competencies – the things they do best and that represent their competitive strengths – are in manufacturing, or retailing, or services, or some other function. IT is complex, expensive, and constantly changing. It is difficult to manage IT, even for organizations with above-average management skills. Because of these considerations, some organizations decide to use outside vendors rather than internal IS units as their primary source of IT services. The idea of outsourcing itself is not new. Organizations have always faced the choice of buying goods and services externally or producing them internally. The decision usually considers two factors: (1) which source is less expensive? and (2) how much control is necessary? Most organizations buy their power from electric utilities because this is cheaper than operating their own generating equipment. However, a few (hospitals, for example) have their own emergency power-generating systems. These internal systems give them more control over their operations; they do not have to shut down if a disaster causes a power failure.

Outsourcing IT functions, such as payroll services, has been around since the early days of data processing. Contract programmers and computer timesharing services are longstanding examples. What is new is that, since the late 1980s, as illustrated in the opening case, many organizations are outsourcing the majority of their IT functions rather than just incidental parts. The trend became very visible in 1989 when Eastman Kodak announced it was transferring its data centers to IBM under a 10-year, \$500 million contract. This example, at a prominent multibillion-dollar company, gave a clear signal that outsourcing was a legitimate approach to managing IT.

In a typical situation, the outsourcing firm hires the IS employees of the customer and buys the computer hardware. The cash from this sale is an important incentive for outsourcing by firms with financial problems. The outsourcer provides IT services under a five- to ten-year contract that specifies a baseline level of services, with additional charges for higher volumes or services not identified in the baseline.

Many smaller firms provide limited-scale outsourcing of individual services, but only the largest outsourcing firms can take over large proportions of the IT functions of major organizations. In the mid-1990s, IBM, EDS, and Computer Sciences Corp. were winning approximately two-thirds of the largest outsourcing contracts. PricewaterhouseCoopers (1999), in its 1998-99 survey of European private banks, found that nearly a quarter of survey respondents had either implemented or were considering implementing the outsourcing of their IT systems.

2 Offshore outsourcing of software development has become a common practice in recent years. About one-third of Fortune 500 companies have started to outsource software development to software companies in India. This trend of offshore outsourcing is largely due to the emphasis of Indian companies on process quality by adhering to models such as Software Engineering Institute's Software Capability Maturity Model (SW-CMM) and through ISO 9001 certification. India has fifteen of the twenty-three organizations worldwide that have achieved Level 5, the highest in SW-CMM ratings. For further details on offshore outsourcing, see Cusumano (2000) and Paulk (2000).

In addition to the traditionally outsourced services, Brown and Young (200) identify two more scenarios for future outsourcing: creation of shared environments (e.g., exchanges, portals, e-commerce backbones), and providing access to shared environments (e.g., applications service providers (ASPs), Internet data centers). For example, Flooz.com, an online gift-currency store, outsourced its storage requirements to StorageNetworks, a storage service provider (Wilkinson, 2000). See outsourcing-center.com for details on advances and practices in outsourcing of various types of services such as ASPs and business process outsourcing (BPO).

### **3 How to manage the risks of outsourcing?**

Organizations should consider the following strategies in managing the risks associated with outsourcing contracts (Clemons, 2000):

Understand the project. Clients must have a high degree of understanding of the project, including its requirements, the method of its implementation, and the source of expected economic benefits. A common characteristic of successful outsourcing contracts is that the client was generally capable of developing the application but chose to outsource simply because of constraints on time or staff availability.

Divide and conquer. Dividing a large project into smaller and more manageable pieces will greatly reduce programmatic risk and provides clients with an exit strategy if any part of the project fails.

Align incentives. Designing contractual incentives based on activities that can be measured accurately can result in achieving desired performance.

The implementation of Internet-based technologies and especially of e-commerce encourages the use of outsourcing, as can be seen in the following IT at Work example below.

A special outsourcing consideration is the implementation of extranets. Implementing an extranet is very difficult due to security issues and the need to have the system be rapidly expandable. (Some companies report 1,000 percent growth for EC activities in a year; e.g., see hotmail.com.) General Electric Information Services (geis.com), an extranet outsourcer, charges between \$100,000 and \$150,000 to set up an extranet, plus a \$5,000/month service fee. However, users of these services admit an ROI of 100 to 1,000 percent. For details see Duvall (1998).

#### 4 Outsourcing Contracting

Some organizations may decide to outsource because they just do not have the skills to manage the IT function, or because they need to sell off IT assets to generate funds. Various sources (e.g., Marcolin and McLellan, 1998) offer outsourcing recommendations, such as:

Write short-period contracts. Outsourcing contracts are often written for five- to ten-year terms. Because IT and the competitive environment change so rapidly, it is very possible that some of the terms will not be in the customer's best interests after five years. If a long-term contract is used, it needs to include adequate mechanisms for negotiating revisions where necessary.

Subcontracting. Vendors may subcontract some of the services to other vendors. The contract should give the customer some control over the circumstances, including choice of vendors, and any subcontract arrangements. Selective outsourcing. This is a strategy used by many corporations who prefer not to outsource the majority of their IT (like Kodak), but rather to outsource certain areas (such as connectivity or network security).



## **Biilage VII    Tekst 4: 'Strategic dimensions of IT outsourcing'**

### **1 Strategic dimensions of IT outsourcing**

by Leslie P. Willcocks and Mary C. Lacity

IT outsourcing has outlived the five-year period typical of a management fad. Global market revenues, which were \$9bn in 1990, look set to grow to \$121bn by 2001. The underlying compound annual growth rate was 15-20 per cent between 1992 and 1999; in the US and the UK, the leading markets, there have been more dramatic rises in years when several major deals have been signed.

From an initial focus on cost reduction, outsourcing is fast becoming a complementary, routine mode of managing IT. On our own estimates, 30-35 per cent of most large organizations' IT budgets will be managed through outsourcing arrangements by 2002. The question "why not outsource IT?" is no longer, if it ever was, an adequate base from which to make outsourcing decisions. The real question now has to be: "How do we exploit the ever maturing external IT services market to achieve significant, strategic business leverage?"

### **2 The outsourcing report card: practices and risks**

It is clear that some organizations seek to address this issue through total outsourcing. Deals such as those at Xerox, McDonnell Douglas (now Boeing) and Continental Bank in the US, Commonwealth Bank, Lend Lease Corporation and the South Australia Government in Australia, and British Aerospace, BP Amoco and the Inland Revenue in the UK, are often referred to as "strategic alliances". Going by media coverage, one could be forgiven for believing this type of outsourcing to be the dominant approach. But upon investigation, a much richer picture emerges of the IT sourcing paths organizations are taking (see Figure 1), and many still seek tactical IT gains rather than strategic business advantage.

The dominant mode is selective sourcing, especially in the US (82 per cent of organizations) and the UK (75 per cent). A mixed portfolio, "best-source" approach typically puts 15-30 per cent of the IT budget under third-party management, with other IT needs met through buying in resources that are under in-house management ("insourcing"), and through internal IT staffing.

Many organizations (10 per cent in the US and 23 per cent in the UK) have no significant IT outsourcing contracts. These organizations perceive IT as a core strategic asset; they believe that their IT employees will strive to achieve business advantage in ways that external providers cannot match, as well as being more loyal to the business.

Total outsourcing (in which 80 per cent or more of the IT budget is under third-party management) is a minority pursuit. In the US some 8 per cent of organizations take this route, in the UK about 2 per cent; worldwide there are just over 120 such deals.

All IT sourcing arrangements have inherent risks (see Figure 1). A mainly in-house function needs to be continually assessed against the market if it is not to become expensive and unresponsive. Much depends on strong business/IT relationships if IT activity is to underpin strategic business direction.

One common underestimated factor in IT outsourcing is the associated management overhead cost. This is typically around 4-8 per cent of the total outsourcing cost even before the effectiveness of the management arrangement has been assessed.

Large-scale outsourcing deals focusing primarily on cost reduction can achieve these, but often at the expense of flexibility in IT operations and business strategy. Alternatively, incomplete contracts or negligible profit margins through over-tight contracts, can (and have) led to hidden costs or opportunistic vendor behaviour. Finally, as will be made clear below, underdeveloped approaches to "strategic partnerships" can be very risky. Indeed, many such deals have had to undergo significant restructuring after 18 to 24 months.

### **3 From tactical to strategic uses of outsourcing**

Against this background, most organizations that have outsourced IT during the present decade have pursued one of two approaches. In incremental/tactical outsourcing the primary focus has been on cost reduction, efficiencies and improved service in the way IT is run. Our 1998/99 UK and US survey suggests that such benefits have tended to come through. Moreover, learning about outsourcing matures organizational capability, positioning organizations for larger-scale, more strategic outsourcing in future.

However, one consequence of incremental IT outsourcing has been that a large minority of organizations have encountered problems impinging on business strategy. These problems are: failure on the part of the supplier to understand the business (37 per cent); misalignment between corporate strategy and IT (35 per cent); and poor strategic IT planning (24 per cent).

Other organizations, through some combination of competitive, financial, business, technical and IT skills pressures, have taken a "hard learning" route; that is, they have outsourced to a significant degree without really developing a strategic focus on their outsourcing objectives, let alone the capabilities to achieve them.

However, in our most recent research into global IT outsourcing practices, we have discerned six strategic foci, in combination or in isolation, through which organizations are increasingly attempting (sometimes with success) to achieve a business advantage.

4 Financial restructuring: improving the business's financial position while reducing or at least containing costs  
Cost efficiency remains a key organizational goal for most selective and total outsourcers. However, organizations such as British Aerospace, Continental Bank and Sears have sought long-term, significant changes in their financial position through outsourcing.

For a company that is in financial trouble, outsourcing provides a significant cash influx from the transfer of IT assets and staff, together with tax advantages; this improves cash flow and makes it possible to restructure the balance sheet and profit and loss account. As such, outsourcing represents a long-term financial strategy, supporting a turnaround in the company's financial and competitive position over time in exchange for annual outsourcing fees representing an interest rate.

Most single-supplier total outsourcing deals in the early 1990s were of this character (see Figure 1); many disappointed. For British Aerospace, protected by a highly detailed contract, outsourcing bought it time to recover its financial position. By the beginning of 1999 much more was being done with the technology, and IT efficiencies were being achieved in operational areas, though the IT budget as a whole had nearly doubled from 1993.

On a wider front, companies pursuing low-cost business strategies will best reflect these by focusing on cost reduction through IT outsourcing. However, there is evidence that this creates inflexibility in both IT operations and business strategy. Companies may well experience constraints when their business strategy moves to seeking faster time to market and rapid market growth. There is also evidence that companies with objectives other than cost reduction tend to view more of their IT as a core competence and pursue relatively less outsourcing.

5 Core competence: redirecting the business and IT into core competences

This strategy is exemplified by British Airways' aim of focusing on transportation of passengers and cargo to achieve £1bn savings and double profits in the 1996-2000 period. To achieve this, the company needed only to own its route structure, brand and (IT-based) yield management system. Almost all other things could be outsourced, leased or bought in.

By 1990, in a worsening financial situation, BP Exploration (BPX) became much more focused on the core competence argument, according to which the company's key competence was exploration, and its key intelligence in its prospectors, not in IT or accounting (say). The accounting function, including its IT, was outsourced in 1991 in a five-year, £55m deal. According to chief executive John Browne, "failure to outsource our commodity IT will permanently impair our business competitiveness". However, this was to be achieved not by traditional outsourcing but by establishing a constellation of partners.

In 1993 BPX signed three suppliers on five-year contracts, worth \$35m annually. Sema Group was selected to handle datacentre operations, SAIC for its strength in developing distributed and leading-edge systems and scientific applications, while Syncordia had the maturity, flexibility and range to manage telecommunications and WANS. These were chosen to support a strategy of becoming a more diversified production company, moving into areas in which BPX often did not have the necessary in-house IT skills.

Subsequently, the in-house team focused on business processes, on information and business value-creation areas, and the vendors on more routine applications and technology. One achievement over the five years was the move to desktop computing throughout BPX, which was widely recognized as impossible without outsourcing.

6 Technology catalyst: strengthening resources and flexibility in technology and service to underpin the business's strategic direction

Outsourcing vendors are regularly used to "fill in" for IT skills shortages, and to provide competitively priced computing power and service. However, some organizations have used outsourcing more strategically, seeing it as a significant investment in transforming a traditional IT function, and achieving a new technology agenda to underpin business strategy.

Thus in 1994 Xerox signed a 10-year global contract with EDS for an estimated \$3.2bn. It had become clear to Xerox that its business restructuring in the early 1990s had left its existing systems and skills base unable to supply the required technological support. It developed a new strategy for which a new infrastructure was required, whose hardware alone would cost \$55m. Internal service was seen as ineffective, costing a lot of money without significant returns. Outsourcing allowed Xerox to refocus IT on business-critical applications, while the vendor facilitated routine operations, applications and their support, telecommunications and the move to a client/server infrastructure.

In the chemicals and speciality products sector, DuPont looked for a similar "technology catalyst" payoff from its \$4bn, 10-year contracts signed with two vendors in 1997. Earlier cost cutting at DuPont had meant a lack of renewal of IT assets and skills, and an IT infrastructure requiring several hundred million dollars of investment. Outsourcing thus became a serious option.

DuPont looked to convert fixed IT costs into variable costs, and also hoped for improved service speed and flexibility, skills renewal, IT career development and further cost reductions. Currently one supplier has 2,600 staff providing global infrastructure and applications, while another has 400 supporting chemical applications. DuPont retained some 60 staff to oversee the contracts, while over 1,000 distributed business and technical people provide business IT leadership, and process control and research and development computing.

7 Business transition: facilitating and supporting major organizational change

Forms of transitional outsourcing have frequently been used and are generally successful. They involve temporarily outsourcing technically mature IT during a period of transition to a new technology. For example, fiberglass manufacturer Owens Corning signed a five-year contract in 1995 for a vendor to maintain legacy systems while it implemented an enterprise resource planning system at 75 sites worldwide.

However, some organizations operate even more strategically, incorporating transitional IT outsourcing into a suite of outsourcing contracts to help major organizational change. Typical areas where outsourcing has been used are during mergers and acquisitions, new business start-ups, and major devolution/restructuring of the business, including privatizations.

British Gas provides an example. Faced with privatization, deregulation and devolution in the mid-1990s, its IT director commented: "Some of the challenges have probably not been met in many other organizations in the world. We are seeing implementations in time scales thought impossible before. There is also incredible downsizing. One of the implications is you cannot do it all in-house, outsourcing is needed just to meet the challenging projects, their size, our lack of resource and the short time scales."

One effective transitional move in 1995 was to contract out all datacentres in an 18-month, £55m deal. On a larger front, in order to devolve into five autonomous businesses, British Gas needed a range of suppliers to assist in re-engineering, and in developing new billing, servicing, retail and distribution IT systems.

8 Business innovation: improving and innovating in processes, skills and technology, while mediating financial risk through the vendor, in order to achieve competitive advantage

Organizations looking to suppliers for IT or even business innovation through traditional outsourcing contracts have often been disappointed. BPX, for example, had to create an innovation fund for suppliers and BPX staff to bid for, such was the lack of innovation coming through. Yet business innovation can be a potent source of competitive advantage that external IT partners can assist in.

This has been recognized in various private finance initiatives in the UK public sector, in the "co-sourcing" deals arrived at by several major companies (for example, Rolls-Royce and Citibank) with suppliers, and in benefits funding mechanisms of the type used successfully by the California Franchise Tax Board in the US. The essential ingredient in all these deals is that, in its IT work, the supplier focuses on the client's business objectives - delivering re-engineered processes at Rolls-Royce, developing a travel agency commission system at Citibank, and improving the accuracy and revenues from tax returns at the CFTB.

The supplier takes the investment risk up front, and is paid according to pre-agreed performance measures as business benefits are realized. Performance-based contracts are perceived by suppliers as high risk. They may require a culture change in suppliers more used to fee-type deals, and at the moment account for a very small percentage of global outsourcing revenues.

9 New market: direct profit generation through joint ventures with vendor partner

Some organizations have experimented with creating a new "spin-off" supplier company from their IT function. Others have contracted with a supplier to sell jointly developed products and services to the external marketplace ("value-added outsourcing"). Another type of arrangement is "equity holding", in which a company and a supplier take share-holdings in each other; the aim is to encourage the supplier to perform well, secure the business relationship, and make investment profit.

Spin-offs are not without difficulties. The new supplier needs to build commercial and marketing skills for a competitive marketplace in which other vendors have better track records and will want to exclude the new entrant. Mellon Bank, Sears Roebuck and Boeing have had limited success with their spin-off IT companies. Like American Airlines' Sabre unit, what is needed is a core competence that will attract external customers.

Electronics multinational Philips took shares in Dutch software house BSO, and transferred its development staff, and subsequently its processing operations, into a jointly owned company, Origin. BSO had an existing reputation and customer base to build on, and Philips was a ready-made large client. Even so, it took Origin several years to build a wider outsourcing market position.

Many deals have a "value-added" element. Thus Xerox-EDS sought future shared revenues for the development of a global electronic document distribution service. Mutual Life Insurance of New York and CSC sought to market software and services to the insurance industry. But partners must truly add value by offering products and services demanded by customers in the market. In several deals this value-added element has been too marginal to the overall fee-based contract to make a difference and refocus attention on mutual business objectives. Also, when they are under pressure in other parts of the outsourcing contract, client and supplier may become less interested when they realize that commercializing and marketing a new product may require an investment nine times greater than the original development cost.

Finally, shared ownership deals such as Lend Lease-ISSC, Swiss Bank-Perot Systems, and Delta Airlines-AT&T have attractions. In principle - and in reality in the BSO-Phillips case - the mechanism of shared risks and rewards overcomes the conflicts inherent in fee-based contracts. However, the incentive of shareholder profit may operate at too high a level to influence performance on the ground. Supporting the supplier's market and profit growth can conflict with, and be at the expense of, service to the original client.

#### 10 Implementing strategic uses of outsourcing

Strategic outsourcing has taken a variety of routes to success. It requires a lot of management maturity and experience of IT outsourcing. It needs complete and creative contracting; a less long-term focus in the contracting arrangements, but a more long-term one in the relationship dimension, and very active post-contract management. A cardinal insight from our research is that organizations still expect too much from vendors and not enough from themselves; put another way, vendors are still much better at selling IT services than their clients are at buying them.

Five capabilities are necessary to pursue IT outsourcing for strategic advantage:

Business responsibility, and the ability to identify business models and value-drivers that will make innovation possible and create business advantage.

Aligning the use of the market with such analysis is fundamental. Applying supplier activity to an existing business model can at best only achieve efficiency gains.

The ability to make sourcing decisions and arrive at a long-term IT sourcing strategy; this must build in learning and take into account business, technical and economic factors.

On this front, our research identifies two proven practices in IT outsourcing. First, selective outsourcing decisions and total in-house / insourcing decisions succeed more often than total outsourcing. Second, senior executives and IT managers who make decisions together achieve success more often than when either group acts alone.

The ability to understand the IT services marketplace and suppliers' strategies and capabilities.

Our research shows two proven practices here. First, informed buying is a core IT capability for all organizations today. Second, organizations that invite both internal and external bids succeed more often than organizations that merely compare a few external bids with current IT performance.

The ability to contract over time in ways that provide incentives for the supplier(s) and ensure that you get what you think you agreed to.

Two things to note here are that short-term (4 years or less) contracts are successful more often than long-term contracts (7 years or more), and that detailed fee-for-service contracts are successful more often than other types of contract.

The ability to manage the contract in ways that secure long-term IT objectives, and achieve the required service performance from the supplier.

This is one of the weakest areas in IT outsourcing practice. IT outsourcing requires a great deal of in-house management - to determine and deliver business requirements, to ensure technical capability, and to manage external supply and IT governance. Without these, the potential for strategic gains from outsourcing can be eroded.

#### 11 Summary

While some have no significant contracts and a few put their entire budgets out to third parties, most organizations adopt a mixed approach to IT outsourcing. There are risks with both extremes - keeping everything in-house can be expensive and inflexible, while going outside can involve unanticipated overhead costs. Cost reduction, efficiency and improved service continue to drive many initiatives in the 1990s - but according to Leslie Willcocks and Mary Lacity, companies are increasingly seeking ways to gain strategic advantage from outsourcing. Examples include long-term financial restructuring, the ability to focus on core competences, the "catalyzing" effect of an external service provider, help during organizational transitions (such as mergers or restructuring), support for innovation, and joint ventures with a vendor partner. The authors conclude with a list of the capabilities required to pursue IT outsourcing for strategic advantage.

## **Bijlage VIII    Samenvattingen Tekst 1: 'Outsourcing Collaborative Partnerships'**

### **Samenvatting van de volledige tekst**

More than 400 people from Merrill Lynch, Thomson Financial (a large market data vendor), and a number of other vendors have been working feverishly on Merrill Lynch's biggest outsourcing initiative ever. With the new system, Merrill Lynch's financial advisors can obtain more of the assets of their high-net-worth customers. The new system also represents a major shift in the way Merrill approaches IT initiatives. In the 1990s, Merrill developed its previous system, Trusted Global Advisor (TGA), as it did any other major system - it developed it in-house. The thought of outsourcing a critical business system to a vendor was viewed as highly unfavorable by most financial services organizations. In the high-speed global business environment, an organization needs to maximize its profits, enlarge its market share, and restrain its ever-increasing costs. INSOURCING Insourcing is a common approach using the professional expertise within an organization to develop and maintain the organization's information technology systems. Michael Palmer, the COO (and former CIO) of Allied Office Products, decided to bring in-house nearly all of the IT functions that had been outsourced. Ever since Eastman Kodak announced that it was outsourcing its information' systems function in 1988 to IBM, DEC, and Businessland, large organizations have found it acceptable to transfer their IT assets, leases, and staff to outsourcers. Organizations should consider outsourcing in order to achieve the following benefits: - Financial savings - Increase technical abilities - Market agility Financial Savings Cost competitiveness has driven the Big Three - Ford, DaimlerChrysler, and General Motors - to collaborate in a unique way by building an IT service. For a business whose IT department is a non-core function, maintaining a "best-of-breed" status under these conditions is next to impossible, especially for small and medium-sized enterprises where cost is a critical factor. Nike's footwear business generates \$6 billion in revenues per year. It requires an achievement in global communication to keep its global Web of Nike designers, product developers, marketing teams, sales staff, distributors, and dealers in line with product developments and marketing plans. Therefore, Kodak moved its focus from being in the film business to being in the picture-processing business, and from this developed Kodak's exciting e-service, Photonet, a password-protected application that allows users to post photographs on the Internet, order copies, send picture postcards, and customize postcard greetings. IT outsourcing is a fast-growing industry since it provides access to state-of-the-art technologies with expert guidance, thus curtailing the need to open up expensive in-house departments. Business process outsourcing is increasingly becoming the strategic choice of companies looking to achieve cost reductions while improving their service quality, increasing shareholder value, and focusing on their core business capabilities. Accenture's first BPO deal came in the early 1990s, when it outsourced finance and accounting functions for British Petroleum.

### **Samenvatting per paragraaf**

#### ***Paragraaf 1***

With the new system, Merrill Lynch's financial advisors can obtain more of the assets of their high-net-worth customers. In the 1990s, Merrill developed its previous system, Trusted Global Advisor (TGA), as it did any other major system - it developed it in-house.

#### ***Paragraaf 2***

Palmer calculated that he could cut the \$24,000-a-month hardware monitoring and maintenance costs by two-thirds and save several hundred ff on development by insourcing.

**Paragraaf 3**

In some cases, the entire information technology department is outsourced, including planning and business analysis as well as the installation, management, and servicing of the network and workstations.

**Paragraaf 4**

**Increase Technical Abilities** Rapid technological change has become an issue for many businesses throughout all industries, resulting in more expense to upgrade systems, more time to install, and increased complexity to master. For a business whose IT department is a non-core function, maintaining a "best-of-breed" status under these conditions is next to impossible, especially for small and medium-sized enterprises where cost is a critical factor. Therefore, Kodak moved its focus from being in the film business to being in the picture-processing business, and from this developed Kodak's exciting e-service, Photonet, a password-protected application that allows users to post photographs on the Internet, order copies, send picture postcards, and customize postcard greetings. IT outsourcing is a fast-growing industry since it provides access to state-of-the-art technologies with expert guidance, thus curtailing the need to open up expensive in-house departments.

**Paragraaf 5**

- Core competencies- Many companies have recently begun to consider outsourcing as a means to fuel revenue growth rather than just a cost-cutting measure. - The Internet- The pervasive nature of the Internet as an effective sales channel has allowed clients to become more comfortable with the outsourcing process.

**Paragraaf 6**

Business process outsourcing is increasingly becoming the strategic choice of companies looking to achieve cost reductions while improving their service quality, increasing shareholder value, and focusing on their core business capabilities. BPO is divided into two categories: (1) back-office outsourcing that includes internal business functions such as billing or purchasing, and (2) front-office outsourcing which includes customer-related services such as marketing or technical support. However, the market is heating up these days, as companies' need for strategic cost cutting, desire to improve business methods, and comfort with outsourcing arrangements grow.

**Paragraaf 7**

Numerous countries have substantially well-trained IT professionals and clerical staff who have lower salary expectations compared to their U.S. Offshore outsourcing has become a small but rapidly growing sector in the overall outsourcing market.

## **Biilage IX      Samenvattingen Tekst 2: 'Outsourcing Information Systems Functions'**

### **Samenvatting van de volledige tekst**

In the IT world, outsourcing means turning over a firm's computer operations, network operations, or other IT function to a provider for a specified time; generally, at least a few years. Today, CIOs are expected to investigate outsourcing sufficiently to satisfy executive management that their IS operations are as efficient and effective in-house as they would be if they were outsourced; otherwise they should outsource what they do not do well. The Driving Forces Behind Outsourcing At a meeting of the Chicago Chapter of the Society for Information Management, Mel Bergstein of DiamondCluster Int'l. gave an overview of outsourcing. His main message was that outsourcing is another step in the evolution of the IT field. High-yield bonds allowed a few people to buy a company and leverage it with debt. These two drivers – focus and value – are still leading companies to restructure and focus on core businesses by asking themselves, "Where do we really add value?" They manage and finance projects, and outsource everything else. Thus, outsourcing is part of the drive for focus and value, and it is not solely an IT issue, says Bergstein; it is a business issue. They also bought products, which may or may not have included training. They also bought transactions, such as payroll check processing from a service bureau or credit reports from a credit rating service. Purchasing transactions allows buyers to shift fixed costs to variable costs and gives sellers higher margins because they take on the risks. They generally handle the entire life cycle – planning, development, maintenance, and training – for major systems projects. Finally, the most bundled approach to contracting for IT services is outsourcing, where the outsourcer contracts to handle all or most of certain IT activities. Outsourcing's History In 1989, essentially only one kind of outsourcing involving IT was available. Companies wanted to remove the huge IT infrastructure investments from their books and shift those fixed costs to variable costs; and they wanted to save money, generally about 15 percent. Former employees might have kept their same desk, but once they became an employee of the outsourcer, they became a provider and were treated differently. Formerly, they had thought that negotiating the deal was the difficult part, so they had not carefully defined governance structures, that is, how the relationship would be managed. They are much more careful in transition planning. Those early contracts have been renegotiated, and although the clients may not have changed providers, they have generally become more adept at renegotiating because they now know what they really need. They chose one of two routes. Either they outsourced maintenance of their legacy systems so their staff could concentrate on building new client-server systems or they outsourced client-server development to specialists and kept maintenance in-house. In either case, once the new systems were brought in, the legacy systems they replaced were shut down. Then, in the late 1990s, when the immense size of Y2K compliance surfaced – to retrofit old applications so they would work after the Year 2000 – most companies outsourced as much of their Y2K work as they could. As the IT outsourcing field matured, data center outsourcing, desktop outsourcing, and other standard IT outsourcing areas became so well understood that they became like commodity services, hence profit margins dropped as the number of competitors rose. To move into higher-margin services, ESPs began specializing in specific functional areas, offering to handle specific business processes as well as their IT underpinnings. This business process outsourcing (BPO) is defined as outsourcing all or most of a reengineered process that has a large IT component. Companies are outsourcing logistics, customer service, and many other essential, yet peripheral, functions to the experts. Each week, travel agencies must report the tickets they have sold to the Airline Reporting Corporation. Balboa is also able to do data mining on its ticket sales to uncover trends that will help it offer new and better travel services. As is obvious, BPO moves IT-based outsourcing out beyond the IS organization; it involves business units as well. BPO outsourcing is often quite a bit more

complex than IT outsourcing because it requires clients to change their business processes to fit with the processes of the service provider. Whereas IT outsourcing moved suppliers and customers closer to one another in terms of working together, the two parties still do not have the same goals. Clients want to save more money, whereas outsourcers want to make more money. In BPO, though, when providers take over an entire process, they can be measured and paid based on outcomes rather than transactions. The client focuses on "what" needs to be done; the provider focuses on "how" to do it. Generally, IT has not been seen as a core differentiating area when tailorable off-the-shelf products and services have been available. It can draw on best-of-breed as well as change course quickly if need be, swapping out one ESP and swapping in another, to keep pace with the market. Unlike traditional IT outsourcing, with e-business outsourcing, machines do not need to be purchased from the client, personnel do not have to be moved, and software licenses do not have to be transferred. E-business outsourcing and IT outsourcing are combining into a new form of managed services outsourcing that is being referred to by various names: utility computing, on-demand computing, virtual data centers, autonomic (self-healing) systems, and grid computing (a take-off on the term electricity grid). Clients can then scale their IT infrastructure or handle spikes in computing demand without having to buy more IT assets. To educate themselves and their business on the utility computing concept, CIOs can test it by using an IT outsourcer in a high-need area, such as e-mail, security, or redundancy, notes Thickins. It is important to understand the amount of consulting this form of outsourcing requires, the chargeback mechanism (for only paying for use), the contract terms, where it could provide benefits, and so on. Mobil Travel Guide illustrates the convergence of e-business and IT outsourcing into utility computing in one of its high-need areas: its Web site. Thus, in 15 years' time, IT outsourcing has expanded significantly, from outsourcing data center operations to outsourcing business processes and from domestic outsourcing to offshoring, which is discussed shortly. So they "insourced" to themselves, creating a shared services organization to handle such functions as IT, legal, facilities management, real estate, mail room, finance, and on and on. Generally, companies created a center of expertise in each area, with all the centers reporting to one shared services vice president. IT was not always included, but, as in the case of MeadWestvaco Corporation (see Chapter 1), it was. Some executives believe having IT in shared services gives them the ability to leverage the IT underpinnings of the other services. Offshoring To round out our discussion of outsourcing, we turn to the topic receiving much attention today: sending work offshore. In the late 1990s, when labor markets were especially tight and IS organizations needed to retrofit their systems to make them Y2K compliant, the use of offshore outsourcers to maintain IT applications grew dramatically. companies, near-shore means outsourcing to Canadian and Mexican companies, whereas offshore means Ireland, India, the Philippines, and other countries. During the recent economic downturn, offshoring gave companies a way to cut costs. In the early 1950s, each service sector employee produced about \$39,000 in output (in 2000 dollars), he notes; in manufacturing, the output was \$48,000. It should come as no surprise that service companies are tapping cheaper sources of labor now that they can because of globally available telecommunications technology, he notes. This global competition in services will force American companies to increase the productivity of their workforce in the same way and increase the quality of their services, Altman believes. In two reports, the Sourcing Interests Group (SIG) explores IT offshoring and the areas CIOs and their staff need to consider. India has become the premier IT and BPO offshoring country because its huge, highly educated workforce speaks English and is hardworking. Furthermore, in IS, all of the major IT outsourcers in India have achieved the highest level (Level 5) in the Capability Maturity Matrix, which means their processes produce high-quality software. Client companies not wanting to deal directly with unfamiliar overseas providers can now tap offshore resources through onshore outsourcers. Here is a brief description of the program at Syntel, a large U.S.-based IT provider that uses an offshore (India) model, as recounted in one of the SIG reports: We initially hold a 1.5-day workshop, and then a



follow-up orientation a few months later. On the provider side, offshorers often put their own employees through an accent neutralization program if they have an accent the client country will find difficult to understand. Providers also teach their employees about the client's company culture (going so far as to hang the same posters on the walls as the client has on its walls) and the client country's culture (going so far as to familiarize employees with holidays and keep them up-to-date on sports teams and major sporting events). These differences show up immediately when people from two cultures try to negotiate an offshore contract. The word "yes" in Asia typically means, "I hear what you are saying," whereas in the United States it means "I can do what you ask" or "I agree with you." When conversing, an Indian might nod his head "yes," which an American might misinterpret as an agreement.

- Have the offshore provider write a "statement of work," to gauge their understanding of what they think they are being asked to do.
- Include a person on the client negotiating team who knows the offshore country culture so that he or she can explain the country's norms to the client team as issues arise and prevent the offshorer's team from holding side discussions in their own language at the bargaining table.

As an example from the SIG report of a company using offshoring for its IT maintenance work, consider Exult, which is itself an example of a BPO company. Kate Kaiser of Marquette University and Stephen Hawk of the University of Wisconsin-Parkside<sup>8</sup> describe an 8-year arrangement between an unnamed financial services firm in the United States and an unnamed IT outsourcer in India. Both mechanisms increase the cost of outsourcing, note Kaiser and Hawk, because, for one thing, dual leadership requires more Indian staff to be located in the United States. believes that outsourcing of services is inevitable and that the real concern of service firms should be their loss of competitiveness, not the loss of jobs in their own country. He believes that the service economy is in the midst of restructuring itself, which is terribly painful when it happens in any industry. Companies that understand niches of customers, and serve them well, will themselves do well, believes Karmarkar, especially as they move their services to the Web. Many have done it and found that they lose touch with their customers. Service companies have a shot at dominating the screen, because the design of the service and the interface (rather than the technology or the appliance) will determine success, states Karmarkar. Due to this skill, NIT has claimed a disproportionate degree of control over the information chain, notes Karmarkar; it dominates the screen. In short, service companies need to understand what kinds of information various clusters of customer want and cater to them. For many service firms, that catering will be via the Web and mobile devices, using offshoring, self-service, and automation where each can contribute the most value. They hold a formal meeting, say, once a week to once a month; but they are generally in daily contact. Some companies have ongoing standing committees, such as a pricing committee or a change management committee, to oversee the use of formal change management procedures. This job position has not been prevalent in IS departments, but we believe it is going to become the norm as companies move toward IS Lite. A major governance item in the contract is the service level agreements (SLAs) because they are used to gauge supplier performance. In BPO, the situation is even worse; companies do not know how many people are in the process, departmental costs do not reflect overhead or IT costs, and so on. Clients who do not know their own performance levels negotiate from weakness; they know less than the vendor because they have not kept track of details, and vendors are not apt to correct mistaken impressions. Furthermore, they are likely to overlook important details, which will later cost them money. It is only when trust in one another breaks down that they turn to the contract. The outsourcer's staff is no longer under the purview of the client, so command-and-control is not a wise option – it only results in an acrimonious relationship. Users may think twice before requesting changes and prepare better definitions of what they want. Service providers do introduce new discipline; the client should prepare employees for this change and assist them in adapting because it is generally best to use the provider's processes. They were the most important items he had to handle. They kept him on focus and turned out to be his best informal management technique.

It means assisting one's suppliers to improve their products and services, generally by improving their processes. Although supplier development has not been prevalent in IT sourcing, we think it will be. Here is an example of supplier development from manufacturing, not IT, from a Sourcing Interests Group research report. In fact, taken to its extreme, it can mean devoting resources to assist suppliers improve their processes. We believe it is a coming focus in the world of IS Lite.

### **Samenvatting per paragraaf**

#### ***Paragraaf 1***

In the IT world, outsourcing means turning over a firm's computer operations, network operations, or other IT function to a provider for a specified time; generally, at least a few years. Today, CIOs are expected to investigate outsourcing sufficiently to satisfy executive management that their IS operations are as efficient and effective in-house as they would be if they were outsourced; otherwise they should outsource what they do not do well. The Driving Forces Behind Outsourcing At a meeting of the Chicago Chapter of the Society for Information Management, Mel Bergstein of DiamondCluster Int'l. gave an overview of outsourcing. High-yield bonds allowed a few people to buy a company and leverage it with debt. These two drivers – focus and value – are still leading companies to restructure and focus on core businesses by asking themselves, "Where do we really add value?" They manage and finance projects, and outsource everything else. They also bought products, which may or may not have included training. They also bought transactions, such as payroll check processing from a service bureau or credit reports from a credit rating service. Purchasing transactions allows buyers to shift fixed costs to variable costs and gives sellers higher margins because they take on the risks. They generally handle the entire life cycle – planning, development, maintenance, and training – for major systems projects. Outsourcing's History In 1989, essentially only one kind of outsourcing involving IT was available. Companies wanted to remove the huge IT infrastructure investments from their books and shift those fixed costs to variable costs; and they wanted to save money, generally about 15 percent. Former employees might have kept their same desk, but once they became an employee of the outsourcer, they became a provider and were treated differently. Formerly, they had thought that negotiating the deal was the difficult part, so they had not carefully defined governance structures, that is, how the relationship would be managed.

#### ***Paragraaf 2***

Either they outsourced maintenance of their legacy systems so their staff could concentrate on building new client-server systems or they outsourced client-server development to specialists and kept maintenance in-house. To move into higher-margin services, ESPs began specializing in specific functional areas, offering to handle specific business processes as well as their IT underpinnings. This business process outsourcing (BPO) is defined as outsourcing all or most of a reengineered process that has a large IT component. Balboa is also able to do data mining on its ticket sales to uncover trends that will help it offer new and better travel services. As is obvious, BPO moves IT-based outsourcing out beyond the IS organization; it involves business units as well. BPO outsourcing is often quite a bit more complex than IT outsourcing because it requires clients to change their business processes to fit with the processes of the service provider. Whereas IT outsourcing moved suppliers and customers closer to one another in terms of working together, the two parties still do not have the same goals. The client focuses on "what" needs to be done; the provider focuses on "how" to do it. With the arrival of business use of the Internet, outsourcing has enabled companies to quickly get Web sites up and handling business. Generally, IT has not been seen as a core differentiating area when tailorable off-the-shelf products and services have been available. It can draw on best-of-breed as well as change course quickly if

need be, swapping out one ESP and swapping in another, to keep pace with the market. Unlike traditional IT outsourcing, with e-business outsourcing, machines do not need to be purchased from the client, personnel do not have to be moved, and software licenses do not have to be transferred. E-business outsourcing and IT outsourcing are combining into a new form of managed services outsourcing that is being referred to by various names: utility computing, on-demand computing, virtual data centers, autonomic (self-healing) systems, and grid computing (a take-off on the term electricity grid). Clients can then scale their IT infrastructure or handle spikes in computing demand without having to buy more IT assets. To educate themselves and their business on the utility computing concept, CIOs can test it by using an IT outsourcer in a high-need area, such as e-mail, security, or redundancy, notes Thickins. It is important to understand the amount of consulting this form of outsourcing requires, the chargeback mechanism (for only paying for use), the contract terms, where it could provide benefits, and so on. Mobil Travel Guide illustrates the convergence of e-business and IT outsourcing into utility computing in one of its high-need areas: its Web site. Thus, in 15 years' time, IT outsourcing has expanded significantly, from outsourcing data center operations to outsourcing business processes and from domestic outsourcing to offshoring, which is discussed shortly.

### **Paragraaf 3**

So they "insourced" to themselves, creating a shared services organization to handle such functions as IT, legal, facilities management, real estate, mail room, finance, and on and on. Generally, companies created a center of expertise in each area, with all the centers reporting to one shared services vice president. IT was not always included, but, as in the case of MeadWestvaco Corporation (see Chapter 1), it was. Some executives believe having IT in shared services gives them the ability to leverage the IT underpinnings of the other services. Offshoring To round out our discussion of outsourcing, we turn to the topic receiving much attention today: sending work offshore. In the late 1990s, when labor markets were especially tight and IS organizations needed to retrofit their systems to make them Y2K compliant, the use of offshore outsourcers to maintain IT applications grew dramatically. companies, near-shore means outsourcing to Canadian and Mexican companies, whereas offshore means Ireland, India, the Philippines, and other countries. During the recent economic downturn, offshoring gave companies a way to cut costs. economy is improving, offshoring has become a political issue, because companies are not expanding their domestic labor force as rapidly as some had expected. In the early 1950s, each service sector employee produced about \$39,000 in output (in 2000 dollars), he notes; in manufacturing, the output was \$48,000. It should come as no surprise that service companies are tapping cheaper sources of labor now that they can because of globally available telecommunications technology, he notes. This global competition in services will force American companies to increase the productivity of their workforce in the same way and increase the quality of their services, Altman believes. In two reports, the Sourcing Interests Group (SIG) explores IT offshoring and the areas CIOs and their staff need to consider. India has become the premier IT and BPO offshoring country because its huge, highly educated workforce speaks English and is hardworking. College graduates in India apply for jobs that only high school graduates will apply for in the United States, such as working in a customer service center. Furthermore, in IS, all of the major IT outsourcers in India have achieved the highest level (Level 5) in the Capability Maturity Matrix, which means their processes produce high-quality software. Most IS organizations in the United States and in other developed countries are not even up to Level 3. Client companies not wanting to deal directly with unfamiliar overseas providers can now tap offshore resources through onshore outsourcers. Outsourcers and offshore advisory firms (who advise clients on the offshore marketplace) now realize that both parties need cultural training to overcome management and communication gaps. Here is a brief description of the program at Syntel, a large U.S.-based IT provider that uses an offshore (India) model, as recounted in

one of the SIG reports: We initially hold a 1.5-day workshop, and then a follow-up orientation a few months later. In fact, we tailor each workshop to the client's location because Midwest employees differ from New Yorkers, for example. We show a video of these differences, we have people wear the native dress, and we demonstrate how Indians prepare a lunch, which we then serve. On the provider side, offshorers often put their own employees through an accent neutralization program if they have an accent the client country will find difficult to understand. Providers also teach their employees about the client's company culture (going so far as to hang the same posters on the walls as the client has on its walls) and the client country's culture (going so far as to familiarize employees with holidays and keep them up-to-date on sports teams and major sporting events). These differences show up immediately when people from two cultures try to negotiate an offshore contract. The word "yes" in Asia typically means, "I hear what you are saying," whereas in the United States it means "I can do what you ask" or "I agree with you." When conversing, an Indian might nod his head "yes," which an American might misinterpret as an agreement. - Include a person on the client negotiating team who knows the offshore country culture so that he or she can explain the country's norms to the client team as issues arise and prevent the offshorer's team from holding side discussions in their own language at the bargaining table. clients are signed in the United States, and are governed by US. The list of issues in offshoring can be truly mind-numbing. As an example from the SIG report of a company using offshoring for its IT maintenance work, consider Exult, which is itself an example of a BPO company. Kate Kaiser of Marquette University and Stephen Hawk of the University of Wisconsin-Parkside<sup>8</sup> describe an 8-year arrangement between an unnamed financial services firm in the United States and an unnamed IT outsourcer in India. firm's IS staff gain competencies rather than lose them: formalized knowledge transfer between the two firms and a dual project-management hierarchy. Both mechanisms increase the cost of outsourcing, note Kaiser and Hawk, because, for one thing, dual leadership requires more Indian staff to be located in the United States. believes that outsourcing of services is inevitable and that the real concern of service firms should be their loss of competitiveness, not the loss of jobs in their own country. He believes that the service economy is in the midst of restructuring itself, which is terribly painful when it happens in any industry. He suggests looking in five places for determining new strategies for surviving and using offshoring, automation, and self-service to execute these strategies. Companies that understand niches of customers, and serve them well, will themselves do well, believes Karmarkar, especially as they move their services to the Web. Edmunds understands its customers and caters to them, states Karmarkar. Look for underserved niches, like Wells Fargo Bank has done in offering specific services to the Hispanic population in the western United States. Service companies have a shot at dominating the screen, because the design of the service and the interface (rather than the technology or the appliance) will determine success, states Karmarkar. In short, service companies need to understand what kinds of information various clusters of customer want and cater to them. For many service firms, that catering will be via the Web and mobile devices, using offshoring, self-service, and automation where each can contribute the most value.

#### **Paragraaf 4**

Here are just four – organizational structure, governance, day-to-day working, and supplier development. Managing outsourcing is different from managing internal staff because, for one thing, it is a joint effort between parties that may not have the same goals, as noted earlier. Therefore, during contract negotiations, the two parties need to figure out and negotiate how they are going to jointly manage the contract they sign. In fact, governance needs to be explicitly addressed in the contract. They hold a formal meeting, say, once a week to once a month; but they are generally in daily contact. A major governance item in the contract is the service level agreements (SLAs) because they are used to gauge supplier performance. In BPO, the situation is even

worse; companies do not know how many people are in the process, departmental costs do not reflect overhead or IT costs, and so on. Clients who do not know their own performance levels negotiate from weakness; they know less than the vendor because they have not kept track of details, and vendors are not apt to correct mistaken impressions. Furthermore, they are likely to overlook important details, which will later cost them money. In addition to SLAs, parties establish simple governance rules to be used when either party is making a decision, so that both are "singing from the same hymnal." Most parties in strong relationships say they put the contract in the drawer after it has been signed and work from trust and agreed-upon rules. It is only when trust in one another breaks down that they turn to the contract. The outsourcer's staff is no longer under the purview of the client, so command-and-control is not a wise option – it only results in an acrimonious relationship. Once Joe works for the supplier, he may no longer be able to provide that service; he must follow the work authorization process defined in the contract. Service providers do introduce new discipline; the client should prepare employees for this change and assist them in adapting because it is generally best to use the provider's processes. They were the most important items he had to handle. They kept him on focus and turned out to be his best informal management technique. It means assisting one's suppliers to improve their products and services, generally by improving their processes. Although supplier development has not been prevalent in IT sourcing, we think it will be.

***Paragraaf 5***

In fact, taken to its extreme, it can mean devoting resources to assist suppliers improve their processes.

## **Biilage X      Samenvattingen tekst 3: 'Outsourcing as an Economic Strategy'**

### **Samenvatting van de volledige tekst**

If organizations can effectively define and measure information technology costs and benefits, as we discussed in the preceding sections, they are in a better position to effectively manage their IT. On the other hand, they still may not be able to manage it as well as firms that specialize in managing IT. For such organizations, the most effective strategy for obtaining the economic benefits of IT and controlling its costs may be outsourcing, which is obtaining services from vendors rather than from within the organization. However, IT is not the primary business of many organizations. It is difficult to manage IT, even for organizations with above-average management skills. Organizations have always faced the choice of buying goods and services externally or producing them internally. Most organizations buy their power from electric utilities because this is cheaper than operating their own generating equipment. Outsourcing IT functions, such as payroll services, has been around since the early days of data processing. What is new is that, since the late 1980s, as illustrated in the opening case, many organizations are outsourcing the majority of their IT functions rather than just incidental parts. The trend became very visible in 1989 when Eastman Kodak announced it was transferring its data centers to IBM under a 10-year, \$500 million contract. This example, at a prominent multibillion-dollar company, gave a clear signal that outsourcing was a legitimate approach to managing IT. Organizations should consider the following strategies in managing the risks associated with outsourcing contracts (Clemons, 2000):

- Understand the project. The implementation of Internet-based technologies and especially of e-commerce encourages the use of outsourcing, as can be seen in the following IT at Work example below. (Some companies report 1,000 percent growth for EC activities in a year; e.g., see hotmail.com.) General Electric Information Services (geis.com), an extranet outsourcer, charges between \$100,000 and \$150,000 to set up an extranet, plus a \$5,000/month service fee. This is a strategy used by many corporations who prefer not to outsource the majority of their IT (like Kodak), but rather to outsource certain areas (such as connectivity or network security).

### **Samenvatting per paragraaf**

#### ***Paragraaf 1***

If organizations can effectively define and measure information technology costs and benefits, as we discussed in the preceding sections, they are in a better position to effectively manage their IT. On the other hand, they still may not be able to manage it as well as firms that specialize in managing IT. For such organizations, the most effective strategy for obtaining the economic benefits of IT and controlling its costs may be outsourcing, which is obtaining services from vendors rather than from within the organization. However, IT is not the primary business of many organizations. It is difficult to manage IT, even for organizations with above-average management skills. Organizations have always faced the choice of buying goods and services externally or producing them internally. Outsourcing IT functions, such as payroll services, has been around since the early days of data processing.

#### ***Paragraaf 2***

For example, Flooz.com, an online gift-currency store, outsourced its storage requirements to StorageNetworks, a storage service provider (Wilkinson, 2000).

**Paragraaf 3**

Organizations should consider the following strategies in managing the risks associated with outsourcing contracts (Clemons, 2000): - Understand the project. Dividing a large project into smaller and more manageable pieces will greatly reduce programmatic risk and provides clients with an exit strategy if any part of the project fails. (Some companies report 1,000 percent growth for EC activities in a year; e.g., see hotmail.com.) General Electric Information Services (geis.com), an extranet outsourcer, charges between \$100,000 and \$150,000 to set up an extranet, plus a \$5,000/month service fee.

**Paragraaf 4**

Some organizations may decide to outsource because they just do not have the skills to manage the IT function, or because they need to sell off IT assets to generate funds. Vendors may subcontract some of the services to other vendors.

## **Biilage XI      Samenvattingen tekst 4: 'Stategic dimensions of IT outsourcing'**

### **Samenvatting van de volledige tekst**

From an initial focus on cost reduction, outsourcing is fast becoming a complementary, routine mode of managing IT. On our own estimates, 30-35 per cent of most large organizations' IT budgets will be managed through outsourcing arrangements by 2002. The outsourcing report card: practices and risks It is clear that some organizations seek to address this issue through total outsourcing. But upon investigation, a much richer picture emerges of the IT sourcing paths organizations are taking (see Figure 1), and many still seek tactical IT gains rather than strategic business advantage. A mixed portfolio, "best-source" approach typically puts 15-30 per cent of the IT budget under third-party management, with other IT needs met through buying in resources that are under in-house management ("insourcing"), and through internal IT staffing. These organizations perceive IT as a core strategic asset; they believe that their IT employees will strive to achieve business advantage in ways that external providers cannot match, as well as being more loyal to the business. Total outsourcing (in which 80 per cent or more of the IT budget is under third-party management) is a minority pursuit. From tactical to strategic uses of outsourcing Against this background, most organizations that have outsourced IT during the present decade have pursued one of two approaches. In incremental/tactical outsourcing the primary focus has been on cost reduction, efficiencies and improved service in the way IT is run. These problems are: failure on the part of the supplier to understand the business (37 per cent); misalignment between corporate strategy and IT (35 per cent); and poor strategic IT planning (24 per cent). Other organizations, through some combination of competitive, financial, business, technical and IT skills pressures, have taken a "hard learning" route; that is, they have outsourced to a significant degree without really developing a strategic focus on their outsourcing objectives, let alone the capabilities to achieve them. However, organizations such as British Aerospace, Continental Bank and Sears have sought long-term, significant changes in their financial position through outsourcing. By the beginning of 1999 much more was being done with the technology, and IT efficiencies were being achieved in operational areas, though the IT budget as a whole had nearly doubled from 1993. These were chosen to support a strategy of becoming a more diversified production company, moving into areas in which BPX often did not have the necessary in-house IT skills. However, some organizations have used outsourcing more strategically, seeing it as a significant investment in transforming a traditional IT function, and achieving a new technology agenda to underpin business strategy. It had become clear to Xerox that its business restructuring in the early 1990s had left its existing systems and skills base unable to supply the required technological support. It developed a new strategy for which a new infrastructure was required, whose hardware alone would cost \$55m. Outsourcing allowed Xerox to refocus IT on business-critical applications, while the vendor facilitated routine operations, applications and their support, telecommunications and the move to a client/server infrastructure. DuPont looked to convert fixed IT costs into variable costs, and also hoped for improved service speed and flexibility, skills renewal, IT career development and further cost reductions. DuPont retained some 60 staff to oversee the contracts, while over 1,000 distributed business and technical people provide business IT leadership, and process control and research and development computing. For example, fiberglass manufacturer Owens Corning signed a five-year contract in 1995 for a vendor to maintain legacy systems while it implemented an enterprise resource planning system at 75 sites worldwide. However, some organizations operate even more strategically, incorporating transitional IT outsourcing into a suite of outsourcing contracts to help major organizational change. Business innovation: improving and innovating in processes, skills and technology, while mediating financial risk through the vendor, in order to achieve competitive advantage Organizations looking to



suppliers for IT or even business innovation through traditional outsourcing contracts have often been disappointed. Yet business innovation can be a potent source of competitive advantage that external IT partners can assist in. New market: direct profit generation through joint ventures with vendor partner Some organizations have experimented with creating a new "spin-off" supplier company from their IT function. Mellon Bank, Sears Roebuck and Boeing have had limited success with their spin-off IT companies. Even so, it took Origin several years to build a wider outsourcing market position. It requires a lot of management maturity and experience of IT outsourcing. It needs complete and creative contracting; a less long-term focus in the contracting arrangements, but a more long-term one in the relationship dimension, and very active post-contract management. Business responsibility, and the ability to identify business models and value-drivers that will make innovation possible and create business advantage. The ability to make sourcing decisions and arrive at a long-term IT sourcing strategy; this must build in learning and take into account business, technical and economic factors. Second, senior executives and IT managers who make decisions together achieve success more often than when either group acts alone. Second, organizations that invite both internal and external bids succeed more often than organizations that merely compare a few external bids with current IT performance. The ability to manage the contract in ways that secure long-term IT objectives, and achieve the required service performance from the supplier. IT outsourcing requires a great deal of in-house management - to determine and deliver business requirements, to ensure technical capability, and to manage external supply and IT governance.

#### **Samenvatting per paragraaf**

##### ***Paragraaf 1***

From an initial focus on cost reduction, outsourcing is fast becoming a complementary, routine mode of managing IT. On our own estimates, 30-35 per cent of most large organizations' IT budgets will be managed through outsourcing arrangements by 2002.

##### ***Paragraaf 2***

The dominant mode is selective sourcing, especially in the US (82 per cent of organizations) and the UK (75 per cent). A mixed portfolio, "best-source" approach typically puts 15-30 per cent of the IT budget under third-party management, with other IT needs met through buying in resources that are under in-house management ("insourcing"), and through internal IT staffing. These organizations perceive IT as a core strategic asset; they believe that their IT employees will strive to achieve business advantage in ways that external providers cannot match, as well as being more loyal to the business. Total outsourcing (in which 80 per cent or more of the IT budget is under third-party management) is a minority pursuit.

##### ***Paragraaf 3***

Other organizations, through some combination of competitive, financial, business, technical and IT skills pressures, have taken a "hard learning" route; that is, they have outsourced to a significant degree without really developing a strategic focus on their outsourcing objectives, let alone the capabilities to achieve them. However, in our most recent research into global IT outsourcing practices, we have discerned six strategic foci, in combination or in isolation, through which organizations are increasingly attempting (sometimes with success) to achieve a business advantage.

**Paragraaf 4**

For a company that is in financial trouble, outsourcing provides a significant cash influx from the transfer of IT assets and staff, together with tax advantages; this improves cash flow and makes it possible to restructure the balance sheet and profit and loss account. As such, outsourcing represents a long-term financial strategy, supporting a turnaround in the company's financial and competitive position over time in exchange for annual outsourcing fees representing an interest rate.

**Paragraaf 5**

To achieve this, the company needed only to own its route structure, brand and (IT-based) yield management system. These were chosen to support a strategy of becoming a more diversified production company, moving into areas in which BPX often did not have the necessary in-house IT skills. Subsequently, the in-house team focused on business processes, on information and business value-creation areas, and the vendors on more routine applications and technology.

**Paragraaf 6**

Outsourcing allowed Xerox to refocus IT on business-critical applications, while the vendor facilitated routine operations, applications and their support, telecommunications and the move to a client/server infrastructure. DuPont looked to convert fixed IT costs into variable costs, and also hoped for improved service speed and flexibility, skills renewal, IT career development and further cost reductions. DuPont retained some 60 staff to oversee the contracts, while over 1,000 distributed business and technical people provide business IT leadership, and process control and research and development computing.

**Paragraaf 7**

For example, fiberglass manufacturer Owens Corning signed a five-year contract in 1995 for a vendor to maintain legacy systems while it implemented an enterprise resource planning system at 75 sites worldwide. However, some organizations operate even more strategically, incorporating transitional IT outsourcing into a suite of outsourcing contracts to help major organizational change.

**Paragraaf 8**

BPX, for example, had to create an innovation fund for suppliers and BPX staff to bid for, such was the lack of innovation coming through. Yet business innovation can be a potent source of competitive advantage that external IT partners can assist in.

**Paragraaf 9**

The new supplier needs to build commercial and marketing skills for a competitive marketplace in which other vendors have better track records and will want to exclude the new entrant. Mellon Bank, Sears Roebuck and Boeing have had limited success with their spin-off IT companies. Electronics multinational Philips took shares in Dutch software house BSO, and transferred its development staff, and subsequently its processing operations, into a jointly owned company, Origin. Also, when they are under pressure in other parts of the outsourcing contract, client and supplier may become less interested when they realize that commercializing and marketing a new product may require an investment nine times greater than the original development cost. Supporting the supplier's market and profit growth can conflict with, and be at the expense of, service to the original client.

***Paragraaf 10***

It requires a lot of management maturity and experience of IT outsourcing. Business responsibility, and the ability to identify business models and value-drivers that will make innovation possible and create business advantage. The ability to make sourcing decisions and arrive at a long-term IT sourcing strategy; this must build in learning and take into account business, technical and economic factors. Second, senior executives and IT managers who make decisions together achieve success more often than when either group acts alone. Second, organizations that invite both internal and external bids succeed more often than organizations that merely compare a few external bids with current IT performance. IT outsourcing requires a great deal of in-house management - to determine and deliver business requirements, to ensure technical capability, and to manage external supply and IT governance.

***Paragraaf 11***

Examples include long-term financial restructuring, the ability to focus on core competences, the "catalyzing" effect of an external service provider, help during organizational transitions (such as mergers or restructuring), support for innovation, and joint ventures with a vendor partner.

## Auteursrechterlijke overeenkomst

*Opdat de Universiteit Hasselt uw eindverhandeling wereldwijd kan reproduceren, vertalen en distribueren is uw akkoord voor deze overeenkomst noodzakelijk. Gelieve de tijd te nemen om deze overeenkomst door te nemen, de gevraagde informatie in te vullen (en de overeenkomst te ondertekenen en af te geven).*

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:

### **Integratie van text mining-technieken in de ontwikkeling van een e-learningcursus**

Richting: **Handelsingenieur in de beleidsinformatica**

Jaar: **2007**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Ik ga akkoord,

**Sven EVENS**

Datum: **24.08.2007**