The development of an Implementation Methodology for a Conceptual Framework Tool used for the Improved Viewing and Utilisation of Organisational Information.

Dirk J. Kotze

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Industrial Engineering at Stellenbosch University.

Study Leader: **C.S.L. Schutte**

December 2008

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 27 November 2008

Copyright © 2008 Stellenbosch University
All rights reserved

Synopsis

This thesis reports on a research project conducted at Stellenbosch University, focusing on the field of information management, and contributing to the development of an approach for structuring information in such a manner that context becomes apparent, which on its part assists persons (and groups of persons) in selecting useful information from a larger whole, given a certain context. This document is composed of three general sections.

The first section is a literature study focusing on the following topics: (1) Information overload: The amount of information in the world, as well as its rate of generation is discussed, together with the notion of information overload, its causes, symptoms and possible solutions. (2) Information utilisation in organisations: General concepts concerning the utilisation of information in organisations are discussed, together with the manners in which organisations fail in efficiently utilising information, but also some examples where organisations have succeeded. (3) Information modelling: Various well-known information modelling approaches currently employed to improve information utilisation are discussed and compared. It is however mentioned that none of these approaches are sufficiently general, flexible and simple enough to assist typical organisational knowledge workers in efficiently interacting with information.

The second section of this thesis focuses on the conceptual framework information modelling approach currently being investigated at Stellenbosch University as a means for improving the information utilisation efficiency of organisational knowledge workers. The basic concepts of the approach are discussed, and it is compared with the more well-known information modelling approaches mentioned in the literature study. The research being conducted on the approach is further briefly discussed, which includes the purpose of this thesis (which is the development of a methodology for conceptual framework implementation in organisations), the research method followed, and the thesis hypothesis, which is: **Information represented by a conceptual framework implemented through the use of the developed conceptual framework implementation methodology, offer a truthful view on information found in an organisation.** The methodology developed is subsequently discussed in terms of its phases and activities, each of which is discussed in terms of its description, the rationale behind and prerequisites of its execution, details of the manner of its execution, and the eventual outputs thereof.

The final section of this thesis provides information on the validation of the conceptual framework implementation methodology described in the previous section, which were done in two parts. A theoretical validation was done through a comparison of the developed methodology with existing methodologies for the implementation of other information modelling approaches. A practical validation was also done through two case studies by which conceptual frameworks were implemented at two separate organisations. The document is finally concluded through a summary of the contents of this thesis, the outcome of the hypothesis test, and some closing remarks.

Opsomming

Hierdie tesis doen verslag oor ‘n navorsingsprojek, gefokus in die veld van inligtingsbestuur, wat by die Universiteit Stellenbosch uitgevoer is, en ‘n bydrae lewer tot die ontwikkeling van ‘n benadering waardeur inligting op só ‘n wyse gestruktureer kan word dat konteks ooglopend raak en waardeur persone (en groepe persone) gehelp kan word om bruikbare inligting te kan identifiseer vanuit ‘n groter geheel, gegewe ‘n spesifieke konteks. Hierdie dokument is saamgestel uit drie algemene afdelings.

Die eerste afdeling is ‘n literatuurstudie wat fokus op die volgende onderwerpe: (1) Inligtingsoorlading: Die hoeveelheid inligting in die wêreld en die tempo waarteen dit geskep word, word bespreek saam met die idee van inligtingsoorlading, oorsake en simptome daarvan, en moontlike oplossings daarvoor. (2) Inligtingsgebruik in organisasies: Algemene konsepte ten opsigte van die gebruik van inligting in organisasies word bespreek, tesame met die maniere waarin maatskappy daarin faal om inligting doeltreffend te gebruik, maar ook voorbeelde waar maatskappy suksesvol was. (3) Inligtingmodellering: Verskeie bekende inligtingmodellerings-benaderings wat tans aangewend word om inligtingsbenutting te verbeter, word bespreek en vergelyk. Dit word egter genoem dat geeneen van hierdie benaderings voldoende algemeen, aanpasbaar en eenvoudig is om tipiese kenniswerkers in organisasies by te staan in die doeltreffende interaksie met inligting nie.

Die tweede afdeling van hierdie tesis fokus op die konseptuele raamwerk inligtingmodelleringsbenadering wat tans by die Universiteit Stellenbosch ondersoek word om die doeltreffendheid van inligtingsbenutting deur kenniswerkers in organisasies te verbeter. Die basiese konsepte van die benadering word bespreek, en dit word vergelyk met die meer bekende inligtingmodellerings benaderings wat in die literatuurstudie genoem is. Die navorsing gedoen oor die benadering word kortliks verder bespreek, wat die doel van hierdie tesis insluit (die ontwikkeling van ‘n metodologie vir die implementering van konseptuele raamwerke in organisasies), die navorsingsmetode wat gevolg is, en die tesis hipotese, wat die volgende is: **Inligting voorgestel deur ‘n konseptuele raamwerk wat geïmplementeer is deur die ontwikkelde konseptuele raamwerk implementeringsmetodologie, lewer ‘n betroubare blik op die inligting wat in ‘n organisasie gevind word.** Die methodologie wat ontwikkel is, word gevolglik bespreek in terme van die fases en aktiwiteite daarvan. Elkeen word bespreek in terme van ‘n beskrywing, die rationale daaragter en voorvereistes vir die uitvoering daarvan, besonderhede rakende die uitvoer daarvan, en die uiteindelike uitsette.

Die finale afdeling verskaf inligting oor die validering van die konseptuele raamwerk implementerings-metodologie wat in die vorige afdeling bespreek is (en wat op twee maniere plaasgevind het). ‘n Teoretiese validering is gedoen deur ‘n vergelyking van die ontwikkelde metodologie met bestaande metodologieë vir die implementering van ander inligtingmodellerings benaderings. ‘n Praktiese validering is ook gedoen deur twee gevallestudies waardeur konseptuele raamwerke by twee verskillende organisasies geïmplementeer is. Die dokument word laastens afgesluit deur ‘n opsomming van die tesisinhoud, die uitslag van die hipotesetoets en ‘n aantal opmerkings ter afsluiting.

Acknowledgements

I would like to thank the following parties for the various contributions they have made towards the successful completion of this thesis, and the ultimate achievement of my M.Sc.Eng degree:

- My wife, Bernice, for her loving support and understanding;
- My study leader, Corne Schutte, for the excellent guidance he provided with regards to this thesis, as well as the Industrial Engineering Department at Stellenbosch University for general support regarding my M.Sc.Eng degree;
- My employer, Indutech (Pty) Ltd, in providing me with time and resources to complete my M.Sc.Eng degree;
- My family and friends for prayer and support, and especially the small “army” of proof readers who assisted me in delivering a thesis of excellent quality: De Villiers, Margarete, Bernice (wife), Abrie (father), Neil, Marijke, Marietjie, Ernst, Corne (study leader), Wilhelm, Heinz, Du Toit and Johan; and finally
- My Lord and Saviour, Jesus Christ, who afforded me the abilities and endurance to successfully complete this venture.

Table of Contents

List of Figures

List of Tables

Chapter 1 – Introduction

1.1 Background

From the end of the twentieth century, the focus of humanity's endeavours had shifted from large-scale, direct involvement in activities essential to its existence (such as agriculture, manufacturing and services) to interacting with the information surrounding these activities. This has been the consequence of humanity's understanding that an increase in the information concerning a particular situation (which subsequently may lead to an improved understanding, and ultimately improved knowledge) will be of increased benefit at the possible recurrence of such a situation, may be partly beneficial to other, similar situations, and may even have the potential to distinguish a person (or group of persons) within a community, or within the whole of humanity itself. Continuous advances in human knowledge and technological prowess did however have a significant effect on the magnitude of information generated, as well as on the rate of information generation, and have reached such proportions that the members of the human populace currently do not possess the capacity to both perform the essential activities needed for continued survival and assimilate all the information associated with it. Over all spectrums of human activity, extensive interaction with information have evolved into a separate essential activity of humanity, with persons being specifically tasked to interact with, process, and in some cases, assimilate and divulge information. Currently, the majority of (especially) the developed world's workforce is comprised of these "knowledge workers" (Bentley 1998), which is a concept that was already created in the late 1950s by Peter Drucker, and refers to workers who produce abstract work products consisting mainly of information, rather than tangible items like cars or pencils (Inmon, O'Neill & Fryman 2008).

The combination of powerful communication technologies, together with the empowering of large segments of humanity to contribute to information generation activities, is currently causing a veritable world-wide explosion of information. Never before in the history of mankind has information been so accessible and readily available as is currently the case. Each day vast amounts of information are created and channelled through a large number different media for eventual consumption by people, who utilise it for a myriad of different purposes. The sheer amount of information, together with the added complexities such as variation in applicability and quality, has however become too much even for specialised knowledge workers to assimilate. For this reason, researchers have started developing and implementing various different techniques and approaches to assist knowledge workers in distinguishing between useful and useless information in their respective contexts. Today almost no knowledge worker environment can be found that do not employ some kind of tool aimed at lessening the impact that working with large amounts of information may have on individuals. However, before continuing this discussion concerning the utilisation of various techniques and approaches for improving human interaction with information, the interaction of *organisations* with information will briefly be discussed.

Nowhere have the effect of the over-supply of information been experienced as intensely as within organisations. Organisations are groupings of people who gain their economic sustenance from working together as a unit (Bentley 1998), and can be seen as societies of minds, in which actions and decisions are not the simple outcome of any single orderly activity, but emerge from an ecology of information processes (Choo 2000). Its capacity for innovation and the manner in which its activities are performed play a major part in the success of an organisation (Bernard & Tichkiewitch 2008). On their part, these attributes are most often driven and supported by information with a very specific focus and level of quality. Having the "right information at the right time" can therefore enable an organisation to perform its own operations in an efficient manner, as well as to manoeuvre itself within its environment in such a way as to stay profitable and competitive.

In order to effectively interact with the information present in its internal and external environment, an organisation must have a good understanding of the different types of information in its possession, and the uses of these types of information. This understanding can be established through the use of metadata, which is a means of describing information within the organisation, and through which a person can rapidly reach a conclusion on whether specific information can sufficiently satisfy his/her need. Inmon, O'Neill and Fryman (2008) stated that organisational metadata can be broken down into two major categories, based on the audience that it serves, namely business metadata and technical metadata. Technical metadata is typically very structured, which makes it easy for seekers of information to locate and use. It does however have a narrow focus, with usually only persons involved in the technical activities of an organisation benefiting from it. Business metadata on the other hand describes the information most of the persons working for an organisation are involved with on a daily basis, but are much less structured, making the locating of information through the use of business metadata much more difficult.

To assist its knowledge workers in locating the information they require, most organisations have implemented tools and approaches aimed at improving the efficiency of their information searching activities, but these still struggle to live up to expectations, as they have varying degrees of success in providing organisational knowledge workers with the information they seek. These tools and approaches are usually quite adept at identifying business metadata within organisational information and matching this with an organisational knowledge worker's information request, but battle to take the context of the information into account as well, which is required to efficiently and successfully match information to the actual need of a knowledge worker. The result is therefore a subset of organisational information describing various contexts, which the knowledge worker still has to process and assess in order to obtain the information that is relevant to his/her context.

In order to become (and remain) agile and competitive in the world of business, organisations clearly need more effective solutions to the problem of locating and utilising information in an efficient manner. An approach which (it seems) may provide a solution to this problem, is the development of tools and approaches that not only identify the business metadata that exist within organisational information, but

also recognise and utilise the relations that exist between the different business (and where relevant, technical) metadata. By focusing in this way on the context of information as well, more of the information processing could be performed by a tool or approach, yielding a smaller, focused, and more relevant set of information in response to an information query by a knowledge worker. The result of employing such a tool or approach in an organisation will then obviously be a reduction in the time spent by knowledge workers to assess and utilise information, and ultimately enable the organisation to become more agile and competitive.

This thesis describes a research project that focused on an approach currently being investigated at Stellenbosch University for identifying and utilising context within information. The next section will communicate the rationale behind the project, with the final section in this chapter explaining the structure of the thesis document. An image depicting the structure of this chapter is displayed in Figure 1.1.

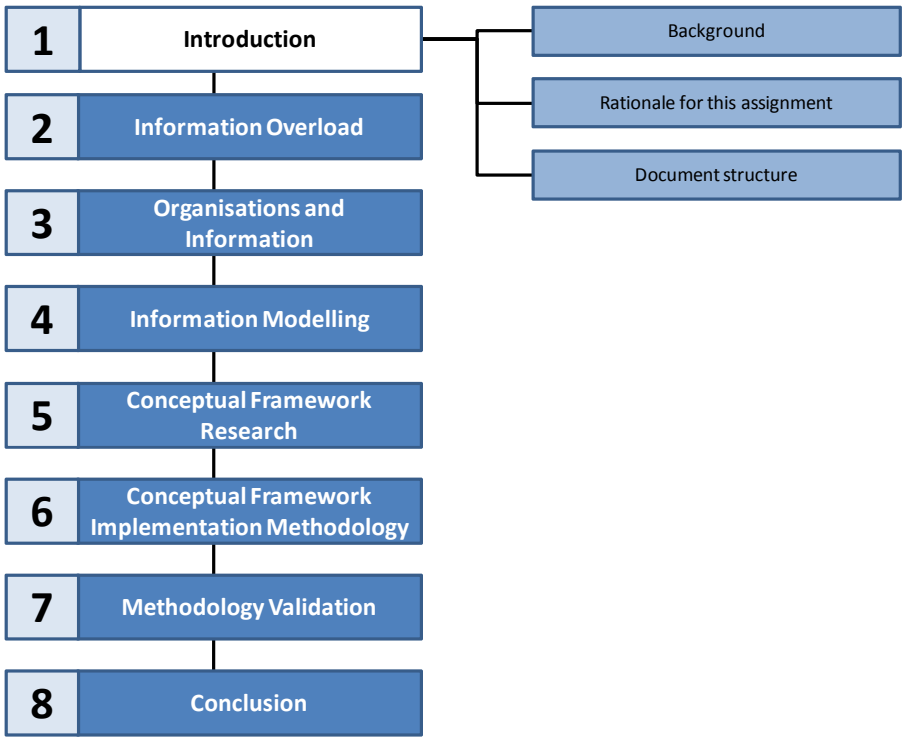


Figure 1.1: Image for navigating thesis and chapter 1

1.2 Rationale for this assignment

Research is currently being conducted on an approach developed at Stellenbosch University through which context in information can be represented in a manner that is easily understood by both humans and machines, and which is aimed at improving the manner in which information can be utilised within organisations. The approach follows the school of thought which states that information can be described

as a collection of concepts (or terms, entities, objects, etc.), and the relations between these terms (similar to ontologies, topic maps, concept maps, etc.). After some initial experimentation it was decided that a formal project (on which this thesis reports) should be conducted through which a methodology was developed for implementing this approach in organisations. Once developed, this methodology was then also to be utilised in the performing of a number of actual implementations in order to refine the methodology and to gain further insight in the manner that this approach may influence the abilities of an organisation to efficiently view and utilise information.

1.3 Document structure

This document consists of eight chapters (as shown in Figure 1.1), which will forthwith be discussed in brief to give an indication of the document structure.

This chapter (Chapter 1) gave a very broad overview of information, and how humanity's creation of larger amounts of information gave rise to, firstly the knowledge worker (which is a person whose sole purpose is to interact with information) and secondly, tools and approaches aimed at assisting knowledge workers in assessing and utilising still-increasing amounts of information. Organisations' dependence on information for survival was discussed next, and the fact communicated that organisational knowledge workers and implemented tools and approaches currently struggle to interact efficiently with the masses of information that an organisation typically possesses. The chapter concluded by indicating that research conducted at Stellenbosch University is focused on addressing this problem, and that this thesis reports on a project contributing to this research. The structure of this thesis is also briefly discussed.

Chapter 2 is the start of this thesis' literature study and begins by briefly explaining the significance of information to people, and the current conditions around information availability and the rate of information generation. It continues by introducing the concept of information overload and its effect on individuals and organisations. Following the introduction, information overload is discussed in more detail by focusing on the different causes and symptoms of the phenomenon, as well as on solutions proposed by authorities on the subject. The chapter concludes by alluding to the purpose of this thesis as a proposal for an approach aimed at solving or at least alleviating the problem of information overload in knowledge workers (especially in an organisational context).

The purpose of Chapter 3 is to provide some background on the manner in which information is utilised by organisations specifically. The introduction discusses the importance of information to organisations as a means to manoeuvre itself within its environment, and as an aid to achieve an advantage over its competitors. Three main themes are then subsequently discussed in the remainder of the chapter:

- Firstly, some general information on information utilisation within organisations;

- Secondly, the shortcomings of the manners in which organisations typically interact with and utilise their information; and
- Thirdly, a brief mention on the manner in which organisational information utilisation has been improved through the implementation of various tools and approaches.

The chapter concludes by stating that the variety of solutions for improving information utilisation may however indicate that people lack an understanding of the true nature of information, and that an improved understanding may be obtained by considering information as a collection of data points which are related to each other in various ways.

Chapter 4 is a continuation of the discussion on how information can be represented as a collection of data points that are connected by the relations that exist between them. The chapter discusses a number of different approaches which are based on this school of thought, and are currently being utilised for various purposes in the world-wide information environment. The approaches discussed are as follows, with each discussion focusing mainly on the concept that drives the approach and the manner in which the approach is typically applied:

- Entity-relationship models (as an introduction);
- Semantic networks;
- Thesauri;
- Taxonomies;
- Ontologies;
- Folksonomies;
- Concept maps; and
- Topic maps.

The chapter concludes by comparing the discussed approaches in terms of their generality and flexibility of application, as well as their complexity in terms of ease-of-use. It is indicated that an ideal approach would be flexible, and could be applied in a large variety of situations whilst still remaining simple for a user to understand. Unfortunately, none of the approaches discussed manages to satisfy more than two of these three requirements. This point finally serves as an introduction to the next chapter which introduces an approach currently being researched at Stellenbosch University, aimed at addressing this very need. Chapter 4 is also the final chapter of the thesis' literature study.

Chapter 5 introduces the concept of a *conceptual framework,* which is the name given to an information modelling approach on which research is currently being performed at Stellenbosch University. The first part of the chapter explains the approach by defining and explaining the components which constitute a

conceptual framework, namely entities, relations, types and facts. It continues by indicating the manner in which the conceptual framework approach differs from the approaches discussed in Chapter 4. This is followed by an explanation on the manner in which conceptual frameworks are applied, and the value of conceptual framework visualisation. This introduction into conceptual frameworks is concluded with information on the research that is being conducted on conceptual frameworks at Stellenbosch University, and the different aspects of this research. The second part of this chapter discusses the hypothesis of the project on which this thesis reports, which is the development of a methodology for the implementation of a conceptual framework in an organisation (focusing on structured information). Chapter 5 concludes by discussing the method that was followed during the research project.

Chapter 6 describes the methodology for the implementation of a conceptual framework (focused on structured information) that was developed through the research project discussed in Chapter 5, and outlines the activities that must be performed in order to successfully implement a conceptual framework in an organisation.

The methodology discussed in Chapter 6 was validated in two ways; firstly from a theoretical perspective, and secondly through practical experience. The first part of Chapter 7 discusses the theoretical validation of the conceptual framework implementation methodology which was a comparison with existing information modelling approach implementation methodologies that was done in order to ensure that the conceptual framework implementation methodology is conceptually sound. The second part of the chapter discusses two case studies that were performed in order to further refine and validate the methodology. The process that was followed to implement each of the conceptual frameworks are discussed, followed by some information on the conceptual framework that resulted from each implementation, and observations by the persons involved on each implemented conceptual framework itself, as well as on the process that was followed to establish the conceptual framework in each particular organisation.

The final chapter (Chapter 8) contains the conclusion of this thesis, and is followed by the bibliography (Chapter 9) and appendices (Chapter 10).

Chapter 2 – Information Overload

2.1 Introduction

From the beginning of time, information and knowledge has proven to be significant in the advancement of the human race. In this modern day and age, almost every aspect of "being human" can be traced back to information that was obtained at some stage in the past, processed, and integrated with a current knowledge base, which in turn influenced humanity's behaviour, beliefs and understanding of the world it inhabits. Even with information being extremely diverse in nature and application (compare science with religion), it is clear that the more information a person possesses, the more he/she will benefit from it, and be empowered to be distinguishable from those around him/her. From this point of view it would seem that there can be no downside to having and discovering information. But, how can it then be that something that is apparently such a benefit to the human race is increasingly being seen as a problem (Carlson 2003)?

The Information Revolution has provided humanity with the capability (both knowledge and technological) to communicate information on a global scale with very little effort. Because humans have always addressed the high cost of finding information by sharing it (Borchers et al. 1998), the act of information sharing has evolved from oral information to the written word, and some years ago, the Internet has provided the ideal spawning ground for the next step in the evolution of information sharing. Due to the technological advancement of the last 50 years, information has become more accessible to people than at any other time in human history (Edmunds & Morris 2000), and according to Bentley (1998), today more people in the Western world work with information than in manufacturing or agriculture.

Additional to information becoming much more accessible, software providers have also started developing tools and services aimed at enabling users to generate and share their own content on the Internet, nullifying this "privilege" that was previously only available to publishing houses and "accepted" authors. This means that these days, internet users can access immense amounts of information on the internet, from very exclusive, high-value and high-quality information sources like peer-reviewed papers, on-line journals and on-line magazine articles, to less exclusive websites with also less value and quality like corporate and interest group websites, and further to the much more social and low quality information sources like forums, blogs and wikis. Figure 2.1 shows this very clearly and further illustrates a tendency that a decrease in accessibility is experienced that results from an increase in control, aimed at ensuring a certain level of quality within information. The ease of information sharing has brought about an abundance of available information with varying levels of value, with the potential to satisfy (almost) any possible need for information that exists in the world today. The challenge in utilising this information

manifests itself in the ability of a person to extract the information he/she needs from the entire knowledge base within a reasonable time span.

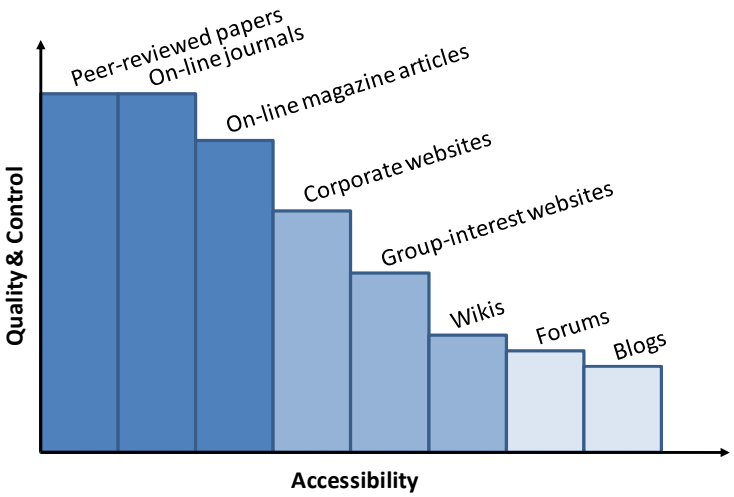


Figure 2.1: Comparison of the quality & control and accessibility of different internet information sources

This chapter (of which the structure is depicted in Figure 2.2) discusses the manner in which interaction with information takes place in the world today, and provides insight into the malady experienced by individuals and groups of individuals (like organisations) in attempting to assimilate vast quantities of information, called *information overload*. In order to impart a better understanding of this problem, causes, symptoms and proposed solutions to the information overload problem will specifically be discussed.

2.2 Available information and rate of information creation

The increase in the generation and availability of information has become ever more apparent over the last few decades as the following quotes clearly indicate:

- 'As we go from grade school to high school we learn only a billionth of what there is to learn. There is enough scientific information written every day to fill seven complete sets of Encyclopedia Britannica; there is enough scientific information written every year to keep a person busy reading day and night for 460 years!' (Siegel cited in Information Overload Statistics 2003)
- 'The daily New York Times now contains more information that the 17th century man or woman would have encountered in a lifetime.' (Wurman cited in Information Overload Statistics 2003)

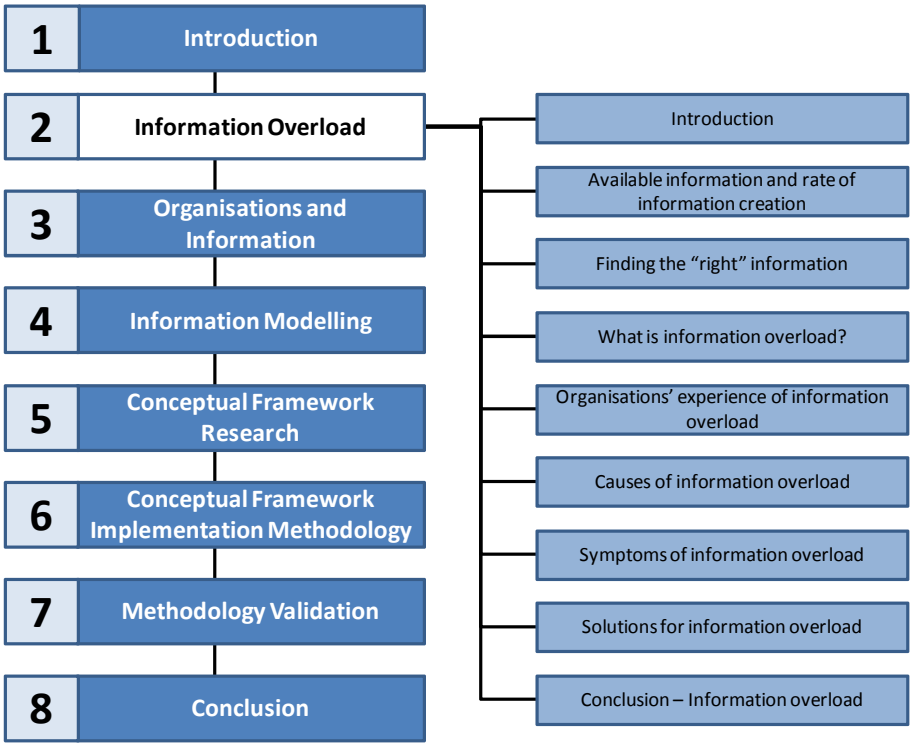


Figure 2.2: Image for navigating thesis and chapter 2

- 'About 1,000 books are published internationally every day, and the total of all printed knowledge doubles every five years.' (Information Overload Causes Stress cited in Information Overload Statistics 2003)
- 'The average Fortune 1000 worker already is sending and receiving approximately 178 messages and documents each day, according to a recent study, "Managing Corporate Communications in the Information Age."' (Boles cited in Information Overload Statistics 2003)
- 'Dr Dharma Singh Khalsa, in his book Brain Longevity,...says the average American sees 16,000 advertisements, logos, and labels in a day.' (Gore cited in Information Overload Statistics 2003)

The University of California Berkeley conducts a project named 'How Much Information', which studies the amount of information produced each year in the world (*How much Information?* 2003), and how that information is utilised by the world's population. Here are some of the findings from the study conducted in 2003:

- ***The (size of the) INTERNET***

  Although the Internet is the newest medium for information flows, it is the fastest growing new medium of all time, and becoming the information medium of first resort for its users. Note that the Web consists of the surface web (fixed web pages) and what Bright Planet calls the deep web (the

database driven websites that create web pages on demand). Table 2.1 shows the estimated size of the different aspects of the Internet:

Table 2.1: The size of the Internet in terabytes (Source: How much information? 2003)

Year: 2002	Terabytes[1]
Surface Web	167
Deep Web	91,850
Email (originals)	440,606
Instant messaging	274
TOTAL	**532,897**

- ## *Amount of information produced per year*

Print, film, magnetic, and optical storage media produced about **5 exabytes**[2] of new information in 2002. 92% of the new information was stored on magnetic media, mostly in hard disks.

- ## *Flow of information*

Information flowing through electronic channels – telephone, radio, TV, and the Internet – contained almost **18 exabytes** of new information in 2002, three and a half times more than is recorded in storage media. 98% of this total is the information sent and received in telephone calls – including both voice and data on both fixed lines and wireless.

- ## *How many web searches are conducted per day?*

According to SearchEngineWatch.com, as of January 2003, there were 319 million searches performed per day by the major search engines. This figure is calculated using the Nielsen/NetRatings "search hours"; the total time spent by all visitors searching at each engine. Table 2.2 shows search statistics per search engine:

[1] *1 terabyte* = 10^{12} *bytes*
= *1,000 gigabytes*

[2] *1 exabyte* = 10^{18} *bytes*
= *1,000,000,000 gigabytes*

5 exabytes = *All words ever spoken by human beings. (Source: How much information? 2003)*

Table 2.2: Search statistics per search engine (Source: How much information? 2003)

Search engine	Search hours per month (in millions)	Search minutes per day (in millions)	Searches per day (in millions)
Google	18.7	37	112
AOL Search	15.5	31	93
Yahoo	7.1	14	42
MSN Search	5.4	11	32
Ask Jeeves	2.3	5	14
InfoSpace	1.1	2	7
AltaVista	0.8	2	5
Overture	0.8	2	5
Netscape	0.7	1	4
Earthlink	0.4	1	3
Looksmart	0.2	0	1
Lycos	0.2	0	1
TOTALS	**53.2**	**106**	**319**

- ***How many hours do individuals spend online?***

  The average global Internet user spends 11 hours and 24 minutes online per month, according to Nielsen/NetRatings. The average user in the United States spends more than twice that amount of time online: on average, 25 hours and 25 minutes at home and 74 hours and 26 minutes at work.

It can be seen that information is definitely not scarce, and new information is clearly also being generated and renewed at an astounding rate, as new technologies like hand-held PCs make the generation of information (often by untrained users) increasingly simpler (Paes 1994). It is further clear that people realise that the information they seek can be found on the Internet, and therefore spend large amounts of time searching for it. The problem is that there currently exists too much information on most topics for any person to review, organise or even just to absorb. Instead of being starved for information, we find ourselves overloaded (Borchers et al. 1998).

2.3 Finding the "right" information

Because of the overabundance of information in the world, finding any information on a certain topic is very little trouble. In some cases systems (like computer-mediated communication systems or CMCS's) have been employed to push information to users, negating the need to go look for information. Based on a number of user-defined criteria, these systems scour the Internet, finding information coinciding with the criteria, and pushing it to the user. If the user however does not require this information at a specific point in time, he/she may be flooded very rapidly with superfluous information if it does not fit the situation in which the user currently finds him/herself (Hiltz & Turoff 1985). In most cases, the need for certain

information is subject to a specific situation or set of circumstances, associated with certain time aspects and dependant on a certain level of quality to make it useful to its target recipients. The information must therefore conform to certain *contextual requirements* to satisfy the information seeker's needs. Carlson (2003) put it very aptly when he said that a query about any given matter stands in a sort of competition with vast amounts of information which would be suitable to answering some other query, but not this particular one.

When searching for information that conforms to these contextual requirements, users searching for electronic information are further confronted with a dimension to the information that was taken for granted in the days when research was mostly done for libraries: quality of content. Books, journals and other reviewed information found in libraries (and now on selected websites on the Internet) are evaluated by experts before being shared with a wider audience, to ensure that a certain level of quality and correctness exist in the information. When using this information, one knows that it is accepted by experts on the topic, and represents the views of an authoritative community. The Internet on the other hand is credibility- and value-neutral (Berghel 1997), and in itself does not enforce any rules or a process to ensure the quality of its content. Any person can create and publish a website, which means that documents with varying levels of quality, written by authors with varying levels of credibility are portrayed to have equal value. Excellent resources therefore usually reside alongside the most dubious (Kirk 1996).

Borchers et al. (1998) stated that when information is abundant, the knowledge of which information is useful and valuable, matters most. Having this knowledge is obviously the key to finding the correct information very quickly, but as this is very seldom the case, searching for information can become a very arduous task. With the Internet providing some assistance to find information of sufficient quality, and little to no assistance in providing information with the right context, information seekers are usually tasked to assess great amounts of information to find a specific piece of information. The problem with having to assess vast amounts of information (aside from having cost and time implications), is that the sheer volumes can have an adverse effect on the information locating effort, leading to a 'condition' called *information overload.*

2.4 What is information overload?

Various definitions and explanations for information overload have been put forward, always aimed at conveying the notion of receiving too much information to process efficiently (Eppler & Mengis 2004). A very descriptive definition for information overload is the 'provision of information in excess of the cognitive and emotional ability of an individual to process that information.' (Canadian Food Inspection Agency n.d.)

'Researchers across various disciplines have found that the performance (i.e., the quality of decisions or reasoning in general) of an individual correlates positively with the amount of information he or she receives — up to a certain point. If further information is provided beyond this point, the performance of the individual will rapidly decline (Chewning & Harrell, cited in Eppler & Mengis 2004). The information provided beyond this point will no longer be integrated into the decision-making process and information overload will be the result (O'Reilly, cited in Eppler & Mengis 2004). The burden of a heavy information load will confuse the individual, affect his or her ability to set priorities, and make prior information harder to recall (Schick et al., cited in Eppler & Mengis 2004). Figure 2.3 provides a schematic version of this discovery. It is generally referred to as the inverted U-curve, following the initial work of Schroder, Driver, and Streufert' (Schroder et al, cited in Eppler & Mengis 2004).

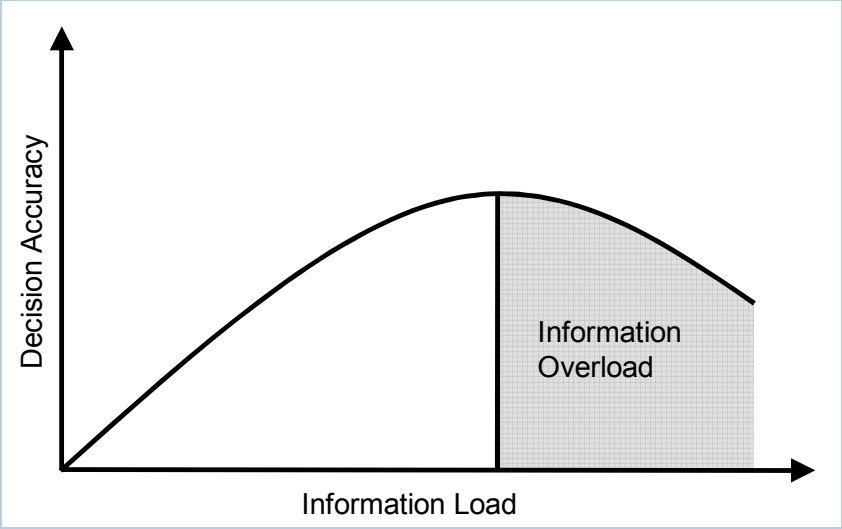


Figure 2.3: Information overload as the inverted U-curve (Source: Eppler & Mengis 2004)

Different approaches have been defined to determine when a party trying to locate specific information, is likely to experience information overload. The most general of these is by comparing the volume of information supply with the information processing capacity of an individual (Eppler & Mengis 2004). A more "classic" definition of information overload, based on the information-processing view of the organisation suggested by Galbraith (cited in Eppler & Mengis 2004) and expanded by Tushman and Nadler (cited in Eppler & Mengis 2004), is that information overload can be explained through the following formula:

$$\boldsymbol{information\ processing\ requirements} > \boldsymbol{information\ processing\ capacities}$$

The terms "requirements" and "capacities" in this definition can be measured in terms of available time (Eppler & Mengis 2004). In other studies (Iselin; Keller & Staelin; Owen; Scheider, all cited in Eppler & Mengis 2004), not only the amount of information and the available processing time (i.e., the quantitative

dimension), but also the characteristics of information (i.e., the qualitative dimension) are seen as major overload elements. This aspect might be seen as the overall quality, or usefulness of the available information (Eppler & Mengis 2004). Iselin (cited in Hwang & Lin 1999), has noted that an information cue set can be defined as having two distinct information dimensions, these being the number of different dimensions (or diversity), and the number of repeated dimensions (or repetitiveness). A study conducted by Hwang & Lin (1999) demonstrated that decision quality suffers with an increase (i.e. overload) in either the diversity or repetitiveness of an information cue set. Yet another viewpoint is that information overload can be regarded on the basis of subjective experience (as opposed to the more objective aspects of the phenomenon) (Eppler & Mengis 2004). Finally, information overload can of course be seen as a combination of both objective and subjective factors, as illustrated by the features characterising information overload provided by Meyer (1998):

- High information load;
- Decision situation and thus pressure to receive information;
- Limited processing capacity;
- Resulting stress and suboptimal or dysfunctional decision behaviour.

It is clear to see that a great number of different influences can contribute to the manifestation of information overload within a person (or persons) searching for specific information. Before looking more closely at the causes of information overload however, the nature of information overload in organisational environments will first be considered.

2.5 Organisations' experience of information overload

Analogous of the "Time is money" locution associated with the business world, "Information is money" (Bentley 1998) has become very applicable to organisations in the information-rich world of today. Because of the vast amounts information available on markets, competition and technology, organisations are in a better position than ever before to make good decisions, ensuring their continued existence and even becoming market leaders in their particular market segment. With the explosion of available exploitable information however, organisations have had to learn to adapt in very short periods. To illustrate, Edmunds & Morris (2000) stated that 'computer processing speed has doubled every two years for the past thirty years and today, the speed of change constantly increases the pressure to adapt at a rate that becomes almost impossible (Shenk cited in Edmunds & Morris 2000).' To successfully grasp opportunities,

organisations need to adapt quickly and to do that, information of high value and quality must be acquired in a very short space of time.

Information is seen by most organisations as the key to success, which means that many people have to deal with overwhelming amounts of information in their daily work (seemingly) for the benefit of the organisation. People simply cannot afford to ignore information in the workplace (Edmunds & Morris 2000), for fear of missing something of importance. Their ability to provide information of high value and quality for effective decision making is unfortunately hampered by the nature of the information being gathered. Information both internal and external to the organisation needs to be signed off by a source of sufficient authority before it can be taken seriously and included in the organisation's decision processes, as information that is irrelevant or of dubious quality can be harmful to employees and the organisation (Hagel & Singer cited in Ho & Tang 2001). The ease of information creation and sharing (see section 2.2), have therefore created a situation where organisations also have to thoroughly assess information (even information generated internally) before utilising it. Again, the problem of information overload comes to the fore, but unlike casual Internet users, organisations rarely have the luxury to sift through masses of information to acquire the information necessary for making immediate business decisions (Kotze et al. 2007), as this consumes valuable time and resources for an apparently non-value-adding task.

Information overload (and the quality and value of information in general) is currently being perceived as a very real problem in organisations, especially with regards to the management of information, as the survey results shown in Figure 2.4 illustrate. The survey on various information management issues was developed by VNU Exhibitions Europe with the support of the Information Management Professional (IMP) Group and yielded 648 respondents, 30% of which were information professionals, 20% were information managers, 14% were IT professionals and the remainder were made up of researchers, analysts, consultants, publishers, project managers and senior management. 35% of respondents described 'information overload' and 'improving the quality of internal information' as their company's main information management challenge, whereas 25% cited 'poor or no document management' and 24% 'sourcing the right external information' as the main challenge (Information Management Independent Research Results 2006).

It is very obvious that information overload, and the overall way information is managed, is a very significant problem in modern-day organisations who, in a continuous effort to maintain their foothold in a specific market segment while trying to out-manoeuvre their competition, are gathering copious amounts of information to be processed and stored, and finally used in their decision making processes. Unfortunately it would also seem that the processing and storing of information actually yields very little information that actually make it to the decision making stage. Carlson (2003), in quoting some organisational information overload statistics, makes this very apparent, stating that 80% of organisational information filed is never used. Still, 71% of workers say their main job is tracking down information

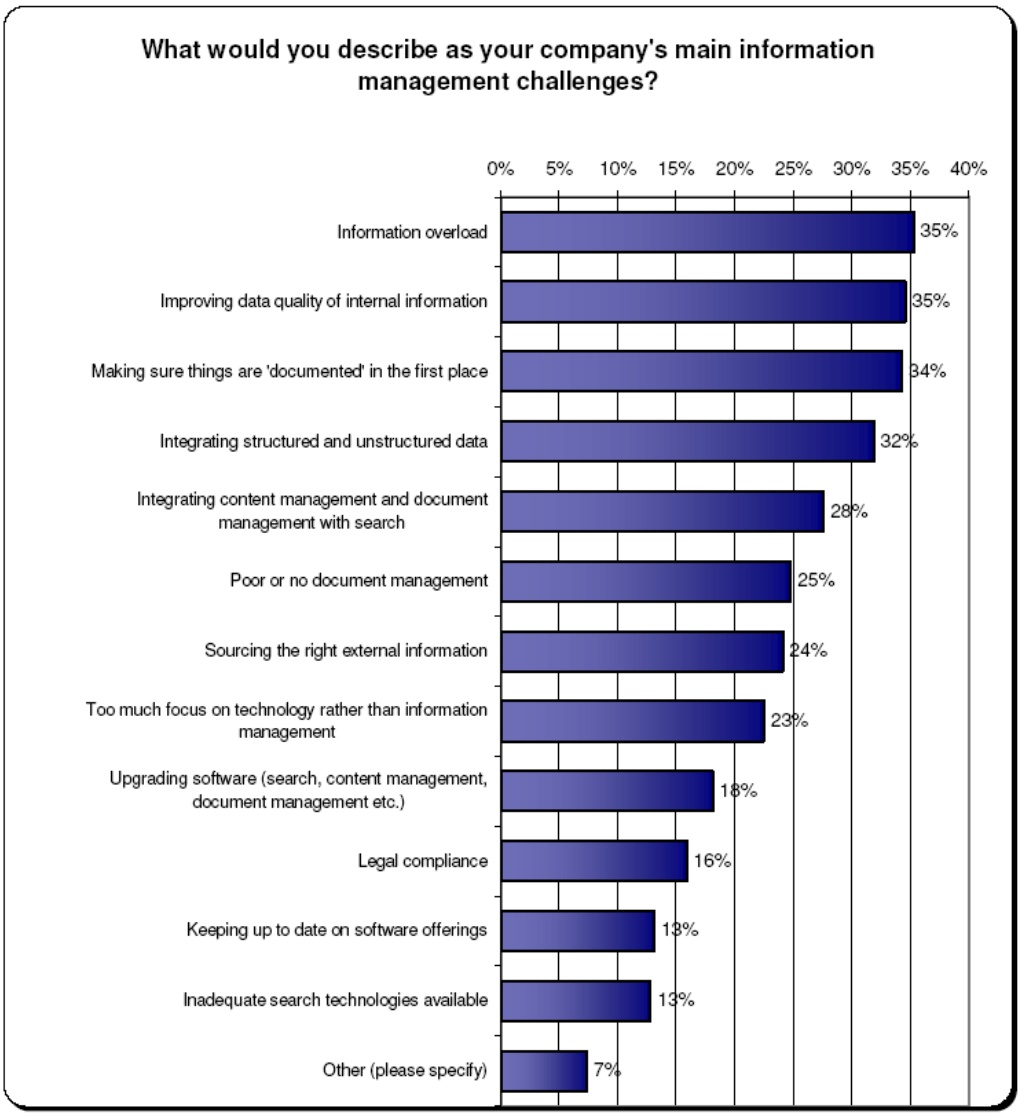


Figure 2.4: Survey results - Organisational information management challenges (Source: Information Management Independent Research Results 2006)

(Carlson 2003), which seems like a terrible waste of resources on something that only has a 20% chance of contributing to the organisation.

The issues of information management in organisations and the support of decision making in organisations will be discussed further in Chapter 3.

2.6 Causes of information overload

Carlson (2003) mentioned that 'technostress (i.e. information overload) results in reduced intellectual performance and poor judgment which, in a sort of negative feedback loop, partly causes, and is also partly

a result of, haphazard and random use of ICT' (Information and Communication Technologies), and that furthermore, the 'lack of a coherent conceptual knowledge management framework can also act as an aggravating factor.' The common perception is that ICT is the main cause (and the main driver) of the problem of information overload. The reality however, is that there are many disparate factors that can cause or contribute to the phenomenon of information overload.

In a review of literature from organisational science, accounting, marketing, MIS (management information systems) and other related disciplines on the problem and experience of information overload, Eppler and Mengis (2004) have grouped the causes of information overload into the following 5 main groupings: the *information* itself (with regards to quality, quantity, frequency, and intensity), the *person* receiving, processing or communicating information, the *tasks or processes* that needs to be completed by a person, team or the organisation, the *organisational design* (i.e. the formal or informal work structures), and the *information technology* used (and how it is used) in the organisation (see Figure 2.5).

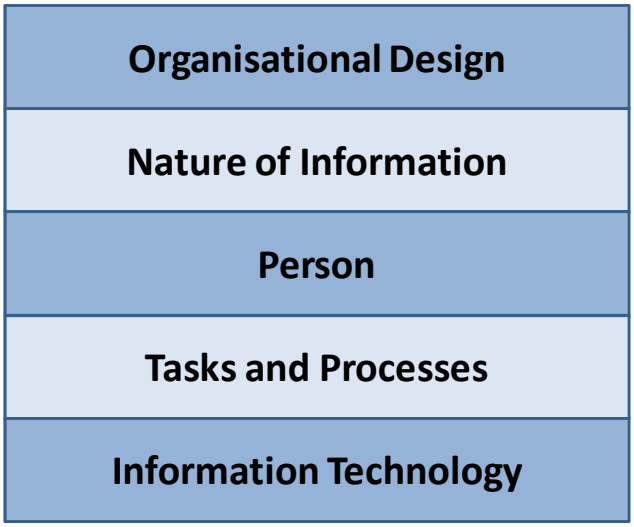


Figure 2.5: Causes of information overload

Of course, more often than not, it is not any single cause giving rise to information overload in a certain situation, but rather a combination of the five causes mentioned. Eppler and Mengis (2004) further note that all five of the causes for information overload influences the two fundamental variables of information overload, namely the information processing capacity (IPC) – which is (for example) influenced by personal characteristics – and the information processing requirements (IPR) – which are usually associated with the nature of the task or process to be performed (mentioned in section 2.4). The five main groupings of information overload causes will now be discussed in more detail:

- ***Organisational design***

Changes to the organisational design may result in existing information, communication, control and reporting structures not being capable to provide the information required to effectively manage the organisation in its new form, placing an increased IPR burden on the employees while the structures are being adapted. Disintermediation or centralisation (Schneider cited in Eppler & Mengis 2004), or because of a change to an interdisciplinary team approach (Bawden cited in Eppler & Mengis 2004) for instance, can lead to an increase in IPR as a need is created for more intensive communication and coordination. Conversely, effective coordination through the use of standards, common procedures, rules or dedicated coordination centres (Galbraith cited in Eppler & Mengis 2004) can reduce the IPR and positively influence the IPC of an organisation (Galbraith; Schick et al.; Tushman & Nadler; Schneider, all cited in Eppler & Mengis 2004).

- ***Nature of information***

It is not only the quantity of information that may cause information overload, but the characteristics of the information as well. These characteristics include the level of uncertainty associated with information and the level of ambiguity, novelty, complexity and intensity of the information (Schneider cited in Eppler & Mengis 2004). Ho & Tang (2001) presented a consolidated view on the nature of information, by stating that information overload can be influenced by the following dimensions of information (differing from the dimensions described by Hwang & Lin in section 2.4): the quantity of the information, the quality of the information, and the format (or diversity) of the information (see Figure 2.6).

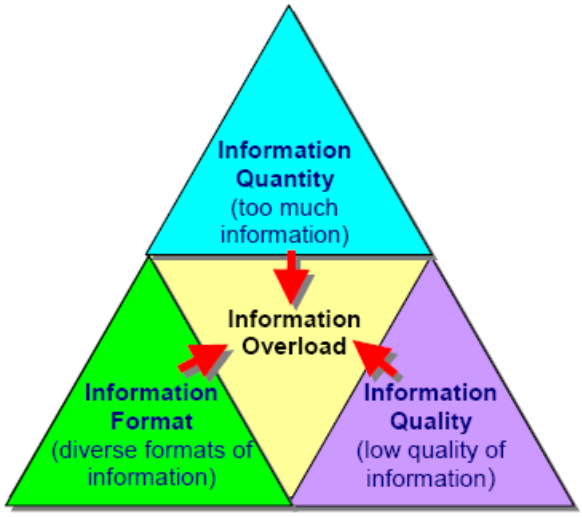


Figure 2.6: Dimensions of information overload (Source: Ho & Tang 2001)

An improvement in quality (e.g. conciseness, consistency, comprehensibility, etc.) can potentially improve the IPC of an individual, as the consistent and improved quality of information being experienced will allow the individual to perform information tasks in a much shorter time, as well as performing it according to a higher standard than would have otherwise been the case. Information quantity has an obvious relation to IPC, and finally, very diverse information may require a person to change context with every piece of information being processed, which limits his/her IPC, thus increasing the amount of time required to complete specific work, which in turn may lead to an increase in work pressure.

- ***Person***

The nature of the person working with information is another cause of information overload. Attitude, qualification and experience can all contribute to the person's ability to cope with large amounts of information, and a lack thereof can contribute to the possible experience of information overload. Studies have shown that any person has a limited capacity to process information (Jacoby et al.; Galbraith; Malhotra; Simon; Tushman & Nadler, all cited in Eppler & Mengis 2004), while further limiting factors were also identified, such as personal skills (Owen cited in Eppler & Mengis 2004), the level of experience (Swain & Haka cited in Eppler & Mengis 2004), and motivation (Muller cited in Eppler & Mengis 2004). It is therefore clear that personality traits can directly impact a person's IPC.

- ***Tasks and processes***

The tasks and processes that need to be completed with the help of information is another important factor that can cause or contribute to information overload. It has been found that the less a process is based on reoccurring routines (Tushman & Nadler, cited in Eppler & Mengis 2004) and the more complex it is in terms of the configuration of its steps (Bawden; Grise & Gallupe, all cited in Eppler & Mengis 2004), the higher the information load will be and therefore the greater the time pressure will be on the individual (Schick et al. cited in Eppler & Mengis 2004). These two factors can obviously cause an increase in the IPR, which can lead to information overload. Information overload is especially likely if the process is frequently interrupted and the concentration of the individual suffers as a consequence (Speier et al. cited in Eppler & Mengis 2004). This means that complex tasks or processes directly increase the IPR, a situation that can further be aggravated by a reduced IPC as a result of frequent context switching or distraction.

- ***Information technology***

The development and deployment of new information and communication technologies, such as the Internet, intranets, and extranets, but especially e-mail, are universally seen as one major cause of information overload (Bawden cited in Eppler & Mengis 2004). As researchers have indicated that the lack of effective tools such as high quality search engines (Kehoe cited in Ho & Tang 2001) for information retrieval may contribute to the overload problem (Hiltz & Turoff cited in Ho & Tang 2001), different technologies and approaches have been developed in an attempt to manage the flow of information and provide users with the exact information they require – with varying success. At this stage searching and relevancy ranking algorithms still struggle to quickly provide a user with the information requested, and pushing technologies, in an attempt to provide the user with information that might be valuable, very easily end up flooding the user with unwanted information and causing a great deal of interruptions (Edmunds & Morris; Speier et al., all cited in Eppler & Mengis 2004). It can therefore be seen that, while ICT has drastically increased the individual's IPC, IPR has also increased at the same time (Eppler & Mengis 2004).

For a more detailed list of information overload causes, see Table 2.3.

Table 2.3: Causes for information overload (Source: Eppler & Mengis 2004)

Causes for Information Overload	
Personal factors	Limitations in the individual human information processing capacity
	Decision scope and resulting documentation needs
	Motivation, attitude, satisfaction
	Personal traits (experience, skills, ideology, age)
	Personal situation (time of the day, noise, temperature, amount of sleep)
	Senders screen outgoing information insufficiently
	Users of information adapt their way of interacting with computers too slowly with respect to the technological development
	Social communication barrier break down
Information characteristics	Number of items of information rises
	Uncertainty of information (info needed vs. information available)
	Diversity of information and number of alternatives increase
	Ambiguity of information
	Novelty of information
	Complexity of information
	Intensity of information
	Dimensions of information increase
	Information quality, value, half-life
	Overabundance of irrelevant information
Task and process parameters	Tasks are less routine
	Complexity of tasks and task interdependencies
	Time pressure
	Task interruptions for complex tasks
	Too many, too detailed for complex tasks
	Too many, too detailed standards (in accounting)
	Simultaneous input of information into the process

Causes for Information Overload	
	Innovations evolve rapidly – shortened life cycle
	Interdisciplinary work
Organisational design	Collaborative work
	Centralisation (bottlenecks) or disintermediation (information searching is done by end users rather than by information professionals
	Accumulation of information to demonstrate power
	Group heterogeneity
	New information and communication technologies (e.g. groupware)
Information technology	Push systems
	E-mails
	Intranet, extranet, Internet
	Rise in number of television channels
	Various distribution channels for the same content
	Vast storage capacity of the systems
	Low duplication costs
	Speed of access

Finally, an aspect which does not fit neatly into any of the five types of causes for information overload given by Eppler and Mengis, is the fact that the reasoning behind the information processing requirements might be the cause for information overload. In an organisational setting a manager might request for much more information than he/she actually needs for a task, for the wrong reasons. Butcher (cited in Edmunds & Morris 2000) provided a list of seven reasons why managers obtain so much information that they can easily be overwhelmed by it all:

- They collect information to indicate a commitment to rationalism and competence which they believe improves decision-making;
- They receive enormous amounts of unsolicited information;
- They seek more information to check out the information already acquired;
- They need to be able to demonstrate justification of decisions;
- They collect information just in case it may be useful;
- They play safe and get all information possible; and
- They like to use information as a currency – not to get left behind colleagues.

2.7 Symptoms of information overload

It is very clear that a heavy information load can have a negative effect on the performance of an individual (whether measured in terms of accuracy or speed). According to Eppler & Mengis (2004), when information supply exceeds an individual's information-processing capacity, a person has difficulties in identifying the relevant information (Jacoby cited in Eppler & Mengis 2004), becomes highly selective and

ignores a large amount of information (Bawden; Herbig & Kramer; Sparrow, all cited in Eppler & Mengis 2004), has difficulties in identifying the relationship between details and the overall perspective (Schneider cited in Eppler & Mengis 2004), needs more time to reach a decision (Jacoby cited in Eppler & Mengis 2004), and finally does not reach a decision of adequate accuracy (Malhotra cited in Eppler & Mengis 2004).

Various symptoms of information overload have been identified over the years, ranging from psychological conditions experienced by persons due to the detrimental effect of information overload, to adaptive mechanisms being utilised by persons in order to deal with excessive amounts of information, unfortunately at a loss to effectiveness. Some of the more frequently experienced psychological symptoms identified through information overload studies in organisations are *a general lack of perspective* (Schick et al. cited in Eppler & Mengis 2004), *cognitive strain and stress* (Malhotra; Schick et al. all cited in Eppler & Mengis 2004), *a greater tolerance of error* (Sparrow cited in Eppler & Mengis 2004), *lower job satisfaction* (Jacoby cited in Eppler & Mengis 2004), and the *inability to use information to make a decision* (Bawden cited in Eppler & Mengis 2004) — the so-called "paralysis by analysis". Some of the most common reactions of persons attempting to deal with excessive amounts of information were already identified in 1970 by Milgram (cited in Eppler & Mengis 2004) as being:

- The allocation of less time to each input;
- The disregard of low-priority inputs;
- The redrawing of boundaries in some social transactions to shift the burden of overload to the other party of the exchange;
- The reduction of inputs by filtering devices;
- The refusal of communication reception (via unlisted telephone numbers, unfriendly facial expressions, etc.); and finally
- The creation of specialised institutions to absorb inputs that would otherwise swamp the individual (Weick cited in Eppler & Mengis 2004).

Miller (cited in Meyer 1998) identified *adaptive mechanisms* utilised by information overloaded persons in dealing with overload situations:

- *Chunking*: gathering and processing certain generic terms instead of single pieces of information;
- *Omission*: occasional, unsystematic skipping of information;
- *Queuing*: the deferment of information processing at peak work load;
- *Filtering*: systematic neglect of irrelevant information; and
- *Capitulation* (escape from the task): prematurely discontinuing the information processing.

These mechanisms, whilst allowing the individual to manage masses of information, unfortunately prevent him/her from properly assessing the information in question, and fully realising its full value. It is clear that solutions are required to provide individuals with access to the information they require whilst protecting them from the psychological effects of information overload, and preventing them from employing techniques which, even though it allow them to process much more information, cause them to assess the information much less thoroughly than is actually required.

2.8 Solutions for information overload

Over the years, many solutions have been proposed to mitigate the effect of information overload on individuals and groups (such as organisations). Some suggestions for combating information overload are provided below, in terms of the five groupings identified in section 2.6 (Eppler & Mengis (2004):

- ***Organisational design***

On the topic of solutions for information overload in terms of organisational design, it would appear that there are conflicting viewpoints on how the problem could be solved. Eppler & Mengis (2004) have actually found that authorities in this field take on conflicting positions, as some feel that *self-contained tasks* and *lateral relationships* (Galbraith cited in Eppler & Mengis 2004) can solve or at least mitigate the problem of information overload, whereas others feel that the focus on collaborative and interdisciplinary work is a cause rather than as a countermeasure of information overload (Bawden; Wilson, all cited in Eppler & Mengis 2004).

- ***Nature of information***

With regard to the information itself, information overload can be reduced if efforts are made to *assure that information is of high value* (Eppler & Mengis 2004), that it is *delivered in the most convenient way and format* (Simpson & Prusak cited in Eppler & Mengis 2004), that it is *visualized, compressed, and aggregated* (Ackoff; Meyer, all cited in Eppler & Mengis 2004), and that *signals and testimonials are used* to minimize the risks associated with information (Herbig & Kramer, cited in Eppler & Mengis 2004). Hwang and Lin (1999) stated that common techniques such as aggregation and summarization should be used to keep the number of information dimensions (both diversity and repetitiveness, see section 2.4) to a minimum, and that the optimal number of information dimensions for each task should be determined empirically in an effort to prevent information

overload from occurring. Ho and Tang (2001) felt that an infomediary[3] approach could be utilised to mitigate the risk of information overload. They evaluated some known infomediary models in terms of the dimensions of information they have defined, namely information quantity, quality and format (see section 2.6, Nature of Information), the result of which is shown in Table 2.4[4]. The results clearly show that all these "agents" can control information quantity very well, but (as a group) are much less successful with ensuring information format and quality, giving a clear indication that the infomediary approach can unfortunately merely mitigate the effect of information overload, and not solve the problem completely.

Table 2.4: An infomediary approach to information overload solutions (Source: Ho & Tang 2001)

Infomediary Models	Information Quantity	Information Format	Information Quality
Portal	✓		
Virtual Community	✓		✓
Transaction Aggregator	✓		✓
Syndication	✓	✓	
Personalisation	✓		
Comparison	✓		✓

- ***Person***

On the individual level, *training programs to augment the information literacy of information consumers* (Bawden; Koniger & Janowitz; Schick et al., all cited in Eppler & Mengis 2004) was deemed very important, together with *providing employees with the right tools for improving their efficiency* (Bawden cited in Eppler & Mengis 2004) and *information management skills* (Edmunds & Morris cited in Eppler & Mengis 2004).

[3] *'Consumers won't have the time, the patience, or the ability to work out the best deals with information buyers on their own. In order for consumers to strike the best bargain with vendors, they'll need a trusted third party - a kind of personal agent, information intermediary, or infomediary - to aggregate their information with that of other consumers and to use the combined market power to negotiate with vendors on their behalf.' (Hagel cited in Ho & Tang 2001)*

[4] *Brief descriptions of the indicated infomediary models are as follows: Portal - a web site that provides broad services including searching, yellow pages, and link to other sites; Virtual Community - a group of individuals that share common interest and ideas among community members through e-mail, chatting, or newsgroups. Transaction aggregators create electronic market places to enable critical mass of vendors and customers to connect with each other and conduct transactions based on the rules and procedures. Syndication - a form of infomediary that collects and packages digital information, aggregating content from many sources. Personalisation - the concept of deciding – given a large set of possible choices – which piece of information has the highest value to an individual. Comparison sites, such as Gomez.com (http://www.gomez.com), or Bizrate.com (http://www.bizrate.com), allow users to search for a product by name, category (e.g. computers, insurance, or travel) or a number of specific parameters (e.g. price, speed, or service quality) (Ho & Tang 2001).*

- ***Tasks and processes***

For information overload experienced as a result of process problems, the following have been proposed as possible remedies: *standardisation of operating procedures* (Bawden; Schick et al.; Schneider, all cited in Eppler & Mengis 2004), *collaboration with information specialists* within the process teams (Edmunds & Morris cited in Eppler & Mengis 2004), and *the use of facilitators or collaborative tools (such as virtual team rooms) as "process enablers" for cognitive support* (Grise & Gallupe cited in Eppler & Mengis 2004).

- ***Information technology***

Finally, when looking at the impact that information technology can have on addressing information overload, *the use of intelligent information management systems for fostering an easier prioritisation of information* (Bawden; Meyer; Schick et al., all cited in Eppler & Mengis 2004) and *the providing of quality filters* (Ackoff; Edmunds & Morris; Grise & Gallupe, all cited in Eppler & Mengis 2004) are some of the main solutions advocated. Examples of such intelligent systems are *decision support systems* (DSS) aimed at portraying large quantities of information in a manageable manner and size (Cook cited in Eppler & Mengis 2004). Advocating a less-traditional solution to the problem of information overload, Belfourd and Furner (cited in Edmunds & Morris 2000) feels that intelligent agents should be used because of the fact that an intelligent agent acts autonomously by making decisions on the basis of data it acquires about the environment, rather than as a result of direct instruction from the user, as well as the fact that an intelligent agent has the facility to learn about the personal preferences of an individual so that it is able to gradually predict the likelihood of items that will be of interest to the user (see also Paes (1994)).

A more detailed list of countermeasures against information overload compiled by Eppler and Mengis (2004) is provided in Table 2.5.

Table 2.5: Countermeasures against information overload (Source: Eppler & Mengis 2004)

Countermeasures against Information Overload	
Personal factors	Improve personal time management skills and techniques
	Training programs to augment information literacy: information-processing skills such as file handling, using e-mail, classification of documents, etc.
	Improve personal information management
	Systematic priority setting
	Improve the screening skills for information
Information characteristics	Raise general quality of information (i.e. its usefulness, conciseness) by defining quality standards
	Focus on creating value-added information
	Promulgation of rules for information and communication design (e.g. e-mail etiquette)

Countermeasures against Information Overload	
	Compress, aggregate, categorize and structure information
	Visualisation, the use of graphs
	Formalisation of language
	Brand names for information
	Form must follow function must follow usability
	Simplify functionalities and design of products
	Customisation of information
	Intelligent interfaces
	Determine various versions of an information with various levels of detail and elaborate additional information that serves as summaries
	Organise text with hypertext structures or gophers
	Interlink various information types (as internal with external information)
Task and process parameters	Standardise operating procedures
	Define decision models developed for specific decision processes (e.g. decision rules)
	Install an exception-reporting system
	Allow more time for task performance
	Schedule interrupted blocks of time for completing critical work
	Adequate selection of media for the task
	Handle incoming information at once
	Collaboration with information specialists within the teams
	Bring decisions to where information exists when this information is qualitative and ambiguous
	Install process enablers for cognitive support
	Use simpler information-processing strategies
	Regulate the rate of information flow
	Search procedures and strategy
	Define specific, clear goals for the information in order to contextualize it and turn it meaningful
	Communicate information needs to providers
	Provide incentives that are directly related with decisions in order to make decision relevant information be processed more efficiently
	Install a measurement system for information quality
Organisational design	Coordination through interlinked units
	Augment info processing capacity through changes in org. design
	Creation of lateral relationships (integrate roles, create liaisons between roles, teamwork etc.)
	Coordination by goal setting, hierarchy and rules depending on frequency of exceptions (uncertainty)
	Creation of self-contained tasks (reduced division of labour, authority structures based on output categories) → autonomous groups
	Reduce divergence among people (e.g. with regard to expectations) through socialisation (e.g. frequent face-to-face interactions)
	Install appropriate measures of performance
	Hire additional employees
	Create slack resources
Information technology application	Intelligent information management (prioritisation)
	Install voting structures to make users evaluate the information
	Prefer push to pull technologies
	Facilitator support through (e-)tools
	Decision support systems should reduce a large set of alternatives to manageable size

Countermeasures against Information Overload	
	Use natural language processing systems (search with artificial intelligence)
	Information quality filters
	Intelligent data selectors (intelligent agents)
	Use systems that offer various information organisation options (e.g. filing systems)

2.9 Conclusion – Information overload

People's innate ability to generate and share information has led to massive amounts of information that is increasing continuously, and lacks consistency in terms of quality. Workers having to assess information are overwhelmed by the sheer volume and diversity of this information, and it is therefore quite clear that they require assistance in performing their work efficiently.

Different causes for the occurrence of information overload have been identified in this chapter as being the person him/herself, the characteristics of information, the parameters of tasks and processes to be completed, organisational design and the application of information technology. The rest of this thesis will focus solely on addressing the problem of information overload in terms of information technology and the nature of information itself, and any reference to information overload from this point on will signify only that.

This aim of this thesis is to discuss a way to solve the problem of information overload for knowledge workers and decision makers in an organisation, or at least to cushion aforementioned persons against the full impact of information overload through an implementation of a combination of the following solutions (mentioned in the previous section):

- Delivering the information in the most convenient way and format;
- Visualisation, compression and aggregation of the information;
- Using methods of intelligent information management for easier information prioritisation; and
- Using quality filters.

The next chapter will explain the manner in which information is used within an organisation, which has a direct influence on its decision making ability.

Chapter 3 – Organisations and Information

3.1 Introduction

> 'Information is not merely a necessary adjunct to personal, social and organisational functioning, a body of facts and knowledge to be applied to the solution of problems or to support actions. Rather it is a central and defining characteristic of all life forms, manifested in genetic transfer, in stimulus response mechanisms, in the communication of signals and messages and, in the case of humans, in the intelligent acquisition of understanding and wisdom' (Kaye, cited in Rowley & Farrow 2000).

More and more humanity have started to realise the importance of information for its continued survival and success. Some of the more significant ways in which information is employed by humans are the following (Rowley & Farrow 2000):

- Decision-making;
- Problem-solving;
- Communication and interpersonal relationships;
- Learning;
- Entertainment & leisure;
- Citizenship; and
- Business & professional effectiveness.

> 'Without an uninterrupted flow of the vital resource, society as we know it would quickly run into difficulties, with business and industry, education, leisure, travel, and communications, national and international affairs all vulnerable to disruption. In more advanced societies this vulnerability is heightened by an increasing dependence on the enabling powers of information and communications technologies' (Martin, cited in Rowley & Farrow 2000).

Information has been an integral part in the world of business from the beginning, assisting decision makers in determining the exact specifications of a product or service to be provided to a customer, how to ensure that the value offering of the organisation is superior to that of the competition, and how to deliver it in the most effective and efficient manner. Internal to the organisation, information is more than a

strategic resource; it acts as the "glue" that keeps the organisation together, a meta-resource that coordinates the mobilisation of the other assets in order for the organisation to perform (Choo 2000). Externally, a much larger information environment exists which appears to be very turbulent in nature, as various persons, organisations and other entities vie for the information they require to elevate themselves above their peers. Choo (2000) very aptly states that competition is the result of an 'unequal distribution of information among organisations and their differential abilities to acquire, absorb, and actuate information.' Technological advances however, have (through the advent of integrated networks resulting in the Internet) managed to level the playing field somewhat, providing organisations with the means to have much more direct access to customers. As interaction with customers increases, organisations are able to improve their service levels and enhance their competitive intensity in a concurrent manner (Chen & Tai 2003).

For a long time organisations saw themselves as closed systems whose only interaction with the "outside world" was through the delivery of products and the acceptance of supplies, whilst buffering its operations from variations experienced in the outside world and focusing on improving internal structures and functions (Choo 2000). The truth of the matter is that an organisation actually behaves as an open system that shares many features with living biological systems and interacts on a very fundamental level with its environment. The application of information is the means through which organisations navigate in the environment, and organisations extract information from the surrounding environment through perceiving stimuli, interpreting and storing it for later use, retrieving and transmitting it when needed, and using information to generate judgements and solve problems (Larkey & Sproull, cited in Choo 2000). Through these means, organisations who (much like species of organisms) basically seek sustenance and growth in a dynamic environment (Choo 2000) utilise information to adapt to new and changing conditions.

In order for the organisation to successfully navigate this environment, it needs effective interpretative and decision-making capabilities to better understand the environment and to determine the optimal combination of information, resources and variation (all being part of the environment) (Choo 2000) that will be most beneficial to its cause. The effective managing of unpredictability, probability, responsibility, risks and experience in the organisation's environment all form part of making decisions in an effective manner (Baupin & Zreik 2000). In order to make good decisions, humans need specific information from their environment to support their decisions, and thus (as an extension of this human trait), organisations do the same, collecting information, whether from the external market-place or internal processes and organisations, in order to contribute to the quality of decisions made (Rowley & Farrow 2000).

If one considers that the acts of both possessing information and creating information point to the concept of having intelligence (Gregory, cited in Choo 2000), it appears that (similar to biological organisms) an increase in information will induce an increase in the intelligence of the organisation and will ultimately also start influencing its behaviour. Through the gathering, processing, interpreting and communicating of information obtained for decision making (Wilensky, cited in Choo 2000) and applying knowledge

developed on the current and future conditions of its environment (Choo 2000), such an *intelligent organisation* will start to develop the ability to better deal with complexity and be able to better react on signals from its environment (Haeckel & Nolan, cited in Choo 2000).

It is very obvious that information is indispensable, and that it can have a significant impact on the success of organisations. The potential severe consequences of the loss or inability to find specific information are driving organisations to seek better and more versatile methods and systems to efficiently extract information from the environment, and structure it to maximise its value to the organisation (Choo 2000). The extraction of information from the environment do however also carry the risk of inducing information overload (discussed in Chapter 2) within an organisation, which is why the manner in which an organisation uses and interacts with the information should be very carefully defined, as the incorrect utilisation and interaction with information can very easily have an adverse effect on organisational efficiency and productivity (Choo 2000). This chapter will consider some of the ways in which organisations currently utilise information, the shortcomings of the more commonly used approaches that are employed in organisations, and some improvements that were developed based on these shortcomings, which has improved the extent to which organisations utilise information (see Figure 3.1).

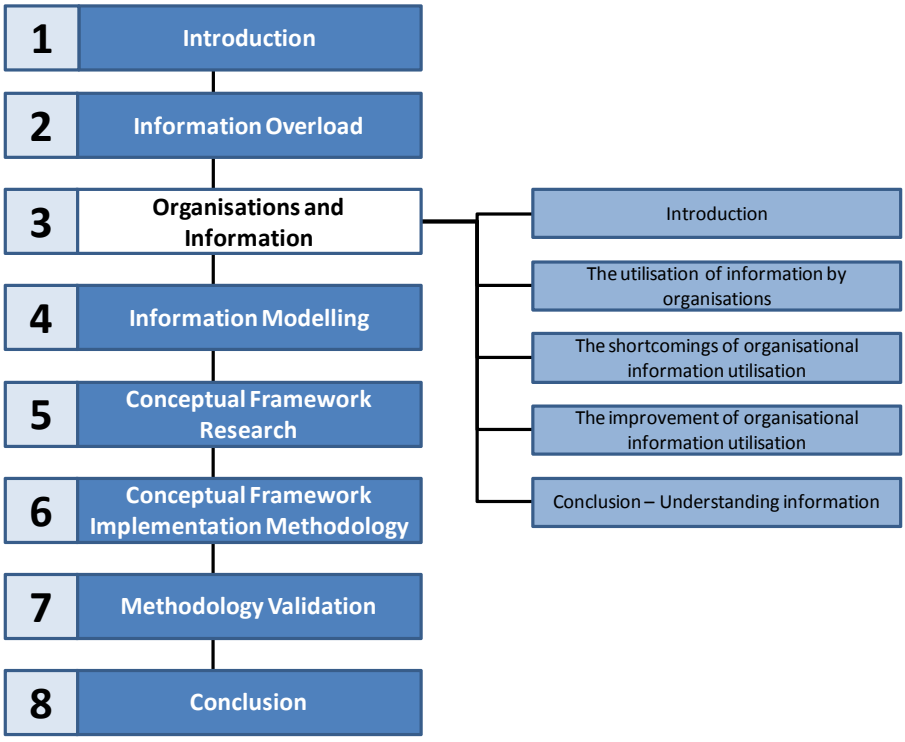


Figure 3.1: Image for navigating thesis and chapter 3

3.2 The utilisation of information by organisations

According to Choo (2000), literature shows that two different research orientations actually exist with regards to organisational information processing (i.e. utilisation). The first views organisations as rational, *decision making systems*. The organisation utilises information to maintain the decision premises that influences the manner in which decisions are made. Information is processed in order to reduce or avoid uncertainty. The organisation first sets its goals, then searches for alternatives, and selects courses of action through which the goals will be achieved. The second orientation sees organisations as loosely coupled social systems, in which individual actors enact or create the environment to which the organisation then adapts. The organisation basically works towards developing a shared interpretation of the environment, and then acting on the basis of this interpretation. Therefore, in the organisation as an *interpretation system*, information is utilised in order to reduce or resolve ambiguity. Actions are often taken first and then interpreted retrospectively.

In both of these viewpoints, it is obvious that information is indispensable to organisations. Most organisations understand this and, through the use of technology, are able to quite effortlessly increase their information (from both internal and external environments (Choo 2000)) at a fierce rate. Today, the majority of organisational information exists as documents and transactions in information systems, rather than on physical media and in the minds of people. 'Much of … organisational information, perhaps most of it, takes the form of documents (e.g. reports, messages, letters, journal and magazine articles, memos, minutes of meetings, research bulletins, etc.), and it is easy to see why documents are important; they are often the organising and interpretive medium that gives data, figures and other information meaning within an organisational context.' Furthermore, 'the Gartner Group has estimated that as much as 90% of an organisation's information is contained in its documents' (Blair 2002). It is therefore imperative that an organisation is well-equipped to handle the information residing in its systems, and be able to provide it to its users in the most efficient manner, especially since decisions made in the organisation are usually based on this information (Choo 2000).

The field of *information management* focuses on the collection and management of information, with its basic goal being the harnessing of information resources and information capabilities so that the organisation learns and adapts to its changing environment (Choo 2000). It can be described in a simple manner by the phrase 'Getting the right information to the right person at the right place at the right time' (Wikipedia, the free encyclopedia 2007c). In the organisation, information management takes place as a cycle of processes designated by Rowley and Farrow (2000) as 'The 7 R's of Information Management' (see Figure 3.2). Starting with the Reading process, the cycle works as follows:

1. 'A person Reads a collection of *Relevant knowledge*' recorded as documentation (electronic or printed) or extracted from the environment through some other data collection method.

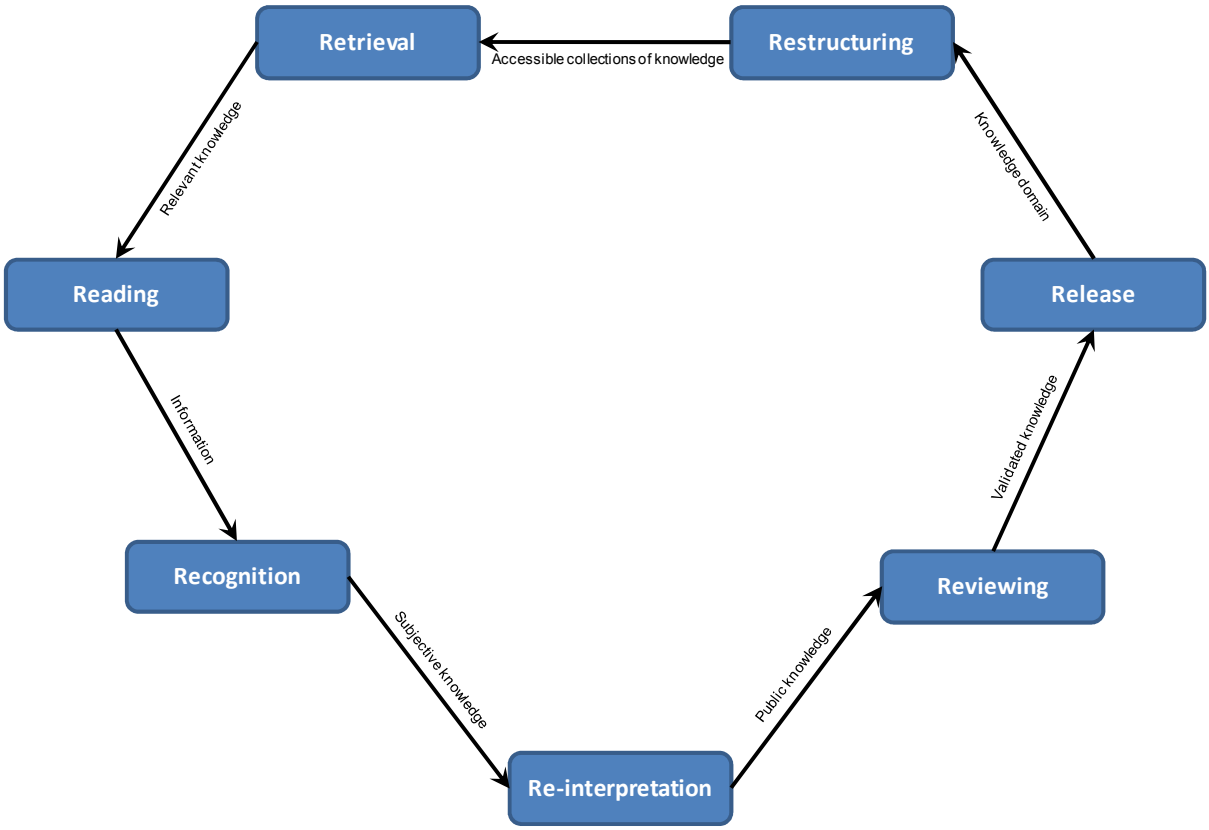


Figure 3.2: The 7 R's of Information Management (Source: Rowley & Farrow 2000)

2. 'Once read, the *Relevant knowledge* becomes *Information* which is absorbed into the cognitive framework of the individual.' 'This process of Recognition is concerned with matching the concepts in the user's cognitive framework with those in the document that is read. Recognising is concerned with converting information into *Subjective knowledge*.'

3. 'Re-interpretation is concerned with the conversion of knowledge into a form than can be easily communicated, such as in a document.' 'We describe this information as *Public knowledge*.'

4. 'Reviewing or evaluation is concerned with the conversion of public knowledge into *Validated knowledge*.'

5. 'Release or distribution is concerned with making public knowledge available within the community, organisation or market-place that might find it to be of value.' 'Once validated knowledge has been released, it enters the *Knowledge domain* upon which individuals, organisations and communities can draw.'

6. 'Organisations will interact with this knowledge domain, select items from it and collect or provide routes of access to a subject in the knowledge domain that they judge to be of specific interest in meeting their objectives.' Various processes can be executed in the 'Restructuring of knowledge to meet a specific purpose. This *Collection of knowledge* will be supplemented within organisations by

information that emerges from the collection of transaction-based data, such as sales data, within the organisation.'

7. 'This accessible collection of knowledge will then be used by individuals as a resource from which they can Retrieve *Relevant knowledge*.'
8. '*Relevant knowledge*, once retrieved, must be Read before the knowledge recorded in documents of various types can be converted into information and the cycle can recommence.'

For a better background on the topic of information management in organisations, some of the aspects regarding it will forthwith be discussed, especially

- Documents vs. transactional information (i.e. the different types of information found in the organisation);
- Information silos vs. integrated information (i.e. the accessibility of organisational information); and
- Information management tools and approaches (i.e. the application of organisational information management).

3.2.1 Document vs. transactional information

Information that resides within an organisation's systems can roughly be grouped into two main categories: documents and transactions.

A *document* (in electronic format) usually acts as a record of knowledge (Rowley & Farrow 2000), and is a complex unstructured grouping of information generated by a person needing to capture context-specific information, and communicate this information to another person. 'They are often the organising and interpretive medium that gives data, figures and other information meaning within an organisational context' (Blair 2002). Examples of internal documents are meeting memoranda, service documentation such as job schedules, progress reports, delivery notes, sales figures and projections, educational material, and administrative documents. External documents include presentations, advertising and promotional material, orders and invoices, queries, complaints, technical drawings of components and parts, price lists, product ranges, legal and safety regulations and standards (Zantout & Marir 1999). Although these types of documents all have different formats and functions, each one can sufficiently be described by a general definition supplied by Schamber (cited in Rowley & Farrow 2000), for an electronic document within an organisational context:

'A unit

- consisting of dynamic, flexible, nonlinear content,
- represented as a set of linked information items,

- stored in one or more physical media or networked sites,
- created and used by one or more individuals
- in the facilitation of some process or project.'

Organisational documents are usually stored in a myriad different ways at a myriad different locations within the organisation, ranging from documents stored in a random manner (e.g. with some or no meta-data attached) at a random location (e.g. some directory on a person's computer), to being stored in a knowledge management system, in a predefined manner (e.g. with meta-data attached) at a predefined location (e.g. team collaboration site).

A *transaction*, on the other hand, only records the smallest unit of business activity (Recordkeeping Manual - Glossary of Recordkeeping Terms 2007), and usually serves as record of human-system interactions, or system-system interactions. The fact that it is the result of system interaction usually also predisposes this information to be very structured in format, which in turn allows the information be processed in a very rapid, and in most cases, in an automated manner. Transactions are usually stored in various kinds of transaction systems, designed to process, store and retrieve information at very high speeds, and which can range from something as simple as a flat text file, to something as complex as a relational database or a decision support system.

Although very different in purpose and form, both types of information can provide context to different situations, and therefore have importance to the organisation. Documents and transactions should thus be utilised in such a way as to ensure that the most benefit are obtained from it.

3.2.2 Information silos vs. integrated information

Organisations are usually made up of different departments or divisions which mirror the organisation's value chain, or are based on some other organisational planning technique. Such a department will typically develop and configure various systems and processes within its boundaries, aimed at optimising the tasks for which it is responsible, and ensuring that it delivers what is required in the most efficient manner. A by-product of this optimisation is usually that the department defines a specific strategy with regards to managing the information that is generated and utilised within its boundaries. Even though information is passed between departments to perform certain tasks, a department is usually more focused on its internal operations than on its external interactions (mainly because of the additional effort required for unspecified information exchange (Carlson 2003)). This means that it becomes much more difficult to access the information from outside the department and that "walls" are inadvertently created around its information sources, making it difficult for other parts of the organisation to be aware of and share its information. These *information silos* cause the organisation's

information to become fractured, as transparency is reduced, which in turn causes duplication of work and sub-optimal decision-making. The silo walls basically act as information barriers that manifest itself by preventing information from being available where it is needed. (Zantout & Marir 1999). A person searching for information in an information silo environment basically have to answer two questions in order to obtain sought-after information: Firstly, does the information exist? And secondly, where is it located?

The problem of information silos is definitively not a new one, and over the years various solutions have been suggested to mitigate or reduce its effects. When studying these solutions, the main theme that always comes to the fore, is the concept of information (or data) integration, which means ensuring that all information in the organisation are integrated to be accessible from any part in the organisation. Information integration aims to ensure that no sectionalising of organisational information takes place, making it possible to have a seamless view on "what the organisation knows." Furthermore, duplication of information is greatly reduced, and the information seeker is always presented with the most current information on a certain topic. In this case, a person searching for information basically only has to answer one question in order to obtain the sought-after information, namely: Does the information exist? The location aspect of information is therefore conveniently hidden from the information seeker. Figure 3.3 shows the difference in complexity between interaction with information in an information silo environment, and interaction with information in an integrated information environment.

Information integration usually necessitates some enabling technology in order to provide users with a single comprehensive information system within the organisation, with the ability to share or aggregate information across the many functions, divisions or departments of the organisation (Goodhue, Wybo & Kirsch 1992). Currently various approaches can be taken in order to improve the integration of information in organisations, and the success of these can vary, depending on the environment and situation to which it is applied.

3.2.3 Information management tools and approaches

As human knowledge about the importance of capturing and managing information increased, various tools and approaches were developed (and are continued to be developed) in order to increase the benefit people can get from information, without increasing the effort of interacting with the information. Some of these tools and approaches will forthwith be briefly discussed in order to indicate some specific ways in which people currently manage and interact with information (each tool or approach to be discussed in terms of its domain of application – documents or transactions).

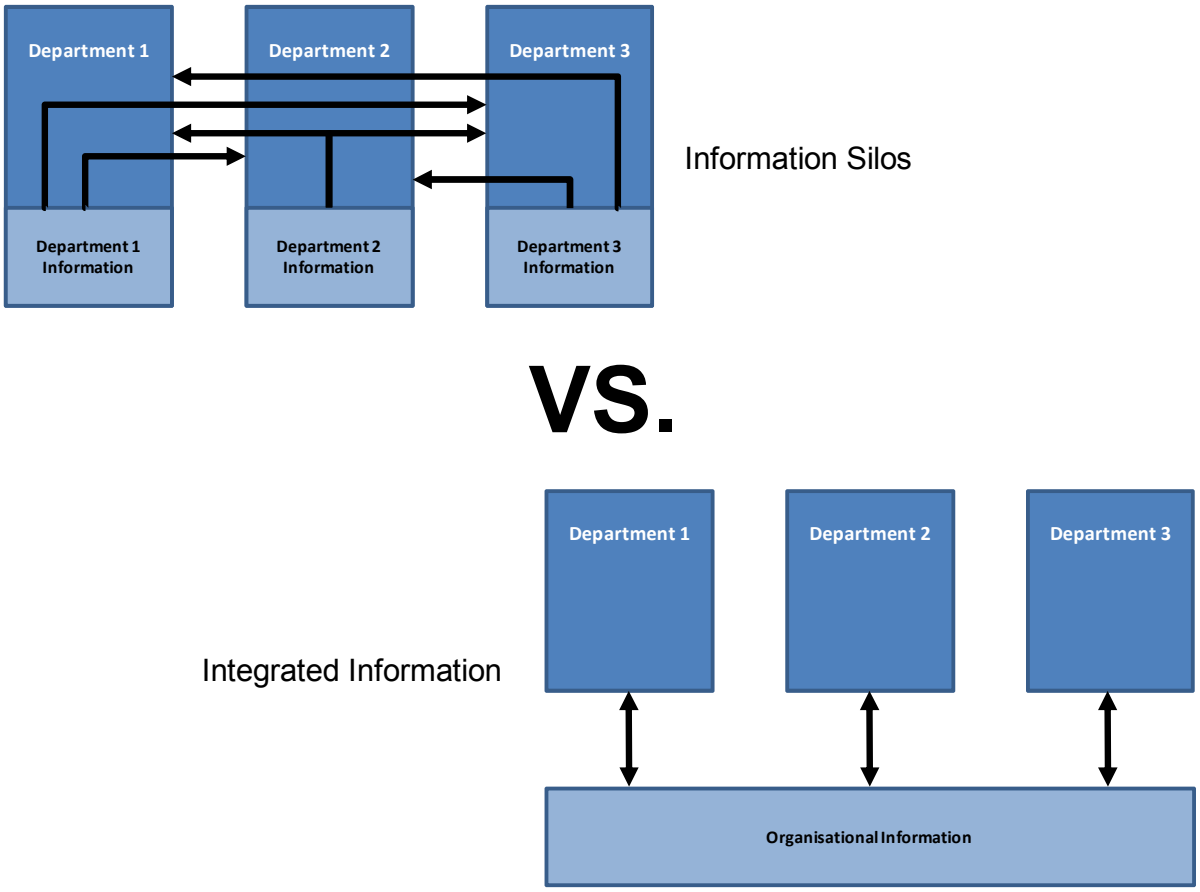


Figure 3.3: The difference in complexity between interaction with information in an information silo environment, and interaction with information in an integrated information environment

- ***Shared directories (Documents)***

Having shared directories for document storage is possibly one of the simplest and most effortless approaches for interaction with information that exists. It basically entails the creation of directories or folders (terminology used by Microsoft, Apple, and some other software development corporations) at an agreed-upon location on a computer network, and giving specific users access to use it for the storage and retrieval of electronic documentation (and therefore, information). In a highly context-specific environment this approach should work very well, as only a limited number of documents will be stored in the shared directories, and because the context is known to the users, the retrieval of information should take place without any difficulty. It is, however, not advisable to use shared directories for major information storage, as the storage of documents are done very subjectively, and very little or no assistance is usually provided for the locating and retrieval of information.

- ***Databases (Documents and transactions)***

A vast improvement on shared directories, or the notion of storing information in a random manner at an accessible location, is the concept of a database. A simple definition of a database (according to WordNet (n.d.)), is 'an organised body of related information.' A more descriptive definition for this context is provided by Wikipedia (2007b) as 'a structured collection of records or data that is stored in a computer system so that a computer program or person using a query language can consult it to answer queries.' It provides an organisation with the functionality to store information in an agreed-upon manner, but also the ability to easily locate information through search algorithms and querying. In response to different challenges, various database models have been developed over the years, with the relational and object database models currently being employed the most. Such is the versatility of databases that a large number of user software applications employ databases as a foundation to efficiently manage the information these applications generate. All of the remaining tools and approaches to be discussed in this section apply the concept of databases and are actually specialised applications of the database concept.

- ***Document management systems (Documents)***

Document management systems (DMSs) (also known as electronic document management systems (EDMS's)) are specifically targeted at providing users in the organisation with easy access to the organisation's document (i.e. unstructured information) sources, and usually employ databases to handle the transactions conducted by these systems. These systems can be as simple as being mere data storage and retrieval systems, to being advanced decision support systems which incorporate modelling, prediction and inference facilities (Zantout & Marir 1999). The core functions provided by current document management systems (according to Zantout & Marir (1999)) are enabling users to:

- directly manipulate the documents;
- index and store to retrieve the documents;
- communicate through the exchange of documents;
- collaborate around documents; and
- model and automate the flow of documents.

The functions listed above loosely denote the four most common requirements for document management systems, namely information retrieval, GroupWare, workflow and knowledge management.

Information retrieval refers to techniques employed for the storage and retrieval of documents (and therefore information) from textual databases (Zantout & Marir 1999). Information retrieval is aimed mainly at comparing the parameters of stored documents to user-supplied search requests, and presenting all positive matches to the user. Relevance is obviously of key importance, which is why the performance of information retrieval is measured in terms of two criteria: precision and recall. High precision indicates that a large percentage of documents retrieved and presented to the users are actually relevant (obviously dependent on the search request), and a high recall indicates that a large percentage of relevant documents were actually retrieved by the system.

'Reflecting the importance of collaboration, *GroupWare* is defined (Khoshafian & Buckiewicz, cited in Zantout & Marir 1999) as the multi-user software that supports individuals using computers to communicate and to carry out tasks' (Zantout & Marir 1999). Through GroupWare functionalities, users of the document management system can collectively work on documentation and collaborate on organisation tasks without having to be in the same location. Some examples of Groupware are conferencing tools, e-mail, instant messaging and web publishing.

Workflow belongs to a broader GroupWare category, where more focus is placed on the modelling of the organisation's business processes. 'The enactment of a workflow may include activation, tracking, status monitoring, messaging, queue handling and routing of documents' (Zantout & Marir 1999), and through this, management processes can be executed more effectively, and documents tracked and routed in a simpler manner. Workflow systems can be classified as one of the following:

- transactional workflow, which will be most applicable in situations where business processes are clearly identified and repetitive in nature, where the processes needs to be executed accurately and where controls can be put in place to automate the process, ensuring that it is executed in the correct manner;
- administrative workflow, which is most applicable to administrative processes performed usually in conjunction with e-mailing, and entails the completion and submission of forms in a semi-rigid manner; and
- ad-hoc workflow, in which the process steps can not necessarily be specified in advance, and the routing and tracking of documents take place in a random manner.

'Knowledge management includes all methods, instruments and tools that contribute to the promotion of an integrated core knowledge process – with the following four core activities as a minimum, to generate knowledge, to store knowledge, to distribute knowledge and to apply knowledge – in all areas and levels of the organisation in order to enhance organisational performance by focusing on the value creating business processes' (Mertins, Heisig & Vorbeck 2005). Knowledge management functionality in document management systems focus on the identification, creation, representation and distribution of explicit and tacit organisation knowledge

in order to improve reuse, awareness and learning in the organisation (Wikipedia, the free encyclopedia 2007d).

• *Data warehouses (Transactions)*

A (enterprise) *data warehouse* is typically an organisational database that contains integrated subject-oriented data captured from one or more operational systems or external information providers (Pipe 1997) and is structured in such a way as to promote easy querying and reporting of the data (Greenfield 2006) for decision making purposes. Bill Inmon, seen as the "father" of the data warehouse, defines it as follows:

'A data warehouse is an integrated, subject-oriented, time-variant, non-volatile database that provides support for decision making' (Inmon & Kelley, cited in Rob & Coronel, 2000). The four under-lined components are explained as follows (Rob & Coronel 2000):

- Integrated: The data warehouse is a centralised, consolidated database that integrates data collected from across the entire organisation, irrespective of source or format.
- Subject-oriented: The data residing in the data warehouse is usually organised and indexed around topics and subjects (Kendall & Kendall 2002) that are meaningful to users, as it attempts to provide answers to questions from the diverse functional areas in the organisation.
- Time-variant: Data in the data warehouse usually have a time component, which means that the data warehouse can give a representation of the flow of data through time, and can even perform projections.
- Non-volatile: Once data is placed in the data warehouse, it will never be removed. In this way, the history of the organisation's data is maintained, providing a more complete picture of the organisation than snapshot operational data would have done.

Figure 3.4 gives an indication of manner in which a data warehouse is created from operational data.

With the data represented in terms of topics and subjects, users can easily access and manipulate information in a data warehouse through different tools such as spreadsheets, statistical analysis programs and executive information systems to assist in effective tactical and strategic decision making (Choo 2000). Because of the tremendous presentation and manipulation capabilities of the data warehouse, it can also provide a very comprehensive foundation for business intelligence requirements in organisations (Chenoweth, Corral & Demirkan 2006), and in some cases data warehouses even provide input data to other analysis, reporting and decision making tools.

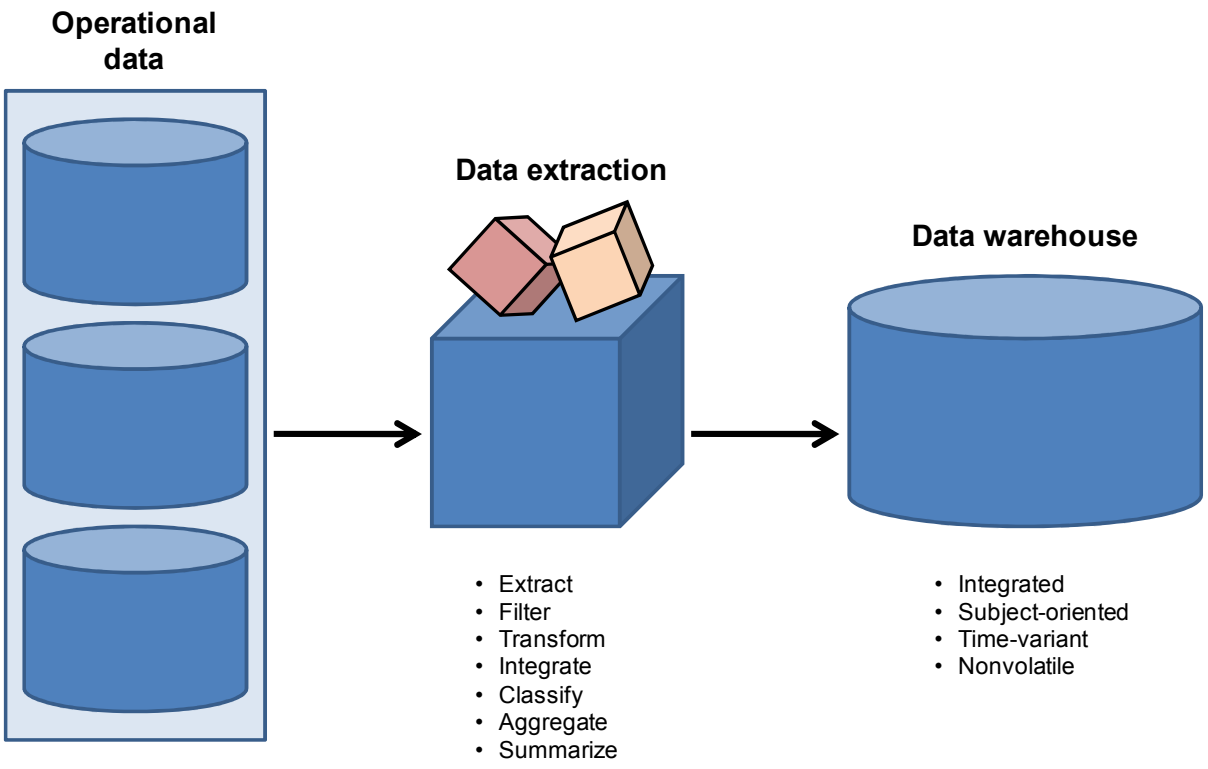


Figure 3.4: Creating a data warehouse (Source: Rob & Coronel 2000)

- ***Data marts and operational data stores (Transactions)***

A data mart is a specialised type of data warehouse that contains only a subset of corporate data that is of value to a specific department or set of users (Pipe 1997). In scope, the information it contains can be derived from an enterprise-wide database (like a data warehouse) or be more specialised (SearchOracle.com 2005) by, for instance only focusing on specific operational systems. Data marts focus on meeting the specific demands of a particular group of users in the organisation, in terms of analysis, content, presentation, and ease-of-use, and therefore users of a data mart can expect to have data presented in terms that are familiar (SearchOracle.com 2005). The smaller size of data marts makes it faster and less expensive to build than data warehouses, but this means that the capability to perform analyses on organisation data across functional area boundaries is severely inhibited (Pipe 1997).

Another specialised application of the data warehouse concept is the *operational data store*. As with data warehouses, information is also captured and integrated from the different operational systems and external information providers of the organisation, but operational data stores focus more on providing knowledge workers with current (or near-current) detailed data for regular day-to-day business querying and reporting (Pipe 1997). Its aim to provide better information for fast decision making, means thus that it contains very little or no summarised or historical data.

Although the tools and approaches discussed here are just a small subset of those that are applied in organisations, it is very clear that great diversity exists in the manner in which human interaction with information takes place. The main reason for this diversity may be that the nature, application and goals of information itself make it very difficult to develop a single solution for human interaction with, and application of, information. This is why, even with multiple information interaction strategies in place, organisations still manage to extract only a fraction of the value that exists within its information resources.

3.3 The shortcomings of organisational information utilisation

When considering the diversity of the tools and approaches currently employed by organisations to interact with information, as well as the different research and development directions in this field, it would seem that the tools and approaches that are currently being advocated and applied do not succeed in completely fulfilling the needs of organisations. A reason for this might be that current solutions do not meet the needs of organisations completely, but also that new requirements are constantly being unearthed, fuelling the need for new solutions. The truth of the matter is that organisations have difficulty in realising the full potential of information within their reach, as well as the correct manner in which to interact with it. Some of the influencing factors that might contribute to the under-utilisation of information in organisations are discussed below:

3.3.1 Environment & information identification

The environment in which an organisation navigates (and especially the amount of uncertainty the organisation has to deal with), can determine how effectively the organisation can utilise the information in its possession. If an organisation finds itself in a very volatile environment where uncertainty and predictability is very high, it will struggle to apply information in its possession to the challenges it faces, and will also experience difficulty in identifying information from its environment which it can apply to further its cause. When uncertainty in the organisation's environment is low, the effective utilisation of its information can be driven by rules and procedures, hierarchies and goals, but with high uncertainty, the organisation must either improve its utilisation by decreasing its uncertainty (i.e. move to a less uncertain environment), or it must increase its information processing capacity in an attempt to generate more applicable information (Galbraith, cited in Goodhue, Wybo & Kirsh 1992).

In order to successfully apply information to a decision, decision makers need to select information (from the external or internal environment) that is completely relevant to their context, or risk possibly arriving at the wrong conclusion. The lability and interconnectedness of the environment make it increasingly difficult however to decide the degree to which an item of information is relevant (Choo 2000), which results in the wrong information being selected very easily. Information selected incorrectly can't contribute to the organisation as it lacks the relevance that makes it useful. The information may reside within the organisation's information systems without being utilised, never to be accessed again and not contributing to the organisation in any way.

3.3.2 Information quantity

The quantity of information that an organisation processes, has a very definite impact on its ability to correctly identify the value inherent to the information. Information overload, which is the result of the amount of information in an organisation exceeding its information processing capability (Schick, Gordon & Haka, cited in Chan 2001), could cause persons performing information processing to experience difficulty in identifying relevant information, be highly selective and ignore large amounts of information, and experience difficulties in identifying the relationship between details and the overall perspective (see Chapter 2). The incorrect identification and processing of information will most likely mean that the information will be incorrectly classified and stored, and which could cause it to never be recalled when it is relevant, and always be recalled when it is not. Because organisations are so reluctant to discard any information for fear that it might be needed at a later stage (Bradley, cited in Elwood 2005), it means that (unless corrected) these incorrectly classified pieces of information will remain within the organisation's information resources, causing it to be transformed into jumbles of information with varying degrees of credibility, which the organisation then employ for decision making. Using incorrectly identified and processed information for decision making could result in the wrong course of action being taken by the organisation, with potentially catastrophic consequences (Chan 2001). In an attempt to mitigate this risk, decision makers in the organisation tend to request even more information to verify the information on which a decision is based (Katzer & Fletcher, cited in Edmunds & Morris 2000). This unfortunately further increases these persons' information load, resulting in a vicious circle where the information utilisation decreases with an increase in information quantity.

3.3.3 Information storage

The manner in which information is stored for use at a later stage directly impacts on its accessibility and therefore on its usability. The inability to locate specific information in its possession can also have severe consequences for an organisation, as the following quote clearly indicates:

'The effects of a failure of "document control" can be dramatic. A major utility company was required to shut down four nuclear reactors because of lost repair instructions (at a loss of $2m per day). The US Department of Defence estimates that half of all military accidents result from missing or inaccurate technical information. A major airline was fined $10k per take-off because of out-of-date maintenance information. A major drug company lost its entire R&D investment owing to inability to provide timely documentation (Fleischer, cited in Blair 2002).'

Although effective information storage does not directly impact on the quality side of organisational decision making, it does impact on how efficiently it is done, which is why it is imperative that information in the organisation be stored in such a manner as to be retrieved as rapidly as possible. To store information in an effective manner, a common understanding must exist that the information should be seen as objective, and not stored in a subjective manner (Rowley & Farrow 2000). In failing to do so, organisations might start finding themselves becoming sluggish and slow to act on opportunities in both operational and strategic environments.

3.3.4 Information retrieval

Together with the manner in which information is stored, the correct information retrieval strategy is also vital to insure that the right information is provided in the shortest possible time. According to Cyert and March (cited in Choo 2000), information search in organisations is problem-motivated, simple-minded and biased. This subjectivity influences the way in which search requests are provided, and might cause the wrong information to be provided. People also expect the information specialists and technology that forms part of the information retrieval process to understand the context of their information request completely, interpret it successfully, and provide the exact information they require, irrespective of the format and level of detail of their request (Choo 2000). Search engines (like those used for Internet searches), are a good example of this, which are technologically very advanced, but will yield much unpolished data when the search request is ill-defined (Chen & Tai 2003). Furthermore, within a single company there are often different identifiers for key business entities, and these inconsistencies can cause major confusion when questions are asked that span multiple systems or multiple subunits (Goodhue, Wybo & Kirsh 1992).

It is therefore clear that, unless an information request is objective, clearly defined and unambiguous, there are no guarantees that the information acquired will correspond with the information requested, and be able to satisfy the information seeker's needs.

3.3.5 Non-integrated organisational information

Although the detrimental effects of information silos are well known, this obstacle to optimal information utilisation still occurs. The manner in which the organisational units perform their daily operations, as well as the information requests processed by them, can result in processes and systems internal to the units being structured to optimise throughput, but unfortunately failing to integrate completely with the processes and systems of the enterprise as a whole.

Various drivers can contribute to information silos originating, but probably the most common driver is a manager. The inner workings of a department or business unit are usually the responsibility of its manager, who are then also measured against the performance of his/her department. In order to ensure that his/her department performs well, a manager is more likely to configure the inside workings of the department to perform optimally and risk non-conformance with organisational requirements, than to conform to the rest of the organisation and sacrifice performance. Through the use of these non-conforming processes and systems, managers and departments risk being unaware of problems similar to theirs possibly being addressed in other parts of the organisation, or that the same problems have been encountered before, and that proven solutions already exist (Choo 2000). Information silos can unfortunately also occur through selfishness on behalf of a manager. Choo (2000) states that organisations are not monolithic, but divided into separate political domains, each with its own manager-leaders, culture, language, and information resources. 'Managers recognise that their positions and their ability to get their way depend on the unique information they hold and the deployment of this information to justify courses of action' (Pfeffer, cited in Choo 2000). This amounts to a type of *information feudalism* where managers and their departments control the manner in which information is obtained, stored, distributed and analysed (Choo 2000), which then prevents the information in their possession from being utilised optimally by the organisation.

Non-integrated organisational information is however not only caused by managers. Technology and information systems are very often the root cause for information integration difficulties experienced by organisations. With the use of decentralised enterprise models for instance, different parties in organisations employ different data models, and integration is only enabled through explicitly representing the meaning of entities. Translation from one entity in one data model to a similar entity in another model is executed by applying the definitions of these entities to draw relationships between them (Kim 2000). When data models change through changes in the business and the meanings of entities and the relationships between them are not updated, the translation will be done incorrectly or

not at all, and the data models can become isolated from one another. Another example where information systems can be the cause for information integration issues in organisations is the continued use of legacy systems, in that these systems tend to become isolated if their interaction with the rest of the organisation's information systems is a problem (Goodhue, Wybo & Kirsch 1992). The integration of legacy systems would seem to be the simplest solution to this problem, however the migration of legacy systems is most often difficult and expensive and bring about problems of their own.

Other reasons for organisation information failing to be integrated (and thereby improving its utilisation) include the difficulty of obtaining top management support or the need for managing expectations for such an endeavour (Goodhue, et al.; Hoffer, et al.; Lederer & Sethi, al cited in Goodhue, Wybo & Kirsch 1992), shortcomings in the methodologies used to attain the expected level of integration (Goodhue, et al., cited in Goodhue, Wybo & Kirsch 1992) and the potential failure of data integration efforts to provide sufficient benefits to offset their costs (Goodhue, Wybo & Kirsch 1992).

3.3.6 Under-performing of information utilisation improvement solutions

It sometimes happens that technological solutions are introduced in the organisation to improve the utilisation of its information, but fail to satisfy the organisation's expectations. Possible reasons for this can be that a particular solution was not implemented according to plan, it is utilised in the wrong manner, or has a detrimental impact on the rest of the organisation's systems, causing people to rather endure the negative effects of under-utilisation of information, than using the implemented solution.

Generally speaking, organisations in different industries, with different sizes and using different applications experience performance problems when new applications are introduced, existing application are grown, or the number of application users are increased, which all put a strain on existing databases and other information infrastructures. With multiple departments using multiple reporting tools and applications to access the same database, performance problems are unavoidable (Appfluent Technology 2002). According to a study done by Appfluent Technology (2002) on reporting and business intelligence application usage, the most frequent reporting problem that affects system performance is when information is extracted from the live production systems. This puts strain on the reporting tools as well as on the production systems themselves. This then increases the extraction time, leading to a drop in productivity, and might influence the utilisation of extracted information.

Data warehouses do not have a good track record when it comes to successful implementation. Especially in the 1990s, it was found that the sheer magnitude of data warehouses were underestimated, resulting in these implementations generally taking longer than planned, overrunning

budget and under-delivering on expectations. 'And that was for the projects that actually were completed; many simply fizzled out' (Sherman 2005a). To implement a data warehouse, data must first be gathered from a variety of sources and translated into a common form. Once the data has been transformed, the data must be summarised (which means the sacrificing of some detail) into queries that might be (or have been) requested. Finally, the data in the warehouse must be organised and ordered by subject, which involves further massive amounts of analysis and design (Kendall & Kendall 2002). At this point the data warehouse exists as a set of integrated and consistent information, but for users to effectively interact with it, various reporting and support functions must also be added, which further inflates the complexity, cost and duration of the implementation (Sherman 2005b). It is therefore very obvious that the implementation of a data warehouse is a monumental task, and the possibility of missing its goals in terms of cost, time and benefit are quite significant.

When the concept of a data mart was introduced, it was thought that a workable solution has been found. Unfortunately, just like data warehouses, many projects ended up failing to deliver what was promised, while also posing a new problem: 'The trade-off for being able to deliver cheaper and faster was to severely limit scope of the breadth and depth of data, its integration, its quality and the analytical capability offered to business users. Initially data marts were perceived as a great success - until business groups realised they were debating in meetings which one of their reports (and associated data marts) had the "right" numbers. In the end, data marts left their business users wanting more – which often meant another round of information silos with more disjointed data marts or data shadow systems and more versions of the truth' (Sherman 2005a).

3.4 The improvement of organisational information utilisation

The previous section (section 3.3) certainly does paint a very bleak picture on the ways in which organisations struggle to get the full benefit from the information in their possession. Various tools and techniques have however been developed to address or avoid these problems, and in a lot of cases are employed to good effect.

In terms of extracting information from the *environment and information identification*, web-tools (which, based on their intelligence and power, can be divided into browsers, search engines, notifications, deductions and push-delivery systems (Chen & Tai 2003)) have improved a great deal in the last few years, and are constantly improving in ability to assist people in acquiring information easily and effectively (Adorf; Laine-Cruzel, et al., all cited in Chen & Tai 2003). Some Web-tools possess the ability to track information on and changes to users surfing the Internet, and can assist users in acquiring information or

documents from the Internet by using this surfing data (Chen & Tai 2003). Carlson (2003) suggested the following approaches for the effective extraction of information from both internal and external environments with relatively little trouble: Intelligent agents, ranking algorithms, cluster analysis, web mining / data mining, web graph algorithms, personalisation, recommendations, and collaborative filtering. Carlson (2003) also listed some attributes of information which can be used to locate information with much less effort: Format, scope, relation to other information, authority, treatment, arrangement and cost.

To address the impact of *information quantity*, Choo (2000) suggests that information variety must be managed to avoid the risk of information saturation, and that ways to do this are making use of specialised knowledge possessed by librarians about information resources, outsourcing the monitoring of specific issues that are of special importance and using information technology to both amplify and attenuate variety. These approaches are however more focused towards identifying applicable groups of information, which then still have to be assessed in order to identify the specific piece of information that will fulfil an information request the best. Because sampling and filtering of information are unfortunately still done best by humans, it is important to involve as many organisational members as possible in information gathering and assessment. These people must however be protected from the negative aspects associated with assessing masses of information. The tools and approaches mentioned in the previous chapter can be applied to also quite successfully counter the effects of information overload. Chan (2001) however, proposed that alternative ways of representing information might be employed to protect people from the effects of information overload. Recently, people have started adding graphs to reports in order to make it easier to read (Beattie & Jones, cited in Chan 2001). With the advent of information and graphics technologies in recent years, graphs have become very popular and are commonly adopted as inexpensive decision aids, and although there is currently no real substantive proof, there is a belief that graphs may be a possible solution to prevent information overload (Chan 2001).

Effective *information storage* basically describes the structuring and organising of knowledge for two main purposes: understanding and cognition, and the finding of information or knowledge (Rowley & Farrow 2000). Organisations have started to realise that information collected from its employees and processes can be turned into value (Conway & Sligar 2002), and various initiatives have been successfully introduced through which information is stored in an effective manner. Examples of such initiatives can include simple solutions such as the storage of information in specifically defined locations or the keeping of metadata on information, but can also be complex solutions such as document management systems and the application of record management strategies, through which automated processes control the manner in which people interact with information, and the consistency of information is ensured. These initiatives ensure that an organisation's information resources are mapped, which allows the organisation to fully harness the value of its information through simple retrieval and by extension, improved utilisation (Conway & Sligar 2002).

Effective *information retrieval* is largely addressed by solutions that are discussed in the paragraphs on environment and information identification, and information quantity. Solutions that are employed for information retrieval can basically be divided into push and pull solutions. Pull solutions are focused on retrieving organisational information based on an explicit request by a user, whereas push solutions attempt to automatically provide a user with information without specifically being asked for it, based on certain criteria provided beforehand by the user. In order for information delivery to mesh well with the user information habits and preferences (Choo 2000), artificial intelligence capabilities of information retrieval applications are continuously being improved, aimed at analysing the user's information utilisation behaviour, and providing better context-specific retrieval results. Furthermore, people are becoming increasingly adept at providing exact specifications of the information they require, which have an obvious positive impact on the correctness of retrieved information.

An increasing number of organisations are starting to experience the benefit of seeking and implementing solutions against *non-integrated organisational information*. Jack Welch Jr., CEO of General Electric, had the following to say on the topic: 'Picture a building. Companies all added floors as they got bigger. Size adds floors. Complexity adds walls. We all built departments – transportation departments, research departments. That's complexity. That's walls. The job all of us have in business is to flatten the building and break down the walls. If we do that, we will be getting more people coming up with more ideas for the action items that a business needs to work with '(Conway & Sligar 2002). Goodhue, et al. (cited in Goodhue, Wybo & Kirsch 1992) mentioned two different impacts information integration can have on organisations: (1) improved managerial information for organisation-wide communication and (2) operational coordination between interdependent parts of the organisation. In terms of the costs and benefits for the information systems, Goodhue, Wybo and Kirsch (1992) listed the following benefits of information integration: (1) increased ability to share information to address subunit interdependence, (2) reduced ability to meet unique subunit information requirements, and (3) changes in the costs of information system design and implementation. Organisations are very much aware of these benefits that integrated information can have for them, and today most organisations have strategies in place to ensure that information is shared between all departments.

The *under-performing of information utilisation improvement solutions* have caused a great many organisations to adopt a very cynical view towards tools and approaches aimed at supposedly improving their utilisation of information. From their point of view, the potential benefit of these tools and approaches are not worth the effort required, especially if the solution has got substantial cost implications. Failures are however invaluable to refine tools and approaches, and also allow an organisation to learn a great deal about itself, the information in its possession, as well as the manner in which it utilises its information. The data warehouse and data mart concepts are a case in point. Failures experienced first in the implementation of data warehouses, and later in the implementation of data marts, caused great resistance in organisations towards such information integration initiatives, and resulted in a serious re-think of both concepts. The benefits and (especially) the limitations of both

concepts were identified, which led to new approaches being developed where the solution entailed the implementation of a combination of data warehouses and data marts which much more significantly improved the utilisation of organisational information. The success rate of implementations improved and organisations have started employing the improved solution with greater success.

3.5 Conclusion – Understanding information

The fact that existing tools and approaches aimed at improving organisational information utilisation only achieve partial success is possibly an indication that the nature of information and its application in organisations are not as clearly understood as one would think. It may be useful at this stage (now that the utilisation of information within organisations are better understood), to take a step back, consider a piece of information like a document or a record, and try to understand how the value offering of such a piece of information can be beneficial to a person or an organisation.

A piece of information can essentially be considered as a collection of data points which give it its value. These data points are typically the names of people, places, products, organisations, dates, numerical values, and so forth. A data point on its own however, does not constitute a meaning, or does not convey any message (except the fact that it is there). Data points are therefore related to one another, and through these relations meaning and purpose are manifested in the information. A document (or any other piece of information) can therefore be explicitly described by a network of data points and the relations that exist between them.

It is to be expected that units of information (e.g. documents, etc.) that exist within the same environment and maybe share the same context, will also share data points and relations. This overlapping quality can be used to construct a single network that describes the information residing in an information store or department, or even describe an entire organisation. Within the organisation, such a network can act as a common language for communication, as well as very effectively facilitate information integration (Goodhue, Wybo & Kirsch 1992). A representation of information as networks of data points and the relations that exist between them, as well as the single network that results from combining such networks, can be seen in Figure 3.5.

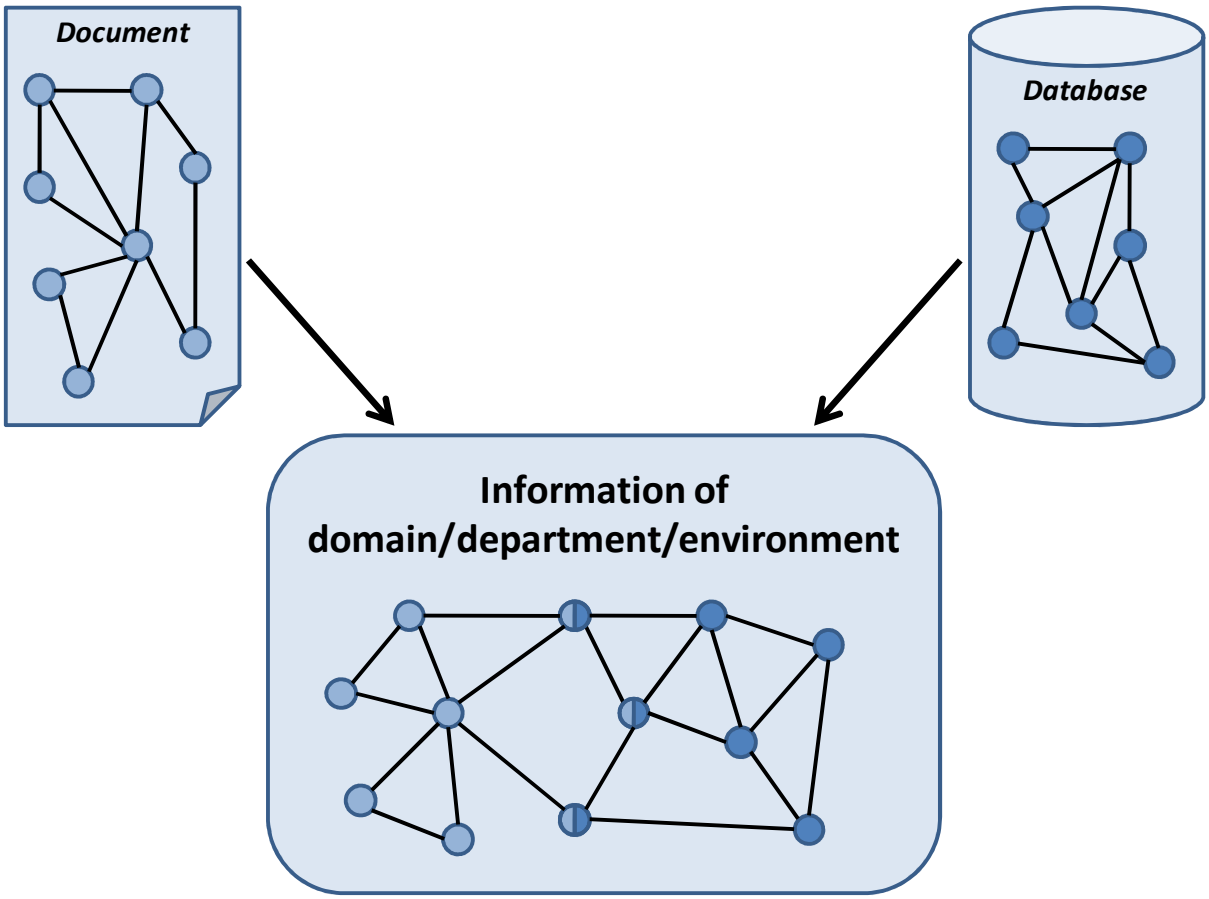


Figure 3.5: Information represented as networks of data points and the relations that exist between them, as well as the single network that results from combining the various individual networks

The idea of modelling the data points (or entities) of an organisation and the relations between them, have already existed for some time. Already in 1976, Chen (1976) have defined the entity-relationship model as a proposed means of conceptually modelling all of an organisation's information, which also provided a basis for integrated information in the organisation (Goodhue, Wybo & Kirsch). It is however currently used more by database and systems developers in the design of organisational information systems.

It can naturally be asked whether normal information users in the organisation can also benefit from having an entity-relational view on information, and navigating these networks (possibly even through visual browsing (Conway & Sligar 2002)) to acquire certain information. This will be further discussed in the following chapters.

Chapter 4 – Information Modelling

4.1 Introduction

The easiest way to find information in a book is usually by utilising its index (typically located at the back of the book). Pepper (2002) makes the statement that an index acts as a map of the knowledge contained within the book, listing the topics covered by whatever name users might be expected to want to reference them by, together with salient (and only salient) references (i.e. the associated page numbers) to those topics. Obviously, this simple technique can greatly reduce the amount of time spent searching for information, and manages to replicate in miniature the structure of its subject, in order to provide a more manageable view of the whole (Pepper 2002). Indexes are applied very successfully to individual printed documents, but are unfortunately much less successful in the electronic information domain. The reason for this is that, instead of only being applied to individual documents, indexes for electronic information are required to span multiple documents and even to cover vast pools of information, which therefore call for individual, content-specific indexes to be merged (Pepper 2002).

The merging of indexes is, however, not the problem. The problem is that indexes on their own lack discrimination, and struggle to contend with situations where (for example) the same subject may be referred to by multiple names (the "synonym problem"), or that the same name may refer to multiple subjects (the "homonym problem") (Pepper 2002). Indexes therefore lack the ability to convey context, and to create a common understanding, especially in the case of combined indexes.

Both semantics and information structuring techniques are currently enjoying an immense amount of attention in the information management research domain, as both techniques can be quite successfully applied to ensure that a common understanding is maintained when different groups of information are merged. Semantics, which is the study of language meaning, helps to bridge nomenclature and terminological inconsistencies, clarifying the underlying meaning of a piece of information (Schreiber 2003). Information structuring on the other hand, entails representing information in terms of the entities (or topics) contained within it, together with the relations that exist between these entities. Such an entity-relation(ship) view of information is very useful in creating common understanding, as well as providing a means to more easily merge different groups of information.

Semantics and information structuring have been found to complement each other very well, resulting in them currently being implemented mostly as a combined solution when merging information, or attempting to achieve common understanding. Although semantics play a significant role in the effective modelling of information, the focus of this document is however aimed more at investigating the

structuring of information, and therefore, for the remainder of this chapter, the focus will be shifted to this aspect of information modelling. This chapter focuses on discussing a number of existing information structuring approaches (see Figure 4.1) which were designed to improve the interaction and utilisation of information, and which are currently employed in a large variety of different environments and contexts.

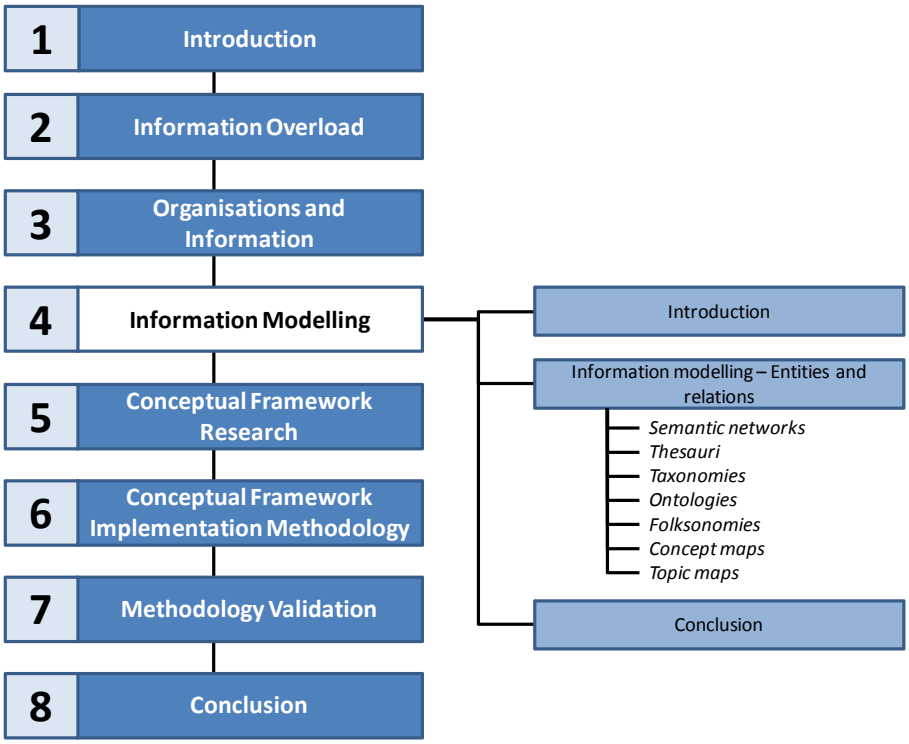


Figure 4.1: Image for navigating thesis and chapter 4

4.2 *Information modelling - Entities and relations*

The concept of representing information in terms of entities and relations is definitely not new. The entity-relationship model was defined already in 1976 by Peter Chen of the Massachusetts Institute of Technology, from which he derived the entity-relationship diagram (ERD), a technique which is very well known to database designers especially, but probably also to most other types of information systems designers and engineers (as an example, see Figure 4.2). The model was specified to be used as a 'basis for a unified view of data' (Chen 1976), and combined the main features and advantages of three other major data models of the time: the network model (Bachman; Bachman; CODASYL, all cited in Chen 1976), the relational model (Codd, cited in Chen 1976) and the entity set model (Senko et.al., cited in Chen 1976). Chen specified that, in the study of a data model, the levels of logical views of data with which the model is concerned should be identified. Four levels of data views were defined:

1. 'Information concerning entities and relationships which exist in our minds.

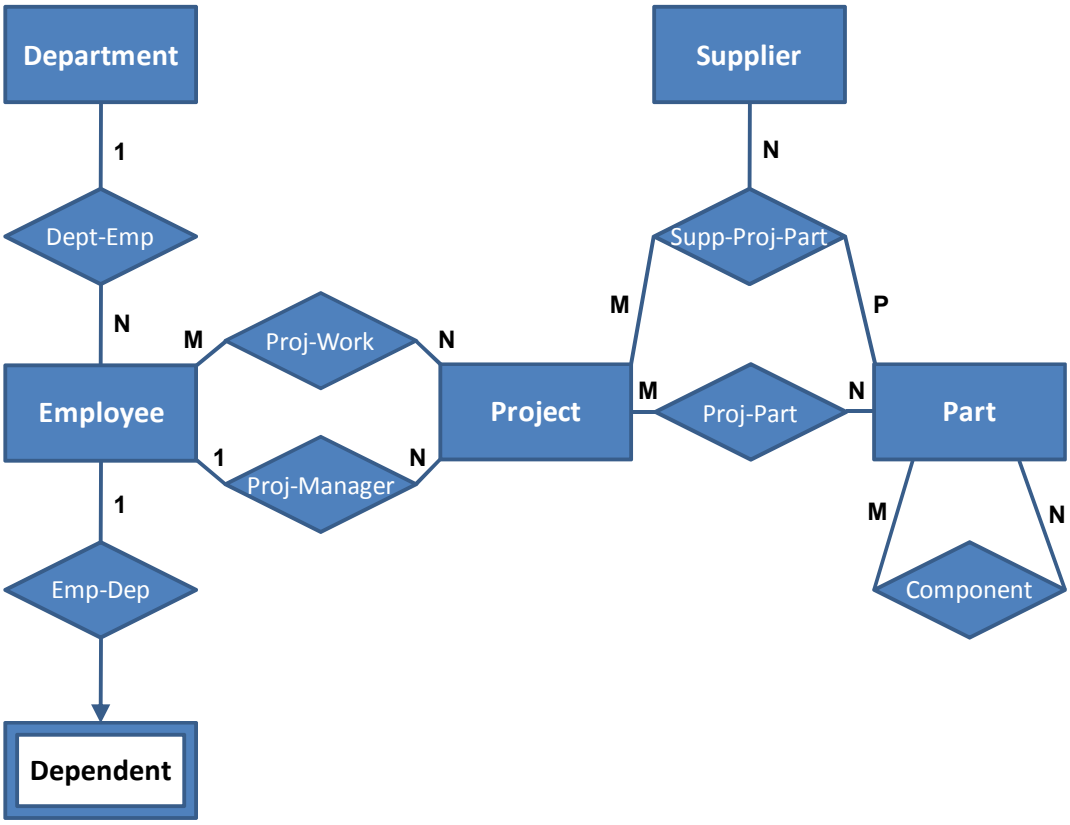


Figure 4.2: An example of an entity-relationship diagram for analysis of information in a manufacturing firm (Source: Chen 1976)

2. Information structure – organisation of information in which entities and relationships are represented by data.
3. Access-path-independent data structure – the data structures which are not involved with search schemes, indexing schemes, etc.
4. Access-path-dependent data structure.' (Chen 1976)

It can be seen that in level 1, Chen proposed the use of entities and relationships as the basis for information modelling, which he defined very concisely as follows:

'An *entity* is a "thing" which can be distinctly identified. A specific person, company, or event is an example of an entity. A *relationship* is an association among entities. For instance, "father-son" is a relationship between two "person" entities.'

These definitions are indeed very rudimentary, but through this understanding of information, Chen managed to define an accepted approach through which the manner in which information exists in peoples' minds could be modelled, so that a common understanding of information could be reached. Not too long after that, in 1982, William McCarthy also defined the REA Accounting Model (so named after its primary components which consist of sets representing economic *r*esources, economic *e*vents and

economic *a*gents), which applied Chen's ideas concerning the representation of information in terms of entities and relationships in the world of accounting (McCarthy 1982).

Today, many approaches and techniques exist through which information is modelled in terms of entities and relations[5]. This is a very clear indication of the validity of Chen's model for addressing information issues. The remainder of this section will concisely discuss some of these approaches and techniques and the manner in which they are applied to specific issues.

4.2.1 Semantic networks

- ***Concept***

Several approaches that are based on the notion of semantics currently exist in the fields of information sciences and informatics. According to Morris (cited in Mertins, Heisig & Vorbeck 2005), semantics are concerned with 'the relationship between the sign and the signified (an object or state of affairs the sign refers to), i.e. with the meaning of signs.' In the information sciences, semantic networks were first introduced in the field of artificial intelligence, and originated from research done in cognitive psychology, which is a branch of psychology dealing with processes of consciousness, and particularly with those of perception (Mertins, Heisig & Vorbeck 2005).

A semantic network (or net) is a graphic notation for representing knowledge in patterns of interconnected nodes (which represent concepts) and arcs (which represent semantic relations between the concepts) (Wikipedia, the free encyclopedia 2007e; Sowa 2006). Some important semantic relations are:

- Meronymy (the semantic relation that holds between a part and the whole, e.g. A is part of B, i.e. B has A as a part of itself) (Dictionary.com 2008e);
- Holonymy (the semantic relation that holds between a whole and its parts, e.g. B is part of A, i.e. A has B as a part of itself) (Dictionary.com 2008b);
- Hyponymy (or troponymy) (the semantic relation of being subordinate or belonging to a lower rank or class, e.g. A is subordinate of B; A is kind of B) (Dictionary.com 2008d);

[5] *Chen's Entity-Relationship model (Chen 1976) refers to the association among entities as "relationships". When regarding the term "relationship" in a linguistic sense, it can be seen that the term refers solely to a relation between people. Because information does not necessarily pertain to people alone, the term "relation" seems better suited for the role, and will therefore rather be used in this document, except where the work of Chen might be pertinently discussed.*

- Hypernymy (the semantic relation of being superordinate or belonging to a higher rank or class , e.g. A is superordinate of B) (Dictionary.com 2008c);
- Synonymy (the semantic relation that holds between two words that can (in a given context) express the same meaning, e.g. A denotes the same as B) (Dictionary.com 2008f); and
- Antonymy (the semantic relation that holds between two words that can (in a given context) express opposite meanings, e.g. A denotes the opposite of B) (Dictionary.com 2008a).

Examples of semantic networks can be seen in Figure 4.3 and Figure 4.4.

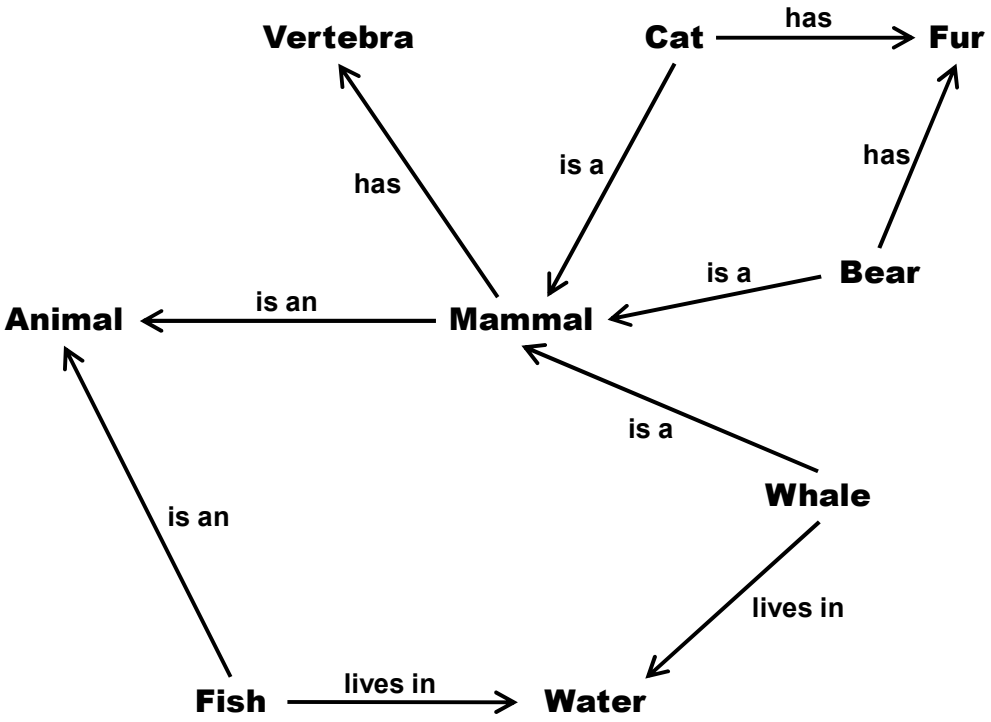


Figure 4.3: An example of a semantic network (Source: Wikipedia, the free encyclopedia 2007e)

• *Application*

Semantic networks are very widely applied, and also serve as the basis for a large number of other information and knowledge representation techniques. Sowa (2006) have defined the six most common kinds of semantic networks as follows:

1. Definitional networks emphasise the subtype or is-a relation between a concept type and its defined subtype. The resulting network, also called a generalisation or subsumption hierarchy, supports inheritance, which means that properties defined for a type is also applied to its children (or subtypes). Definitional networks are typically constructed from

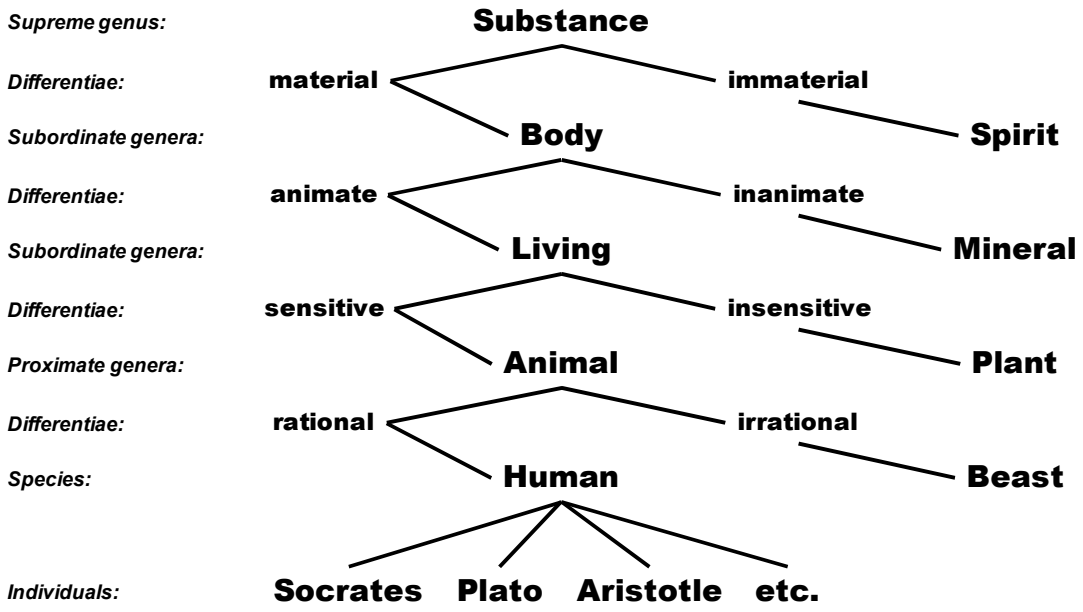


Figure 4.4: An example of a definitional network - Tree of Porphyry, as drawn by Peter of Spain (1329) (Source: Sowa 2006)

definitions, which imply that the information contained in these networks is assumed to be true.

2. Assertional networks are designed to assert propositions. Unlike definitional networks, information in an assertional network is assumed to be contingently true, unless explicitly marked otherwise. Some assertional networks have been proposed as models of the conceptual structures underlying natural language semantics.

3. Implicational networks use implication as the primary relation for connecting nodes. They may be used to represent patterns of beliefs, causality, or inferences.

4. Executable networks include some mechanism, such as marker passing or attached procedures, which can perform inferences, pass messages, or search for patterns and associations.

5. Learning networks build or extend their representations by acquiring knowledge from examples. The new knowledge may change the old network by adding and deleting nodes and arcs or by modifying numerical values, called weights, associated with the nodes and arcs.

6. Hybrid networks combine two or more of the previous techniques, either in a single network or in separate, but closely interacting networks.

Some further approaches applied in the process of knowledge structuring, and which draws on the theory of semantic networks include mind mapping[6], thesauri, taxonomies, topic maps, ontologies as well as process models.

4.2.2 Thesauri

- ***Concept***

The Oxford English Dictionary defines a thesaurus as 'a "treasury" or "warehouse" of knowledge, as a dictionary, encyclopaedia or the like' or 'a collection of concepts or words arranged according to sense; also a dictionary of synonyms and antonyms' (Gilchrist 2003). It can be seen as an expansion of the semantic network concept, because it not only maps the semantic relations between concepts, but also allows for hierarchical, associative and equivalence relations (ARMS 2006). It works on the basis of a structured collection of concepts and their designations (Westphal, cited in Mertins, Heisig & Vorbeck 2005), and basically constitutes a controlled vocabulary through which information retrieval can happen much more efficiently (Mertins, Heisig & Vorbeck 2005). Given a particular term, a thesaurus will indicate which other terms mean the same, which terms denote a broader category of the same kind of thing, which denote a narrower category, and which are related in some other way. It therefore does not only indicate that two terms are related, but also the reason and manner of the relation (Pepper 2002). The following are the most common symbols employed in the construction of thesauri to indicate semantic relations (Rowley & Farrow 2000; Cann 1997):

- SN – Scope Note
- USE – Use (indicates a cross reference from a natural language term to a preferred term)
- UF – Used For (indicates a cross reference from a preferred term to a natural language term)
- BT – Broader Term (indicates a broader term)
- NT – Narrower Term (indicates a narrower term)
- RT – Related Term (indicates related terms)

[6] *A mind map can be considered a very free form variant of semantic networks, in which the use of colours and pictures are aimed at stimulating human creativity. A major difference between mind maps and semantic networks however, is that the structure of a mind map is hierarchical, with nodes propagating from a centre and sub-nodes propagating from nodes, whereas semantic networks, where any node can be connected to any node, have a non-hierarchical structure (Wikipedia, the free encyclopedia 2007e).*

Examples of the manner in which these relations are employed are shown below (from the Thesaurus of Engineering and Scientific Terms) (Cann 1997):

- Excavating machinery

  Use Excavating equipment

- Evaporative cooling

  BT Cooling

  NT Film cooling

  RT Cooling systems

  Cooling towers

- Fixed investment

  UF Capital investment

More examples of thesauri can be seen in Figure 4.5 and Table 4.1 where two internet thesauri were queried for the term “intelligent”. Both examples provide related terms, synonyms and antonyms of “intelligent” The example from Free Dictionary (2007) furthermore also provides different meanings of the term “intelligent”.

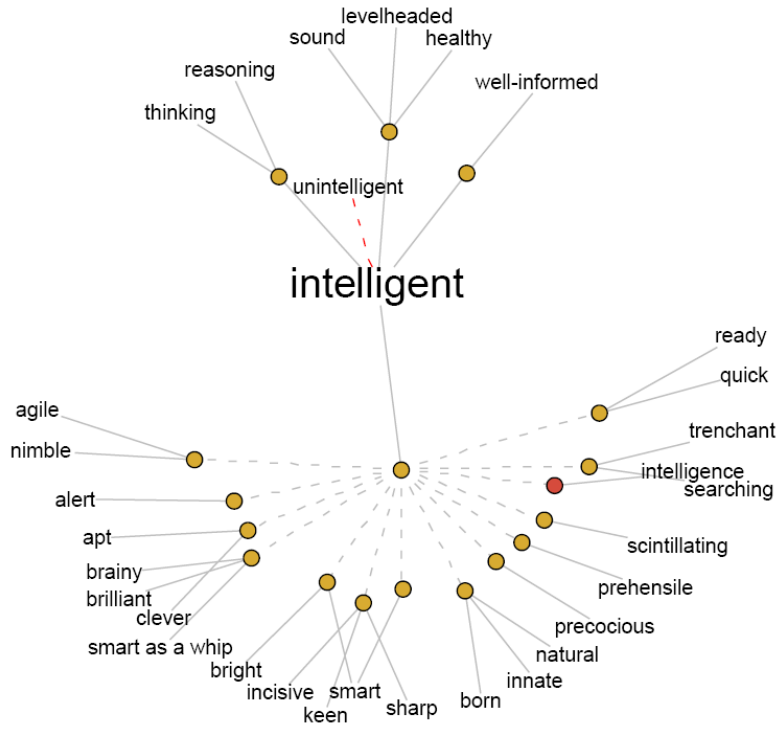


Figure 4.5: Another example of Internet thesaurus result for the term "intelligent" (Source: Thinkmap Visual Thesaurus 2007)

Table 4.1: Example of Internet thesaurus result for the term "intelligent" (Source: Free Dictionary 2007)

Adj.	**1.**	**intelligent** - having the capacity for thought and reason especially to a high degree; "is there intelligent life in the universe?"; "an intelligent question"	
		Related Word	smart - showing mental alertness and calculation and resourcefulness
		Related Word	precocious - characterised by or characteristic of exceptionally early development or maturity (especially in mental aptitude); "a precocious child"; "a precocious achievement"
		Antonym	unintelligent, stupid - lacking intelligence; "a dull job with lazy and unintelligent co-workers"
	2.	**intelligent** - possessing sound knowledge; "well-informed readers"	
		Synonym	well-informed
		Related Word	sophisticated - having or appealing to those having worldly knowledge and refinement and savoir-faire; "sophisticated young socialites"; "a sophisticated audience"; "a sophisticated lifestyle"; "a sophisticated book"
	3.	**intelligent** - exercising or showing good judgment; "healthy scepticism"; "a healthy fear of rattlesnakes"; "the healthy attitude of French laws"; "healthy relations between labour and management"; "an intelligent solution"; "a sound approach to the problem"; "sound advice"; "no sound explanation for his decision"	
		Synonym	level-headed, levelheaded, sound, healthy
		Related Word	reasonable, sensible - showing reason or sound judgment; "a sensible choice"; "a sensible person"
	4.	**intelligent** - endowed with the capacity to reason	
		Synonym	reasoning, thinking
		Related Word	rational - consistent with or based on or using reason; "rational behaviour"; "a process of rational inference"; "rational thought"

- ***Application***

Peter Roget produced the first thesaurus in 1852, which was the *Thesaurus of English Words and Phrases*. Roget referred to the thesaurus as being a “verbal classification” (Gilchrist 2003) in which entries were arranged to facilitate the expression of ideas and to assist in literary composition (Oakes n.d.). This function of a thesaurus is still mostly attributed to it today.

When considering the application of thesauri in the structuring of information however, instead of helping to produce diversity of vocabulary in text, it is tasked to the standardising of vocabulary, which in turn expedites the retrieval of information. Foskett (cited in Oakes n.d.), listed the main uses of thesauri for information retrieval as follows:

- Thesauri provide a map of a given field of knowledge, showing concepts and relations.
- Thesauri provide a standard vocabulary for consistent indexing.
- Thesauri assist users with locating terms for proper query formulation. (An example of a query would be the words chosen as input to a web search engine).

- Thesauri help ensure that only one term from a synonym set is used for indexing and searching: otherwise a searcher who uses one synonym and retrieves some useful documents may think the correct term has been used and the search has been exhaustive, without knowing that there are other useful documents under other synonyms.
- Thesauri provide classified hierarchies for broadening or narrowing a search (selecting query terms which are broader or narrower in meaning) if too many or too few documents are retrieved at the first attempt.

4.2.3 Taxonomies

- ***Concept***

A taxonomy is a classification of information according to a predetermined system, used to provide a conceptual framework for discussion, analysis or information retrieval (Conway & Sligar 2002). The classification is performed through the organising of concepts within a hierarchy, and in this case, the hierarchy only communicates class and authority, that is, an "is-a" association (Mertins, Heisig & Vorbeck 2005). Each concept within the taxonomy has at least one parent-child relation to other concepts in the taxonomy. Different types of parent-child relations may exist within the taxonomy, but it is good practice for all the children concepts of a parent concept to be of the same type (What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model? 2003).

Poly-hierarchy is allowed in some taxonomies, which means that a concept may have multiple parents. Because the concepts found within a taxonomy are expected to be unequivocal and occur only once within the taxonomy (Mertins, Heisig & Vorbeck 2005), it means that, even though a concept appears in multiple places in a taxonomy, it will still be the same concept. Likewise, if a concept has children in one place in a taxonomy, then it has the same children in every other place where it appears (What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model? 2003).

- ***Application***

Taxonomies are frequently used (and well known) in the field of biology, as it provides an excellent means for the classification of living organisms (for an example, see the taxonomy of a hippopotamus shown in Figure 4.6).

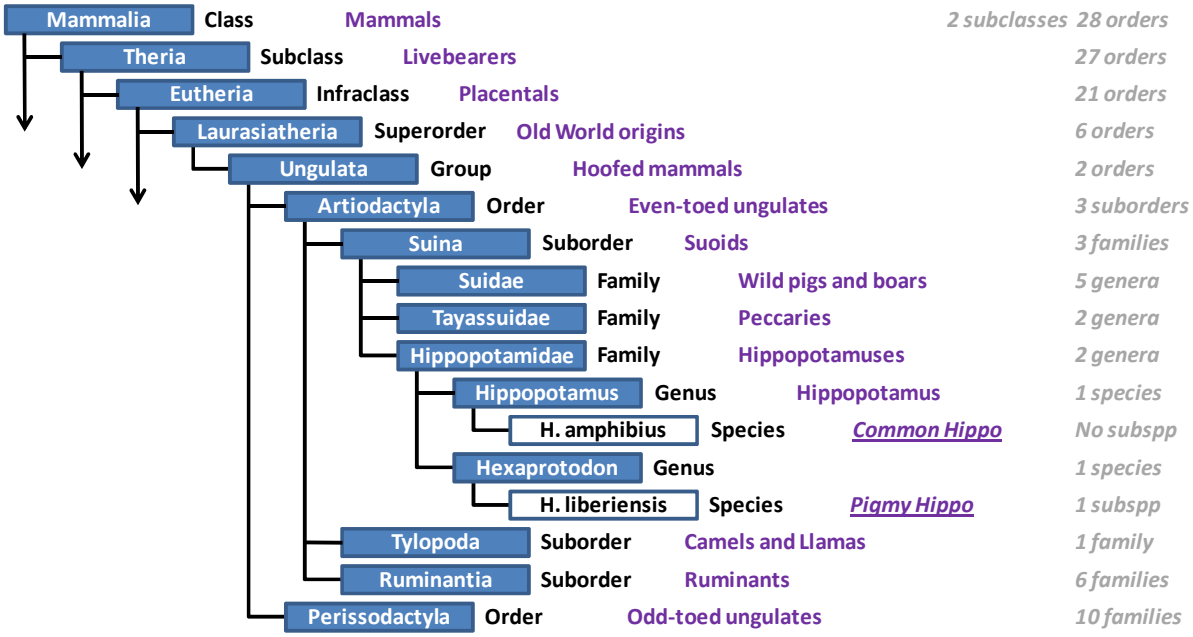


Figure 4.6: An example of a taxonomy - Taxonomy of the hippopotamus (Source: Hippopotamus Taxonomy n.d.)

This however does not mean that the taxonomy approach is exclusive to the biological domain. The most widely used taxonomy in the world, the Dewey Decimal System, is well known in library science (Wyllie, cited in Cheung, Lee & Wang 2005), and according to Mertins, Heisig and Vorbeck (2005), increased use of taxonomies have also been noted in the more abstract branches of economics and science. Gilchrist (2003) have however noted that, through differences in the understanding of what a taxonomy is, a large number of techniques and applications were being implemented in large organisations as taxonomies. From a number of case studies, he found the five most common impressions of what taxonomies are to be, as follows:

1. Web directories: Web directories are commonly used on the Internet, and are increasingly being utilised in intranets. A menu of top terms is presented to the user. Clicking on a term will display a second level, and so on for several more levels, until the user finally arrives at a reference or a link concerning the sought-after information. Web directories are definitely a form of classification, but each level does not have to be hierarchical in the accepted sense. Terms may therefore be repeated at different levels (contrary to the requirements of a taxonomy), providing alternative pathways for the searcher.

2. Taxonomies to support automatic indexing: In this case, a classification of terms exists, with an algorithm associated with each term, comprising sets of words and phrases, synonyms and syntactic variations, weights and instructions. These algorithms then act as a rules base which is used for the automatic extraction of appropriate indexing terms from documents where manual indexing would not be economically viable. These structures can however rather be considered as enriched thesauruses than as actual taxonomies.

3. Taxonomies created by automatic categorisation: Currently there exist software solutions which analyse the text in documents, and create categories based on that analysis. In turn, these categories are then used to classify the analysed documents. A network of related terms is the result, which are usually displayed in a format very similar to that of Web Directories, or two-dimensional maps. As is the case with Web directories, the result of the automatic categorisation process is merely classification, usually based on a statistical analysis of the occurrence and co-occurrence of terms in documents. There is therefore no specific presence of any form of hierarchy, which is a prerequisite for taxonomies.

4. Front end filters: Here, taxonomies are created or imported, and employed for query formulation. It functions as a tool for users to correctly select the most appropriate terms for an information query. Aspects, such as the disambiguation of homonyms, grouping of synonyms, and relations between terms are all presented to the information searcher, who compiles a query from these terms, and submits it to the Internet or intranet. This taxonomy is however still only a thesaurus formatted to enable easy navigation.

5. Corporate taxonomies: In order to stay agile and efficient, organisations are required to provide their staff with easy access to their information sources. A map and user guidelines to the organisation's repositories and their content is therefore often required, and compiled through the deceptively simple-sounding process of "mapping", which usually involves an immense amount of human effort. The result is a corporate taxonomy, which can be seen as a thesaurus/ontology[7] hybrid. This type of taxonomy is usually compiled manually, and provides a high level map on the organisation, guiding staff to whatever organisational information they require. It can contain information about the organisation itself: its processes; methods, guidelines and standards; about the people within and associated with the organisation; and also about the content of various information repositories. In this case there is no clear indication of classification taking place, and any hierarchy construction might happen possibly only by accident.

Although none of these approaches are pure applications of the taxonomy concept, it gives a good indication of the manner in which taxonomies can add value to an organisation. To get a clearer picture of the different manners in which the concept can benefit an organisation, one can consider the three generic types of organisational taxonomies discussed by Conway and Sligar (2002):

- Descriptive Taxonomies: This type of taxonomy is aimed at supporting information retrieval through searching. A core set of controlled vocabularies are developed and maintained, which an organisation can use to consistently label or tag its content with descriptive metadata selected from these authorised vocabularies. Furthermore, these vocabularies can

[7] *Ontologies will be discussed in the next section.*

be used to capture employee terminology as well, which can be mapped to the organisation's preferred terms. At search time, the association between the user-specified terminology and the organisation's preferred terms can then be utilised for more precise searching, or the user's search term can be expanded to also include other variants of the supplied term, resulting in a broader search of the information.

- Navigational Taxonomies: This second type of taxonomy is aimed more at discovering information through browsing. Again a controlled vocabulary is provided by this taxonomy, but rather than being used for the compiling of search queries, the taxonomy is displayed to employees to help them find specific information. The navigational taxonomy consists of labels which describe categories of content, based on the employees' view on how the information is organised. Because of this, category labels may be organised differently from the context-based descriptive taxonomy, and may also contain words or phrases that would not necessarily be found within a descriptive taxonomy.
- Data Management Vocabularies: This type of taxonomy is basically a short list of authorised terms, used to support business transactions. It has, until recently, been used mainly for content management, as it provided the perfect means to collect and categorise aspects of sales activities like contact names, contact types (managers, decision-makers, etc.), names of organisation's being worked with according to certain descriptors, geographical areas, market types, etc. It aims to provide a consistent collection of information to a particular group of people for utilisation in their daily activities, avoiding the duplication of effort and thus saving time and money.

4.2.4 Ontologies

- ***Concept***

The term "ontology" is well known in philosophy, and serves to express 'the metaphysical study of the nature of being and existence' (Princeton University, cited in Kishore, Sharman & Ramesh 2004a). It is perhaps more clearly defined as 'the science of what is, of the kinds and structures of objects, properties, events, processes, and relations in every area of reality' (Smith, cited in Kishore, Sharman & Ramesh 2004a), and the question an ontology basically asks, is: What kinds of things exist or can exist in the world, and what manner of relations can those things have to each other (Shirky 2005)?

This concept was found to be very applicable to the information systems discipline. Information on the various entities, relations, constraints and processes of a particular domain usually reside within

such systems, and information systems professionals and researchers are usually tasked with the efficient identification, capturing and representation of this information, based on the information seekers' requests (Kishore, Sharman & Ramesh 2004a). Ontologies are therefore very well adapted to provide a mechanism for specifying information within a certain domain, in the same way as it would describe the world. When considering the definitions of ontologies put forward by Schreiber (2003), which is 'a study of what exists' or 'a formal specification of part of the real world', these definitions can very easily be changed to 'a study of what information exists' or 'a formal specification of part of the information world', to better fit the information systems context.

Maybe the best-known definition of an ontology in the information systems context was proposed by Thomas R. Gruber (1993): 'An ontology is an explicit specification of a conceptualisation.' This is often confused with epistemology, which is the philosophical theory of knowledge (Gruber n.d.). Zhong (2002) and Gilchrist (2003) both referred to Gruber's definition as being the following: 'An ontology can be defined as a formal, explicit specification of a shared conceptualisation', which they continued to explain as follows:

'"Conceptualisation" refers to an abstract model of some phenomenon in the world by having identified the relevant concepts of that phenomenon. "Explicit" means that the type of concepts used, and the constraints on their use are explicitly defined. "Formal" refers to the fact that the ontology should be machine-readable. "Shared" reflects the notion than an ontology captures consensual knowledge, that is, it is not private to some individual, but accepted by a group. An ontology describes the subject matter using the notions of concepts, instances, relations, functions and axioms. Concepts in the ontology are organised in taxonomies through which inheritance mechanisms can be applied (Benjamins et al., cited in Gilchrist 2003).'

An ontology is therefore an abstract representation of a real world phenomenon (Breu & Ding 2004), and is made up of an abstraction hierarchy of concepts (also called the universe of discourse[8] (Gruber 1993)) which, together with the describable semantic relations (Rubin et al. 2002) existing among them, constitute a representational vocabulary aimed at describing the phenomenon (Carr et al. 2001). It further implies commitment by a set of agents to use the vocabulary (through querying and making assertions) in a way that is consistent (although not necessarily complete) with respect to the theory specified by the ontology, which actually guarantees the consistency of the ontology (Gruber n.d.).

Edgington et al. (2004), notes however that an ontology does not necessarily have to serve merely as a vocabulary or a taxonomy, but can accommodate both content and mechanism

[8] *"The universe of discourse is the aggregate of the individual objects which 'exist,' that is, are independently side by side in the collection of experiences to which the deliverer and interpreter of a set of symbols have agreed to refer and to consider" (Peirce, cited in Kishore, Sharman & Ramesh 2004a).*

(Chandrasekaran, Josephson, & Benjamins, cited in Edgington et al. 2004), making provision for axioms and constraints to be added on top of the hierarchy of concepts and subsumption relations, thereby fully binding the intended interpretation of the ontology. Within the information systems environment, ontologies are therefore not only suitable for describing or modelling the information system, but are also very well suited for interacting with it.

For better understanding of the possibilities offered by ontologies, one should have a good understanding of the function of its components. Ontologies are usually composed of concepts (also known as classes), relations, constraints, axioms, and attributes (also known as roles or properties) (see Figure 4.7):

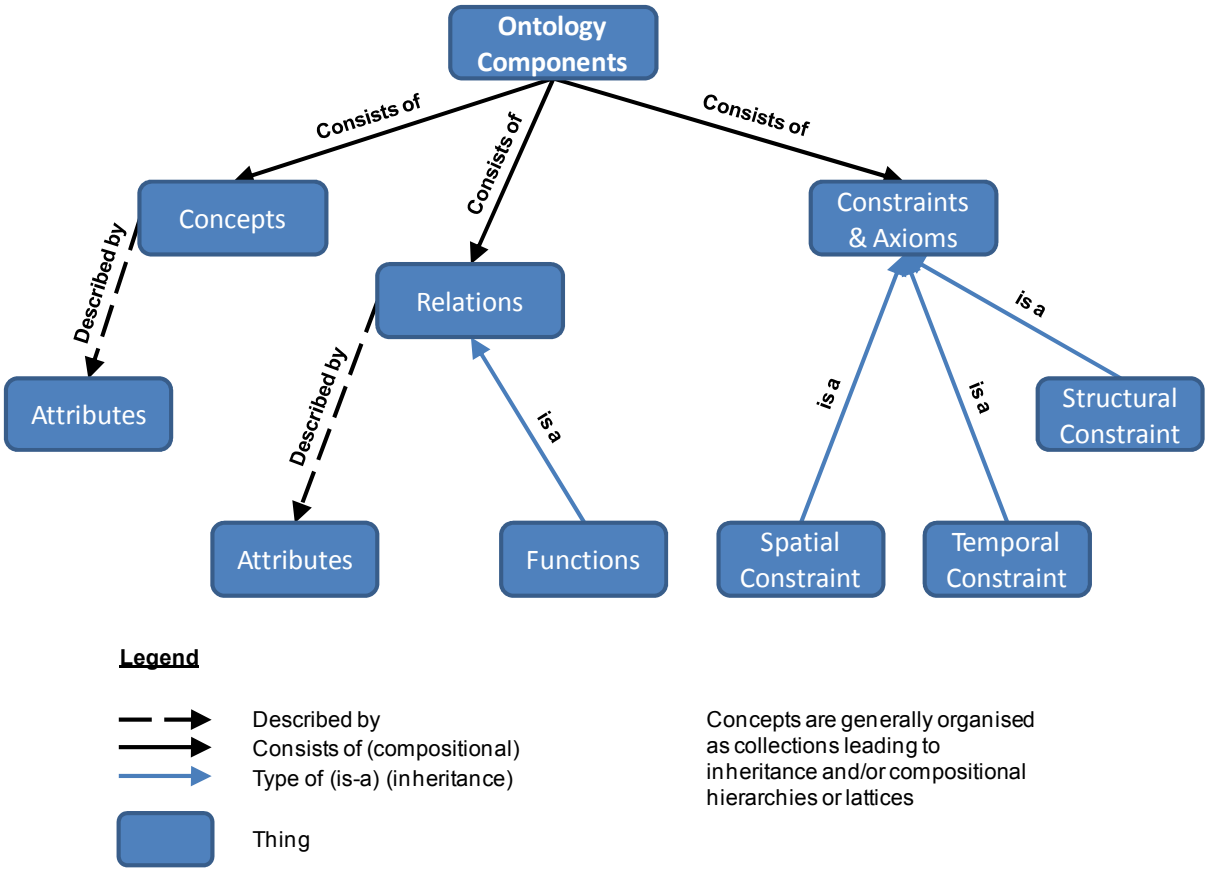


Figure 4.7: The components of an ontology (Source: Kishore, Sharman & Ramesh 2004b)

- Concepts: A concept can be defined as an abstract or general idea inferred or derived from specific instances. Within the information systems environment, these concepts denote terms and other aspects of the domain that are relevant to those who interact with it, and can usually be categorised as being either elementary or composite (Kishore, Sharman & Ramesh 2004b). These concepts are furthermore only labels or identifiers for real "things" found in the domain, with these "things" indicated as instances of the concepts within the ontology. Concepts can be grouped together into categories (Kishore, Sharman & Ramesh

2004b), which implies the existence of generic and specialised concepts (analogous to super-classes and sub-classes (Noy & McGuiness 2002)), resulting in the concepts being structured as a classical taxonomy (Mertins, Heisig & Vorbeck 2005).

- Relations: In the ontology, concepts are related to other ontology components through relations and functions (Kishore, Sharman & Ramesh 2004b). Additional to the hierarchical relations that exist between concepts in the taxonomy structure, other relations between concepts and concept categories are also allowed. A *function* (shown in Figure 4.7), is a special type of relation which relates some number of objects in the ontology to exactly one other object (with an object being any concept, instance, relation or function) (Kishore, Sharman & Ramesh 2004b).
- Constraints and Axioms: An ontology distinguishes itself from a taxonomy by providing possibilities of further specification of the concepts it describes, through the use of constraints and axioms (Mertins, Heisig & Vorbeck 2005). *Constraints* provide the means to place restrictions and boundaries on both the static and dynamic aspects of an ontology, and provide very good assistance for the representing and inferring of information. Structural, spatial and temporal constraints are the most common in ontologies, and these can also be classified as hard (must be satisfied), or soft (should be satisfied). Constraints may also represent concrete or inferred information. An *axiom* is a special type of constraint that can be assumed without proof. It provides factual truth from which useful conclusions can be derived, and are included in ontologies for the verification of correctness and the deducing of new facts (Kishore, Sharman & Ramesh 2004b).
- Attributes: Attributes represent any additional information associated with an ontology object to improve the understanding of the object, and to describe it in a more complete manner. It is usually directly associated with the particular object, and not accessible through a relation.

An example of an ontology can be seen in Figure 4.8.

- ***Application***

An ontology defines a common vocabulary that can be employed for the effective sharing of information within a domain, with the added benefit of also providing machine-interpretable definitions of the basic concepts found in the domain, as well as the relations that exist among them (Noy & McGuinness 2002). Because of this, ontologies are (mostly) implemented for the following reasons (Noy & McGuinness 2002):

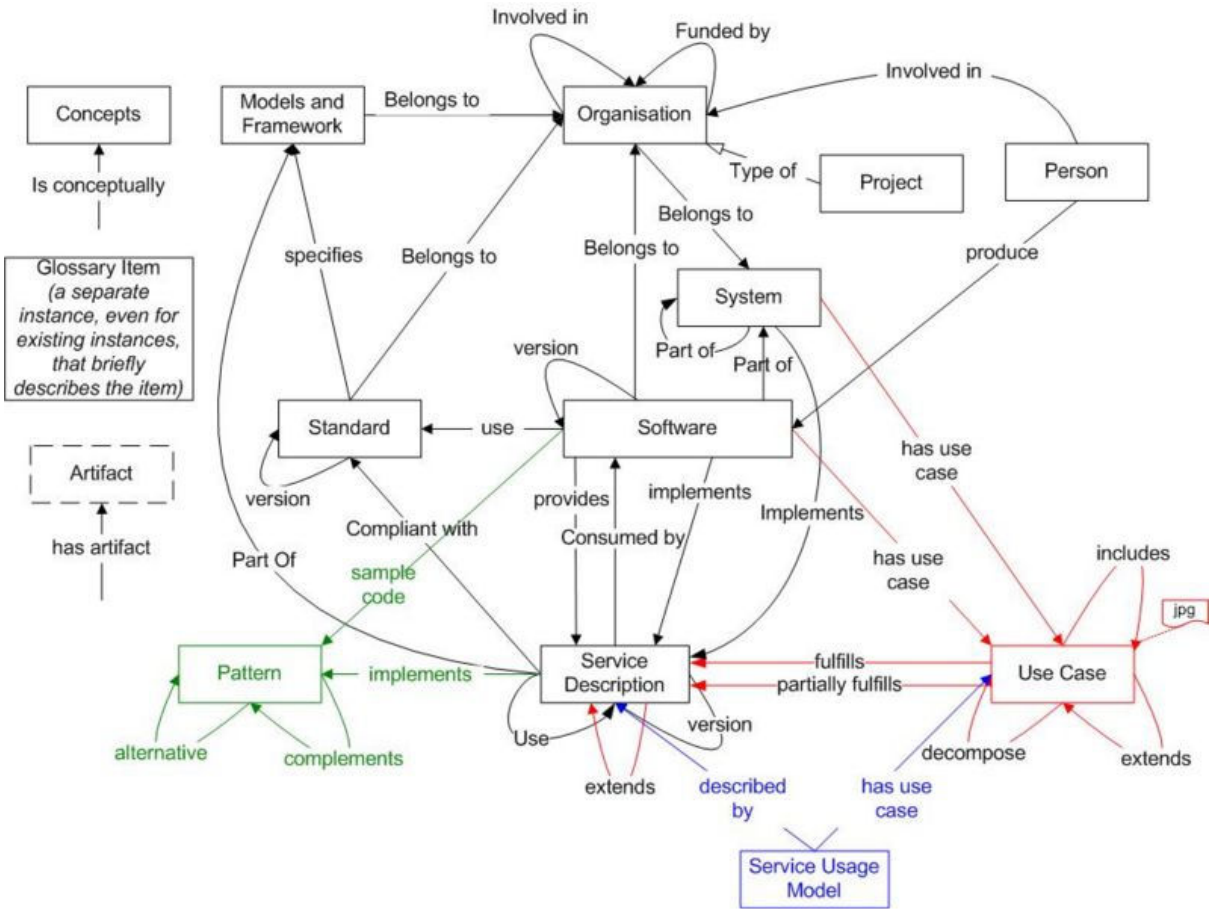


Figure 4.8: An example of an ontology (Source: FREMA Ontology 2006)

- To create a shared understanding of the structure of information among people or software agents;
- To enable the reuse of information in the knowledge domain;
- To make the assumptions underlying the knowledge domain explicit and well-understood;
- To specify the knowledge embodied in an ontology at an appropriate level of granularity (universe, bounded universe, domain, operational); and
- To provide means for analysing domain knowledge.

For a more clearly defined categorisation of ontology types, the research of Jasper and Uschold (cited in Kishore, Sharman & Ramesh 2004a) identifies the following four broad groups of ontologies (based on their manner of application):

- Ontologies for knowledge reuse;
- Ontologies as specification;
- Ontologies as a provider of common access of heterogeneous information; and

- Ontology as a search mechanism.

When one considers the different ways in which it can be applied within a particular domain, the flexibility of the ontology concept becomes very apparent. To illustrate this point, some distinctions made between different kinds of ontologies are briefly explained (Kishore, Sharman and Ramesh 2004a):

- Content vs. Mechanism Ontologies: Content ontologies are aimed at providing declarative knowledge and are therefore purely sources of information, whereas mechanism ontologies are aimed mainly at procedural knowledge, and facilitating the execution of methods and tasks in the information domain.
- Terminological vs. Axiomatic Ontologies: The content in terminological ontologies is not required to be fully specified by axioms and definitions, and is basically only simple collections of categories and terms. Axiomatic ontologies on the other hand, are distinguished by axioms and definitions stated in logic or some computer-oriented language, allowing it to be automatically translated into logic.
- Computational ontologies vs. Meta-ontologies: A computational ontology captures knowledge about a universe of discourse, but a meta-ontology is a language that can be used to represent an ontology. Such a language is furthermore also an ontology itself, and can be informal (like English), semi-formal (like UML), or formal (like First Order Logic).

Because of the formal manner in which ontologies are specified, together with its capability for expressiveness (Breu & Ding 2004), sharing and integrating of information between ontologies can be done with relative ease. The knowledge of two domains is merged (and therefore expanded), allowing members of the two respective domains to increase their own knowledge, and to interact with the new knowledge in a manner that is familiar to them. It is important that the terms within these ontologies (together with their precise definitions and meanings) are accompanied by a set of axioms that constrain the interpretation and well-formed use of these terms (Kim 2000). In this way, ambiguity in interpretation is minimised (thus preserving context) and the understanding of one domain's terms by its users can be transferred in an exact manner to the users of another domain.

The effective integration of ontologies however, necessitates the use of ontology languages, developed to specify the precise meanings of the terms within an ontology in a standardised format (Bechhofer et al. 2000). Just as ontologies can vary quite drastically in its application, the ontology languages development has also yielded some very diverse methods of specifying ontologies, amongst others: Logic-based languages, production rules, semantic nets, frame-based languages, description logics, and mixed formalisms (Kishore, Sharman & Ramesh 2004b). Some of these languages currently being developed and studied are DAML (DARPA Agent Markup Language), OIL

(Ontology Inference Layer), OWL (Web Ontology Language) and RDF (Resource Description Framework) Schema[9] (STARlab – Research n.d.).

4.2.5 Folksonomies

- ***Concept***

The folksonomy is probably the youngest of the information modelling concepts discussed in this chapter, as it only emerged as a system in the last two years (which also explains why it is largely absent from academic literature) (Speller 2007). Existing literature is largely opinion-based, which gives a clear indication that the concept itself still needs to mature. It is however increasingly being employed on the Internet to facilitate user interaction with information, which justifies it being discussed here.

A folksonomy is an information retrieval methodology, through which information is categorised by using collaboratively generated, open-ended labels called 'tags' (Sinclair & Cardew-Hall 2007), attached to information through a labelling process called tagging (Noruzi 2006). The term "folksonomy" was introduced by Thomas Vander Wal as a combination of "folk" and "taxonomy" (Smith, cited in Stock 2007) to indicate the establishing of an information retrieval structure through user input. It was however found to be somewhat inaccurate (Golder & Huberman 2006), as it does not make use of notations and relations like taxonomies do, but is basically a flat list of uncontrolled terms (Stock 2007), with no hierarchy and no directly specified parent-child or sibling relations between the terms (Mathes 2004). Another (and possibly the most important) contrast from taxonomies, is that the structure of the labelling system is derived solely from user inputs, and is not designed and administered by information managers, systems administrators, archivists or librarians (Noruzi 2006).

Together with "folksonomy", other names have also been proposed to communicate the notion of a folksonomy, like "ethnoclassification" (Merholz, cited in Speller 2007), and "distributed classification" (Mejias, cited in Speller 2007). These other names have also met some opposition, like Mathes (2004) feeling that Merholz's use of "ethnoclassification" is inaccurate, because 'what is

[9] *OWL and RDF Schema are both employed in the specification of the Semantic Web, an initiative initially proposed by Berners-Lee et al. (cited in Gilchrist, 2003). 'Berners-Lee et al. described the Semantic Web as being: "not a separate Web, but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in co-operation", in line with his original conception of the Web as a universal, decentralised system; and one which does not discriminate between "the scribbled draft and the polished performance, between commercial and academic information, or among cultures, languages, media and so on"' (Berners-Lee et al. cited in Gilchrist, 2003).*

happening is quite unlike classification and far more like categorisation'[10]. From these different views, it is quite obvious that the concept still requires quite some discussion and research for it to reach maturity, but to get a better understanding of the concept without getting bogged down in semantics, it would be prudent to take a step back and first consider the activity from which a folksonomy originates, namely tagging.

Tagging can be seen as a type of collaborative free keyword indexing, where there are no indexing rules and users can tag content with any words they see fit (Stock 2007). Rather than adhering to a fixed information categorisation structure, these tags are attached to information in a random manner, aimed purely at providing a specific user with the means to retrieve information of a personal interest at a later time. What makes the use of tags so appealing to users, is that, instead of being forced to describe a piece of information by using the author's vocabulary or the vocabulary of a particular classification or categorisation system, users can apply their own language and vocabulary with which to describe the information (Mathes 2004). Tags can be words, acronyms, numbers or anything else that will assist the user in retrieving information, and the user is free to use the tags in any way that he/she feels appropriate, without having to agree with other users about how the specific content should be tagged (Shirky 2005) (see Figure 4.9).

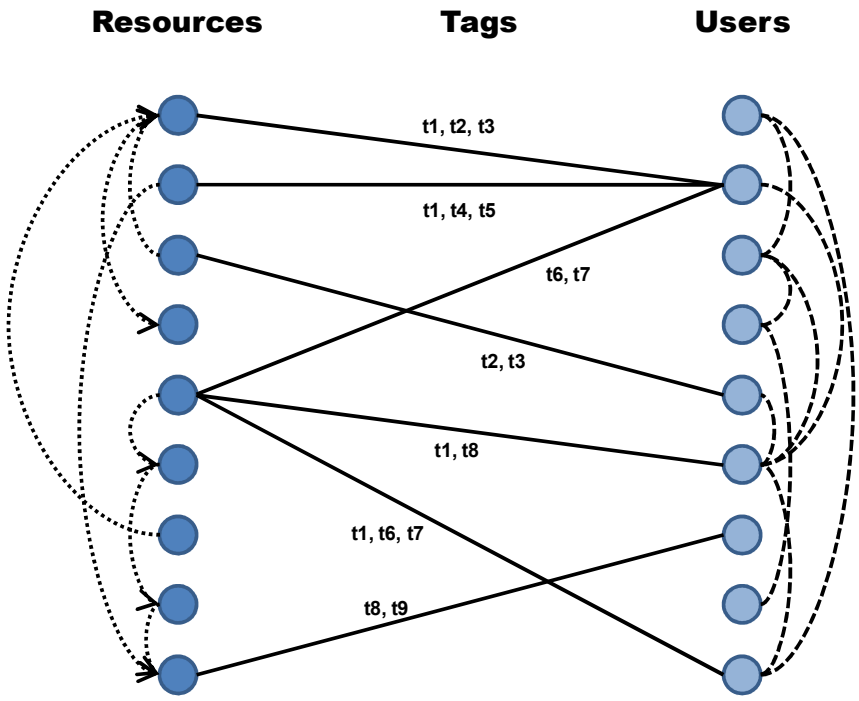


Figure 4.9: A model of a tagging system (Source: Auray 2007)

Through the tagging of various pieces of information content, related information are also grouped together, which happens purely because of the tagging itself and not through the use of fixed

[10] *To maintain coherence, "folksonomy" will be used for this concept in the rest of this section, as the aim is to rather focus on the characteristics of the concept than on the semantics of its naming.*

categories or applying officially approved choices (Noruzi 2006). In this way, a kind of "bottom-up social classification" (Vander Wal, cited in Noruzi 2006) of the information takes place, which is unlike traditional approaches used in library classification, like the Dewey Decimal Classification (DDC), or the Library of Congress Classification (LCC) (Noruzi 2006).

Considering that this chapter focuses on the application of relations within information, it is important to note that a folksonomy is based on a relaxation of the relations between a term and its index and 'unlike, say, the Linnaean system for classifying animals, which is an exemplary system, folksonomies are more horizontal, labelling systems organised on a grid and open to characteristics that are known in linguistics as sememes' (Auray 2007). In fact, at the start of a folksonomy, no index exists; the index is essentially a "cloud" of tags, which is created and grown as users add more and more tags to their content. This is obviously completely different from traditional methods through which indices were compiled by librarians or professional indexers, or else derived from the authors of documents (Spiteri 2007). Being a product of user inputs however, gives the folksonomy the ability to adapt much quicker to user vocabulary changes and needs than a conventional index would have been able to (Mathes 2004).

- ***Application***

The Internet has experienced a massive inflow of amateur content because of a significant increase in cultural capital on the part of Internet users, upgraded education, generalisation of intellectual work and the spread of the use of the Internet among white-collar workers. 'This wealth of amateur content represents a generalisation throughout all cultural life of the phenomenon known as the "democratisation of innovation" (Von Hippel, cited in Auray 2007); with amateurs injecting radical innovations into that realm undreamt of by engineers and whose existence specialists had previously primarily observed in cutting edge areas confined to IT, biotech or chemistry' (Auray 2007). This massive inflow of amateur content has however completely overturned the organisation of information on the Internet (Auray 2007), which left Internet users discovering and developing new ways to index and retrieve useful information. As a result of this, keen observers of the Internet and inventors of social software began to notice that users who don't write computer programmes were applying tags to the Internet content they create or encounter (Gruber 2005).

Tagging is becoming increasingly popular because it is fundamentally about sense making[11] (Golder & Huberman 2006). Web users describe and organise their content (bookmarks, web sites/pages or photos) with their own vocabulary, and assign tags to each unit of content. Tagging improves the searching and information interpretation of users, and helps them to identify the main ideas around

[11] *'Sense making is a process in which information is categorised and labelled critically, through which meaning emerges' (Weick, Sutcliffe & Obstfeld, cited in Golder & Huberman 2006).*

topics on the internet (Noruzi 2006). Also, with bookmarking, tagging helps to counter the spam-induced noise in search engines, and with photo and video sharing, tagging helps to improve the precision and recall of search engines (Gruber 2005). Some popular Internet services which make use of tagging are Del.icio.us (a bookmarking service), Flickr (a photo sharing service) and YouTube (a video sharing service)[12] (Stock 2007).

For a better understanding of the manner in which tagging can be applied, one can consider the following functions performed by tags (specifically for bookmarking), which were identified by Golder & Hubermann (2006):

1. Identifying what (or who) the specific content is about: An overwhelming majority of tags fulfil this function, and include common nouns of many levels of specificity, as well as proper nouns (specifically when describing people or organisations).
2. Identifying what the specific content is: Tags can be used to identify what type of content the specific bookmarked item is, in addition to what it is about, like *article, blog* and *book*.
3. Identifying who owns the specific content: Some content are tagged according to who owns or created it.
4. Refining categories: Certain tags are not used alone and, rather than establishing categories themselves, are used to refine or qualify existing categories (like *25* and *100*).
5. Identifying qualities or characteristics: Adjectives such as *scary*, *funny*, *stupid* and *inspirational* is used to describe the tagger's opinion of the specific content.
6. Self Reference: Tags such as *mystuff* and *mycomments*, aimed at identifying content in terms of its relation to the tagger.
7. Task organising: When collecting information related to performing a task, that information might be tagged according to that task, in order to group that information together (examples include *toread* and *jobsearch*).

A mechanism that further increased the usefulness of tagging is the *tag cloud*, which is an interface element commonly associated with folksonomy datasets (see an example of a typical tag cloud in Figure 4.10). Sinclair & Cardew-Hall (2007) provides the following explanation of a tag cloud:

'A tag cloud (more traditionally known as a *weighted list* in the field of visual design) is a visual depiction of content tags used on a website. Often, more frequently used tags are depicted in a larger font or otherwise emphasized, while the displayed order is generally alphabetical. Thus both finding a tag by alphabet and by popularity is possible. Selecting a single tag within a tag cloud will generally lead to a collection of items that are associated with that tag.'

[12] http://del.icio.us/, http://www.flickr.com/, http://www.youtube.com/

Figure 4.10: An example of a tag cloud (Source: Everyone's Tags 2007)

It is obvious that the concept of the folksonomy (or more generally, the notion of tagging), is very well suited to provide users with the information they require in a manner that is familiar to them. Its novel approach (letting the actions of users determine the structure of information rather than the ideas of a select group of dedicated professionals) appeals to users, but has some shortcomings, such as that it does not enforce precision, may have difficulty with distinguishing between different languages and has no control over synonymy and homonymy (Stock 2007); problems which other approaches have much less difficulty to address. Its benefits however (such as providing an authentic representation of the language used by a specific community (Quintarelli, cited in Stock 2007), allowing for multiple interpretations of information used by the same community (Peterson, cited in Stock 2007) and providing an environment for serendipity to take place (Mathes 2004)), make it far too valuable to discard, which has led some individuals to suggest that it should be used in conjunction with existing (and more structured) methods, which will yield much more complete solutions for information structuring and retrieval (an example of such an integrated solution can be seen in Figure 4.11, which depicts an information environment in which the functionality of ontologies, text-oriented methods and folksonomies are integrated). As the folksonomy concept matures, the robustness of implementations will improve, and will naturally cause the concept to be implemented much more extensively.

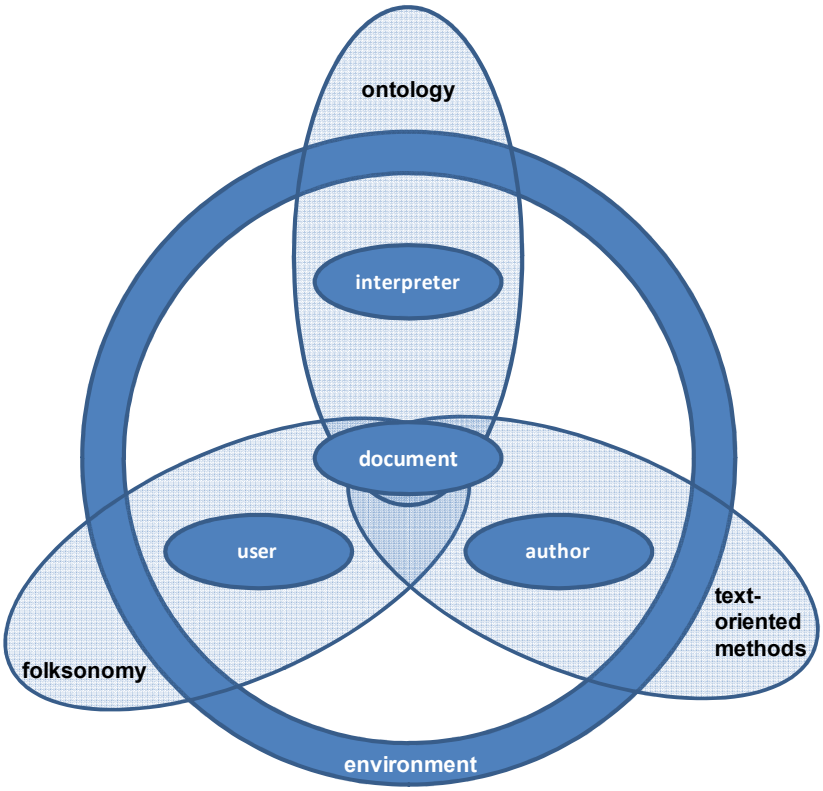


Figure 4.11: An example of an information environment in which the functionality of ontologies, text-oriented methods and folksonomies are integrated (Source: Stock 2007)

4.2.6 Concept maps

- ***Concept***

A concept map is a two-dimensional representation of a set of concepts that is constructed so that the inter-relations among them are evident (Carvalho, Hewett & Cañas 2001), and is aimed at providing a means for people to easily organise and represent knowledge for improved understanding (Novak & Cañas 2006). It consists of concepts (a concept defined by Novak & Cañas (2006) as being 'a perceived regularity in events or objects, or records of events or objects, designated by a label'), and the relations that exist between the concepts (typically specified by linking words or linking phrases). Concept maps represent meaningful relations between its concepts in the form of *propositions*. Propositions are statements about some object or event in the universe (either occurring naturally or having been constructed), each containing two or more concepts which are connected through linking words or phrases (i.e. relations) to form meaningful statements. Propositions are also sometimes called semantic units, or units of meaning (Novak &

Cañas 2006). An example of a concept map (created by NASA for Mars Exploration (Novak & Cañas 2006)) can be seen in Figure 4.12.

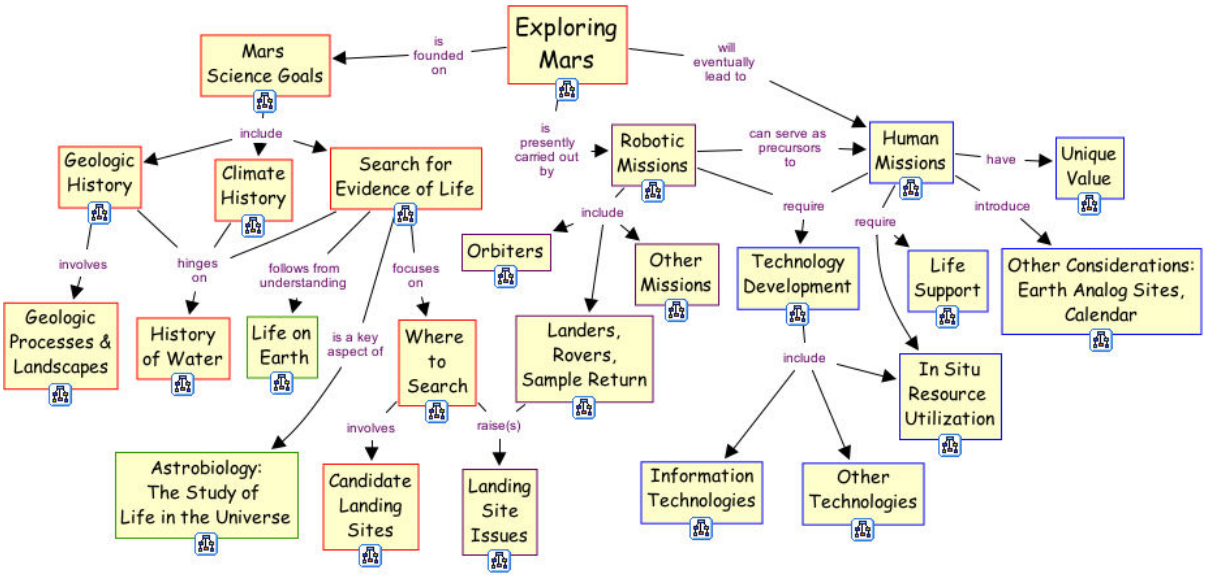


Figure 4.12: An example of a concept map (Source: Novak & Cañas 2006)

Concept maps is structured mainly in a hierarchical fashion, with the most inclusive, most general concepts at the top of the map, and the more specific, less general concepts arranged hierarchical below. Context however, also plays a major role in the structuring of concept maps, and therefore they are generally constructed to answer a specific question (called a focus question), which provides a guideline as to the manner in which knowledge in the concept map should be structured. The context of concept maps is further enhanced through the use of cross-links, which are relations between different segments or domains of the concept map. Cross-links aim to providing insight in the manner in which concepts in different domains of the concept map are related. Finally, specific examples of events or objects may also be added to concept maps in order to clarify the meaning of certain concepts (Novak & Cañas 2006).

Additional to an obvious ability to improve understanding by representing knowledge in a visual manner, concept maps are specifically targeted to act as visual languages. Visual languages are similar in their characteristics to natural language text, in that syntactic and semantic constraints can be applied to them, and their representation capacity can range from fairly informal to extremely formal. Being a visual language that is both comprehensible and formal, concept maps are able to be effectively utilised for both comprehension and editing, and also for parts of the knowledge acquisition process itself (Gaines & Shaw 1995). It is furthermore well suited for the generation of new knowledge, as experts who build concept maps are empowered to not simply externalise pre-existing internal knowledge but also, through phases of comparison and consideration, perform knowledge construction (Leake, Maguitman & Cañas 2002).

- ***Application***

Concept maps were developed by Joseph Novak in 1972 through research done at Cornell University, which was aimed at understanding the changes that occur in children's understanding of science (Novak & Musonda, cited in Novak & Cañas 2006). The research was based on the learning psychology of David Ausubel, in which the fundamental idea is that learning takes place through the assimilation of new concepts and propositions into the existing concepts and propositional frameworks held by the learner. An adequate way to represent the conceptual understanding of children was required, which gave rise to the idea of representing the children's knowledge in the form of a concept map (Novak & Cañas 2006).

Since then it has evolved into a technique that successfully enables the examination of human conceptualisations and assists in human knowledge construction by providing a rich and flexible concept representation that help humans understand domains and revise their domain knowledge (Leake, Maguitman & Cañas 2002). Apart from being used in education, not only as a learning tool, but also as an evaluation tool (Mintzes et al.; Novak; Novak & Gowin, all cited in Novak & Cañas 2006), concept maps have since been employed in various different environments, and have also influenced a number of other well-known knowledge representation approaches, as indicated by Gaines & Shaw (1995):

'Concept maps have been used in education, policy studies and the philosophy of science to provide a visual representation of knowledge structures and argument forms. They provide a complementary alternative to natural language as a means of communicating knowledge. In many disciplines various forms of concept map are already used as formal knowledge representation systems, for example: semantic networks in artificial intelligence, bond graphs in mechanical and electrical engineering, CPM and PERT charts in operations research, Petri nets in communications, and category graphs in mathematics.'

Wikipedia, the free encyclopedia (2007a), lists the following as some of the tasks for which concept maps are currently being used (especially with regards to business and education):

<u>Business</u>

- Note taking and summarising of key concepts, their relationships and hierarchy gleaned from documents and source materials;
- New knowledge creation: e.g., transformation of tacit knowledge into an organisational resource, and mapping team of knowledge;
- Institutional knowledge preservation (retention), e.g., eliciting and mapping expert knowledge of employees prior to retirement;

- Collaborative knowledge modelling and the transfer of expert knowledge;
- Facilitating the creation of shared vision and shared understanding within a team or organisation;

Education

- Instructional design: concept maps used as Ausubelian "advance organisers" which provide an initial conceptual frame for subsequent information and learning;
- Training: concept maps used as Ausubelian "advanced organisers" to represent the training context and its relationship to persons' jobs, to the organisation's strategic objectives, to training goals;
- Increasing meaningful learning;
- Communicating complex ideas and arguments;
- Examining the symmetry of complex ideas and arguments and associated terminology;
- Detailing the entire structure of an idea, train of thought, or line of argument (with the specific goal of exposing faults, errors, or gaps in one's own reasoning) for the scrutiny of others;
- Enhancing meta-cognition (learning to learn, and thinking about knowledge);
- Improving language ability;
- Assessing learner understanding of learning objectives, concepts, and the relationship among those concepts.

Finally, it should be mentioned that concept mapping is not equivalent to the technique of *mind (or idea) mapping*. Mind mapping is used for the capturing of ideas or impressions generated in a spontaneous manner (through brainstorming, for instance). Ideas are usually captured and portrayed in a radial fashion, with the mind map then serving as a focus for discussion (Figure 4.13 shows an example of a mind map). A concept map, on the other hand, captures knowledge in a hierarchical fashion, is generated through consideration and collaboration, and serves to provide a system view of a real or abstract system or set of concepts (Wikipedia, the free encyclopedia 2007a).

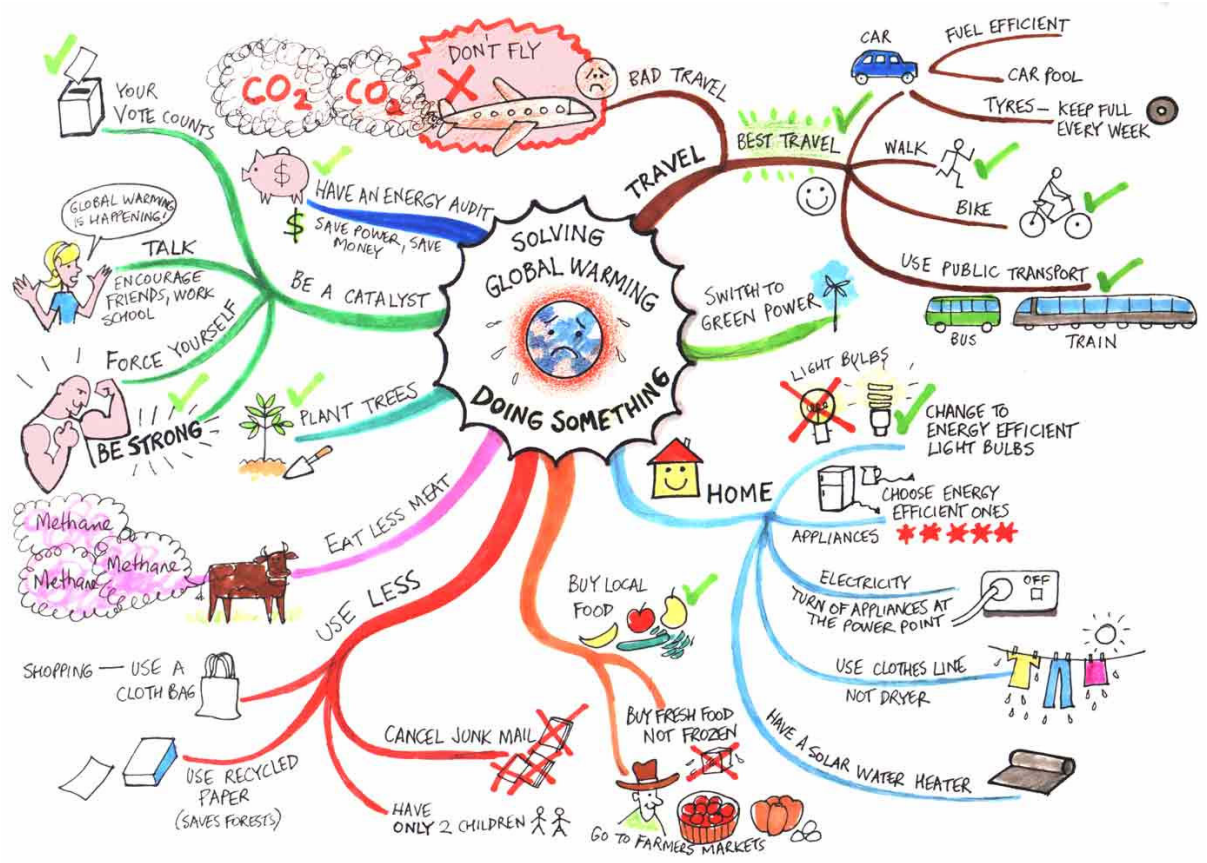


Figure 4.13: An example of a mind map (Source: Combating Global Warming Mind Map 2007)

4.2.7 Topic maps

- ***Concept***

The topic map concept is an ISO standard (ISO/IEC 13250:2003) developed for the representation and interchange of knowledge (Wikipedia, the free encyclopedia 2007f) which specifically focuses on improving the "retrievability" of information. This is done by the topic map acting as a superimposed external layer, aimed specifically at describing the nature (i.e. subject) of the knowledge represented in specific information resources (as opposed to metadata which only really describes the characteristics of the information sources) (Biezunski & Newcomb 2001). It can almost be understood as 'a semantic net stretched over the actual information resources' (Mertins, Heisig & Vorbeck 2005). Being external to the information sources, means that there are no real limitations to the kinds of information that can be described by topic maps. With the ability to describe the nature of knowledge, also comes the capability to identify similar concepts in disparate information objects and, based on this similarity, to associate all the objects relevant to a specific concept with one another (Biezunski & Newcomb 2001).

Rath (2003), when comparing topic maps to three of the most common forms of information location referencing in use today; catalogue reference, table of contents and indexing, said the following:

'A topic map can be seen as an electronic super index, implementing the back-of-book index paradigm and much more.

Topic maps are various technologies in one:

- Complex metadata: A topic map contains information about information resources. It is not part of the information resources, it is created, managed, and stored separately from the resources, but could be closely connected to them;
- Search index: As searching in a back-of-book index is a very precise searching method, searching in a topic map provides better search results (than) searching in a full text index. A topic map can be seen as an intelligent search index;
- Link network: Topic maps are well-organised link networks helping to avoid the "lost-in-hyperspace" syndrome;
- Knowledge structure: Topic maps are a base technology for explicit knowledge modelling and knowledge navigation – hence their value to Knowledge Management.'

Topic maps are constructed through the use of three basic concepts, namely topics, associations and occurrences.

Topics act as the representation of *subjects* within a computer, with a subject being defined by the ISO/IEC 13250 standard as follows (ISO/IEC 13250 Topic Maps 2002):

'In the most generic sense, a 'subject' is anything whatsoever, regardless of whether it exists or has any other specific characteristics, about which anything whatsoever may be asserted by any means whatsoever.'

A note on the subject definition furthermore states the following (ISO/IEC 13250 Topic Maps 2002):

'The invisible heart of every topic link is the subject that its author had in mind when it was created. In some sense, a topic link reifies a subject. The identity attribute of a topic link is provided to allow the author of the topic link to indicate, as unambiguously as possible, the subject he had in mind as the organising principle of the topic.'

Topics act as surrogates for subjects, and have characteristics like names, occurrences and associations with other topics (Newcomb & Biezunski 2001). Being characterised by its associations

furthermore enables a topic to serve as a hub which connects everything that pertains to the particular subject it represents (Biezunski & Newcomb 2001).

The name of a topic serves as an identifier through which applications are able to interact with the contents of the topic map and perform meaningful actions on it (Pepper 2002). A topic also is not restricted to having only one name, but may have multiple names, something that taxonomies and thesauri do not allow. It is therefore possible for different users and applications to refer to a specific topic using their own terminology, but still be sure that all parties are referring to the same topic (Garshol 2004).

An *association* is a topic map construct aimed at asserting the relation between topics (Pepper 2002). Its role is to simulate the manner in which people relate different topics to each other, and in doing this, provides the context that people require to understand the various topics that exist in a topic map (Rath 2003). It sheds light on the general interconnectedness that exists within a topic map, and can show for example, the groupings of topics in the topic map that are similar in meaning. Unlike some of the other information modelling approaches discussed in this chapter, associations can be defined in quite a free manner, and do not necessary have to be defined according to a specific structure, like a hierarchy. This aspect however, requires that each association be defined clearly, which invalidates the existence of generic relations in the topic map (Garshol 2004).

Topic *occurrences* are references (Mertins, Heisig & Vorbeck 2005) to information objects external to a topic map, which bear some relevance to one or more of the subjects which are represented by topics in the topic map (Biezunski & Newcomb 2001). These references are the mechanism by which a topic map connects to a knowledge base, making the information it contains much more visible, and describing the information it contains in such a manner that it can be retrieved much more effortlessly than before. It is however very unlikely that any topic occurrence will be an exact instantiation of a subject represented in a topic map. It should therefore always be remembered that a topic occurrence is a resource declared to contain some kind of information about the topic under consideration, and that it's only useful under certain conditions (that is, within some "scope"). (Biezunski & Newcomb 2001).

A diagram illustrating the various basic concepts of a topic map can be seen in Figure 4.14. Note that in this case, the link between a topic and its occurrence is established through the use of an URI (Universal Resource Identifier) (Internet Engineering Task Force (IETF), cited in Rath 2003).

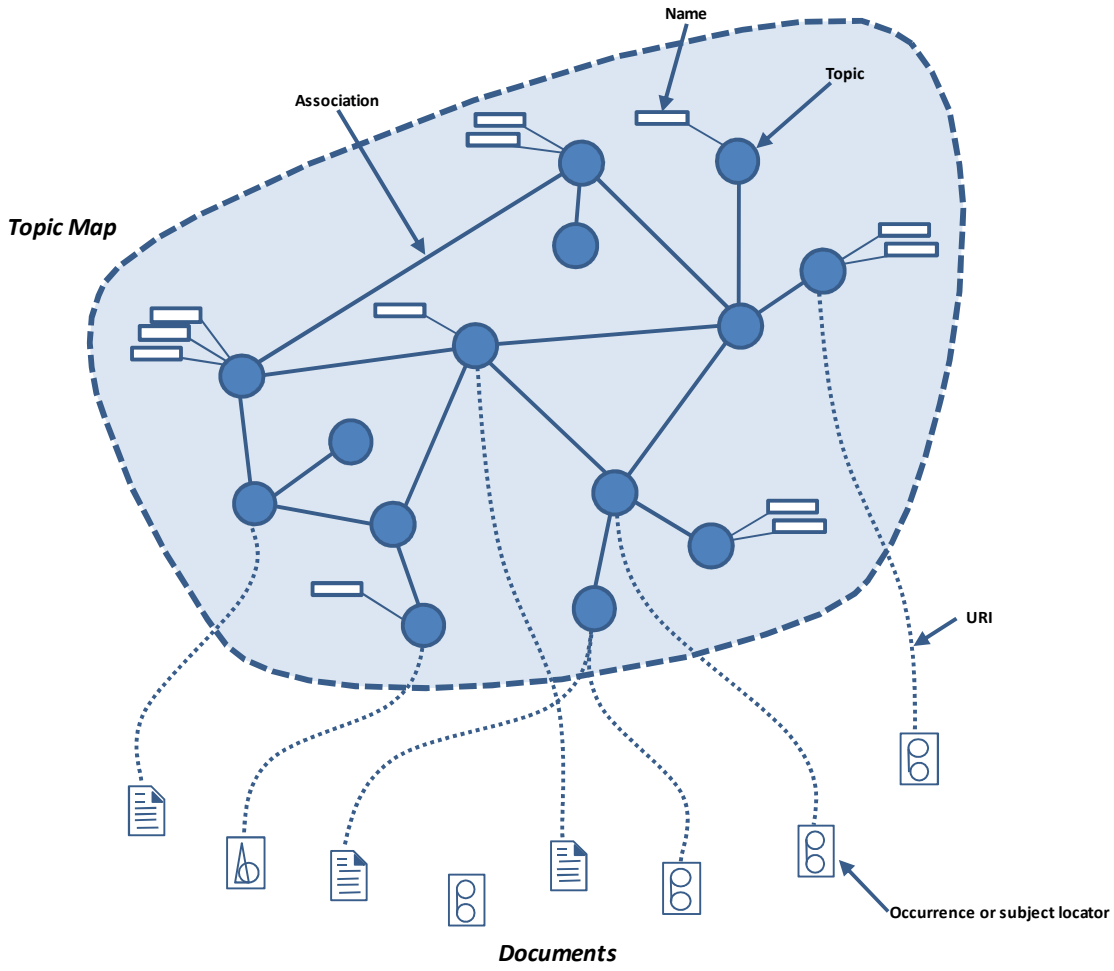


Figure 4.14: The components of a topic map (Source: Topic Map Example 2007)

Finally, topic maps make use of the following concepts for control and ease-of-use purposes:

- Types: Topics, associations and occurrences can all be typed (Biezunski & Newcomb 2001; Pepper 2002), with topic, association and occurrence types[13] also being topics. It is a powerful means to perform classification and grouping of topic map elements, and allows a topic map user to better describe the world from which the topics are taken (Garshol 2004). If necessary, it is ideal for structuring topics in hierarchical or any other type of structure aimed at improving users' understanding of the topic map.

- Scope: In order to address the problem of context, topics and their associated characteristics and associations can be defined to only be valid within certain limits, which may or may not be specified explicitly (Pepper 2002). These limits (i.e. scope) act as a filtering mechanism, ensuring that only information relevant to the user is displayed. Just some of the ways in which scope can be applied to the information in a topic map is the modelling of different

[13] *Rath (2003) refers to types used by topic maps as "classes", but the ISO 13250 standard uses the term "type", which is also preferred in this document.*

languages, access rights, views on information, user skills, and differences in interest (Rath 2003).

- Roles: In some cases ordinary associations do not provide enough context for a user to understand the information contained by a group of topics. This can be remedied by associating roles (which are topics themselves) to the topics, which will then indicate the role that each topic plays in a specific association (Pepper 2002).

• *Application*

Topic maps originated in work on the merging of electronic indexes (Garshol 2004), and were developed to fulfil a requirement that most conventional information retrieval approaches fail to address: the *subject* of an information object. Metadata is very useful to describe the properties of an information object, but has very limited capabilities for conveying what the information object is about (Garshol 2004). Full-text indexing on the other hand is very useful to extract the words that resides within an information object, but fails to provide the context that will give the words meaning. Topic maps focus on addressing the limitations of both of these approaches by performing *subject-based classification*, which is classification that groups objects by the subjects they are about. This approach have resulted in much more precise searching for information being possible, and means that navigation systems driven by topic maps can be much richer and more flexible than simple tree-based navigation (Garshol 2004).

Additional to its capabilities for accurately describing information and performing subject-based classification, the ISO/IEC 13250 topic map standard has also made provision for the merging of topic maps. This ability was built into the topic map standard, because the merging of indices was the main reason why the development of topic maps was initiated (Rath 2003). Topic maps can therefore be developed in different domains, and consolidated into a single topic map, which will provide a unified view on the information contained in the group of domains.

The navigation and retrieval abilities possessed by the topic maps approach (together with its capability for merging information domains) are ideally suited for Internet application. SGML (Standard Generalised Markup Language) was originally specified by the topic map standard as the syntax through which topic maps should be described, but proved to less ideal for application in an Internet environment (Lacher & Decker 2001). To address this shortcoming, the XML Topic Map standard (XTM) was developed in 2001 based on the ISO/IEC 13250 standard (Mertins, Heisig & Vorbeck 2005), and allows topic maps to be implemented for Internet use (Lacher & Decker 2001).

Topic maps have been touted by some as the future of the internet. Already in 1999 Charles F. Goldfarb, the inventor of SGML, branded topic maps as the ‘GPS of the information universe’ (Rath

2003), and both the Gartner Group (in 2000) (Gartner group, cited in Rath 2003) and the META Group (in 2001) (META Group, cited in Rath 2003) have recommended that enterprises that are facing significant challenges in managing information should track the development of topic mapping technology (Rath 2003). Rath (2003) gives the following reasons why topic maps are likely to survive (and excel) in the information technology world:

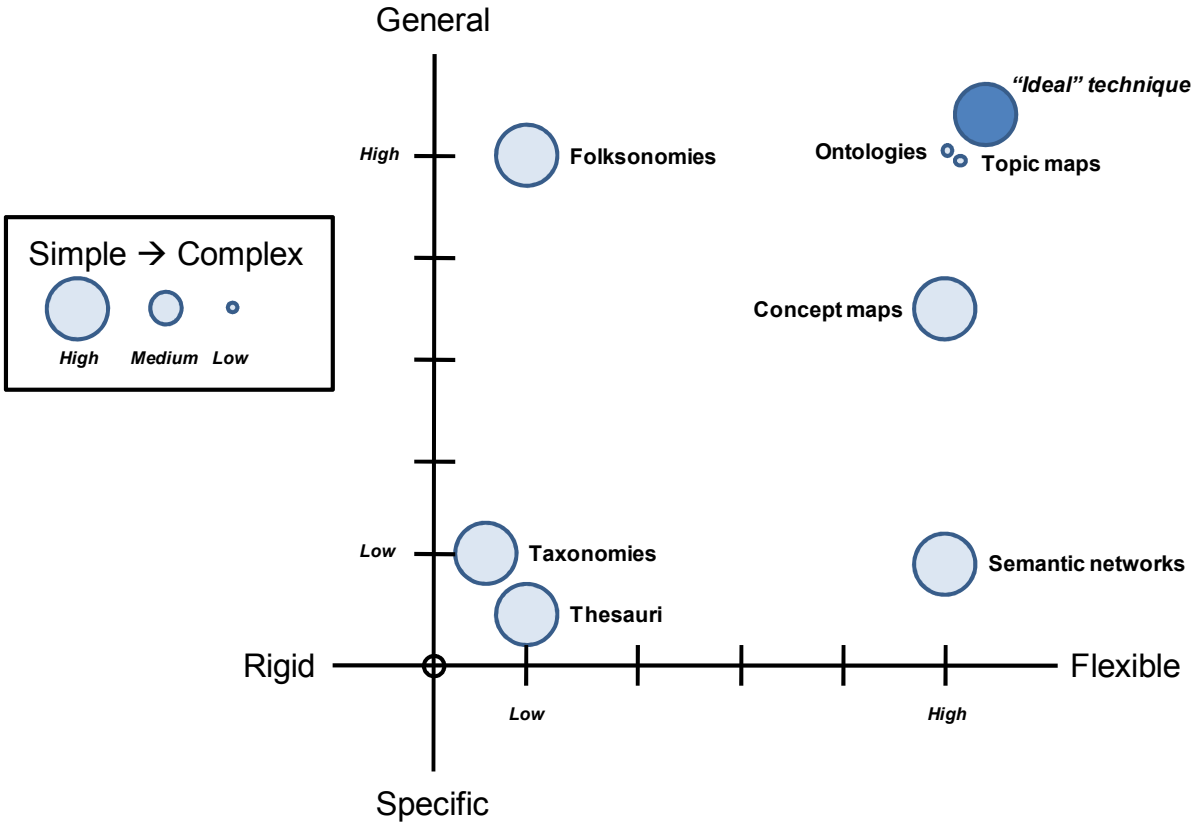
- 'The Internet provides both the infrastructure to deliver information access to everybody's desktop as well as the required organisation of large amounts of information – both not existing when Semantic Networks were developed. So, the world is ready for topic maps – the problems they solve and the infrastructure they need are in place.
- Topic maps are not a proprietary solution but an international ISO standard describing electronic indices, classification schemas, knowledge structures, that are Web enabled, and ready to use in the Internet and Intranets.
- Topic maps are not an academic knowledge representation paradigm. They are built by practitioners, focussing on real-world requirements and on implementable solutions.
- Topic maps not only cover the meta layer above the resources, they are also providing a powerful mechanism to connect resources within the meta layer. They are the bridge between Information Management and Knowledge Management.'

Finally, the topic map approach has the ability to interact and share information with other information modelling approaches, which simplifies the transfer of information. Approaches have been proposed through which topic maps can integrate with both ontologies (Vatant 2004) and RDF Schema (Lacher & Decker 2001), as part of the Semantic Web initiative (see section 4.2.4). This will ensure that efforts to develop a "well-defined Internet" are not constricted to the views of a single group of people, allowing for the existence of different views while still achieving the sought-after outcome.

4.3 Conclusion

This chapter has discussed a number of the entity-relation techniques that can be employed in a variety of ways to model information in such a way that the manner in which humans (and by extension organisations) interact with and share information is improved. These have included techniques tailored for linguistic purposes (semantic networks, thesauri) and information classification (taxonomies, folksonomies), techniques for focused discussion and machine-interpretability (ontologies, topic maps) and for facilitating improved human understanding (concept maps). Each of the techniques has been developed to address a very specific need in the field of human-information interaction, and does this mostly in a very

effective manner. Figure 4.15 shows a comparison between these techniques in terms of three measures, namely generality (i.e. does the technique cater for a specific need, or can it be applied in a variety of different manners), flexibility (i.e. does the technique lend itself toward having its structure configured to address a specific need, or does it have a set structure), and simplicity (i.e. how easy do users find the approach to understand and interact with). The rationale behind the placement of the techniques on the coordinate system is as follows[14]:



Technique	Generality	Flexibility	Simplicity
Semantic networks	Low	High	High
Thesauri	Low	Low	High
Taxonomies	Low	Low	High
Ontologies	High	High	Low
Folksonomies	High	Low	High
Concept maps	Medium	High	High
Topic maps	High	High	Low
"Ideal" technique	**HIGH**	**HIGH**	**HIGH**

Figure 4.15: Comparison of information modelling techniques in terms of generality, flexibility and simplicity

[14] *Interestingly enough, none of the techniques (except for Concept maps) appear in the medium ranges of the landscape, which gives the impression that the development of these techniques is generally focused on addressing specific needs, and not on providing solutions that partly address a large variety of needs.*

- *Semantic networks* are very simple to understand and quite flexible in the manner in which they can be constructed, but are focused on a very specific need, namely the clarification of a specific language.
- *Thesauri* are also very simple to understand, but have quite a rigid manner in which terms are organised within its structure, and are also focused on a very specific need, which is to indicate the manner in which terms of a specific language are related.
- *Taxonomies* are (similar to thesauri) very simple to understand and are constructed in a very definite manner, but in this case are focused on indicating the "natural" hierarchy that exists within a group of concepts.
- *Ontologies* are on the other hand, quite complex to understand, but with the complexity comes the ability to apply the technique in a myriad of different manners, which are greatly helped by the technique providing great flexibility in the structuring of information.
- *Folksonomies* are again very simple to understand and can be applied in various contexts and environments, but are not very flexible as it has to be constructed in a very specific way.
- *Concept maps* are very simple to understand and are very flexible in the manner in which it can be constructed, but cannot be as generally applied as, for instance, ontologies or topic maps. This is because, although different types of information can be represented as concept maps, the technique focuses on addressing a specific need, which is to represent the information in such a manner that a person can easily comprehend it.
- Finally, *Topic maps* are (similar to ontologies) quite complex to understand, but again, provide great flexibility in application, which affords it the ability to be applied in a large variety of different manners.

Unfortunately the application of these techniques sometimes take place without due consideration of its limitations, or the technological prowess that will be required from its users. The problem usually associated with the simpler techniques, is inability to address a great variety of information representation needs, whilst the more complex techniques usually have the problem of requiring too much technical knowledge and input from its users for them to truly experience its value. It would seem that a need still exists for an information representation technique which can be easily understood by most information users (without having an extensive technical knowledge of the technique's inner workings), whilst still providing enough functionality to address information modelling requirements over a broad spectrum of application (see the technique designated **'"Ideal" technique'** in Figure 4.15).

The next chapter introduces another approach to the entity-relation view of information modelling, aimed at addressing this very need, and on which research is currently being conducted at Stellenbosch University.

Chapter 5 – Conceptual Framework Research

5.1 Introduction

The Global Competitiveness Centre[15] (GCC) of the Stellenbosch University's Industrial Engineering department, in conjunction with an industrial engineering company named Indutech (Pty) Ltd[16] , has started doing research in 2004 on an approach for representing information in terms of entities and the relations that exist between them. It originated from a need identified through contact with industry for an approach by which complexity in the products, processes and/or systems of organisations can be visualised and, through improved understanding, be reduced. Some of the more complex techniques discussed in the previous chapter (Chapter 4) were investigated, and although the functionalities offered by techniques such as ontologies would have sufficiently addressed the need, the amount of detail and effort that an acceptable solution would require, was deemed to be unacceptably high. In the end, the need was suitably addressed through the application of an approach that resembled a very simplified variation of these techniques, whilst still having sufficient information modelling capabilities. The approach is based on the assumption that information can be represented in terms of the entities that constitute the information together with the relations that exist between these entities, and is much simpler to understand and apply than the more complex established approaches, such as ontologies and topic maps. The network of entities and relations that is generated through this approach is termed a "conceptual framework", and is aimed at assisting a person in the understanding, navigation, and discovery of information. As indicated by Figure 5.1, this chapter will provide a detailed definition of the concept, as well as the hypothesis for the project which this thesis reports on, and the research method that was followed in the execution of the project.

5.2 Defining conceptual frameworks

This section is aimed at familiarising the reader with a concept that forms the basis of research that was conducted, and of which the outcome is presented by this thesis. It should however be noted that this section is aimed at merely establishing an understanding of the concept, and not at arguing the validity thereof.

[15] *Henceforth referred to as GCC*

[16] *Henceforth referred to as Indutech (Website: www.indutech.co.za)*

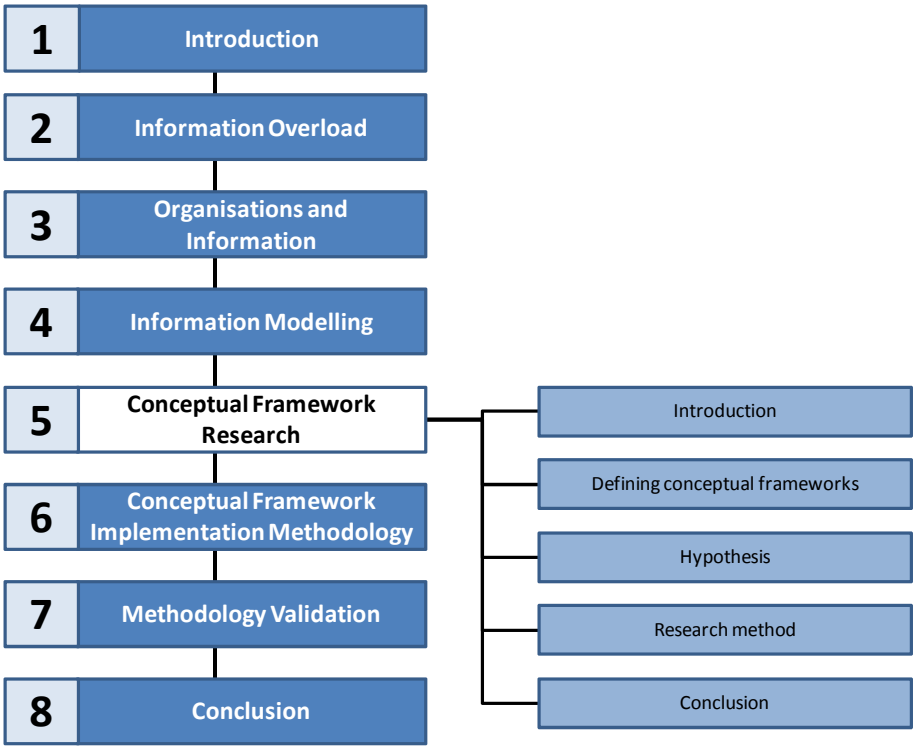


Figure 5.1: Image for navigating thesis and chapter 5

5.2.1 Concept

The Hypothesis Browser website of the Huck Institute for Life Sciences, Penn State University (HyBrow n.d.) defines a conceptual framework as follows:

> 'A conceptual framework is a formal way of thinking (i.e. conceptualising) about a process/system under study.'

It can therefore signify a mechanism or construct used by a person to increase his/her understanding of a certain process, system or (in this case) collection of information. This is however, a very general definition. The conceptual framework concept, on which research is being conducted by the GCC and Indutech, was defined by Uys (2005) as:

> 'A formal model of a given domain, consisting out of the domain entities and the relationships holding among these entities, used for understanding and analysing the domain in question.'

This is still a very general definition which sufficiently defines a number of already existing information modelling approaches (some of which are discussed in Chapter 4). While this is currently a sufficient definition for the purposes of the GCC, a more specific definition is however required for the purpose of this thesis, to distinguish conceptual frameworks from approaches such as concept maps, ontologies and topic maps. A fact which distinguishes the conceptual framework from most information modelling approaches is that in the development of the conceptual framework approach, great emphasis was placed on simplicity. In a conceptual framework, only four concepts are utilised: entities, relations, facts, and types, and the application of the conceptual framework approach is based on certain accepted assertions about these four concepts. The extended definition that will be used in this thesis to describe a conceptual framework is therefore as follows:

'A conceptual framework is a formal model of a given domain, consisting of domain entities and relations that have been identified between these entities. The construction of a conceptual framework is based on only four concepts, namely that of entities, relations, facts and types, which provide the means for improved understanding of the domain in question through visual browsing and navigation of these concepts.'

The concepts of conceptual framework entities, relations, facts and types (within the scope and focus of this thesis) will forthwith be discussed.

- ***Entities***

A conceptual framework *entity* is defined as follows:

An entity is a reference to any specific thing, real or abstract, that forms part of, or has any importance in, a particular information environment. It is represented by a commonly-used label that corresponds to the perception of the real-world aspect which the entity represents.

Entities are the building blocks used to construct a conceptual framework, and typically represents the aspects encountered in a person's (or group of persons') daily environment, for example, "Howard Shore" (which may denote a person), "Budget2005.xls" (which may denote a document), "Procurement" (which may denote a department within an organisation) and "Dubai" (which may denote a geographical location). Additionally, a rather unique feature of the conceptual framework approach which is quite uncommon in other information modelling techniques is the fact that references to real or abstract features of entities that exist in the information domain (i.e. attributes) are considered as entities themselves, rather than as attributes which, in the case of other conventional information modelling approaches, might be seen as information with reduced

importance to the information domain. This means that entity features such as “80” ((kg) i.e. a measure of weight), “15” ((min) i.e. a measure of duration) and “17” ((km) i.e. a measure of distance) are also regarded as entities. This increases the diversity of the environment under consideration, which is useful when searching for specific information.

- ***Relations***

A conceptual framework *relation* is defined as follows:

Any association between two entities represented in a conceptual framework is represented by a relation, which is a text string that states the manner in which the two entities are associated. A relation is not required to be unique to a pair of entities, and may be reused to describe the relation between any number of relevant entity pairs.

Consider the following example: If two entities exist, one denoting a person “Howard Shore”, and one denoting a document “Budget2005.xls”, then a relation indicating that the person “Howard Shore” created the document “Budget2005.xls”, might be “is person who created document” (see Figure 5.2).

Entity	Relation	Entity
Howard Shore	is person who created document	Budget2005.xls

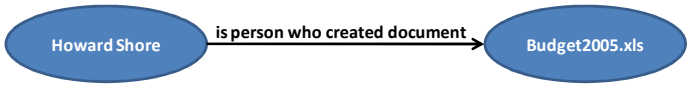


Figure 5.2: An example of a relation between two entities

Relations can be defined very specifically to describe the association between only two entities, or may be defined in a general manner as to be applicable to groups of entities (see also the next section, Types), for instance, in a conceptual framework containing a number of entities denoting persons and documents, the relation “is person who created document” might be applied between various person-document entity pairs, and will successfully convey that specific information.

The construct created through the joining of two entities by means of a relation, is known as a *fact* (defined as follows):

The joining of two entities by means of a relation (i.e. entity-relation-entity) produces the smallest unit of information found in a conceptual framework, known as a *fact*.

When considered on its own, a fact is the smallest unit of information that exists in a conceptual framework, and is deemed to be a truthful representation of some part of the world being modelled

by the conceptual framework. Facts are by implication multi-directional, and can also be grouped together to form *complex facts*. An example of a complex fact may be to state that "Budget2005.xls" (a document) was created by "Howard Shore" (a person) who works in "Procurement" (a department). This fact, which consists of three entities and two relations (and also illustrates the multi-directionality of facts), is actually a combination of the information contained in the two facts: (department) "Procurement" is the department of person "Howard Shore", and (person) "Howard Shore" is the person who created document "Budget2005.xls" (see Figure 5.3).

Entity	Relation	Entity
Procurement	is department of person	Howard Shore
Howard Shore	is person who created document	Budget2005.xls

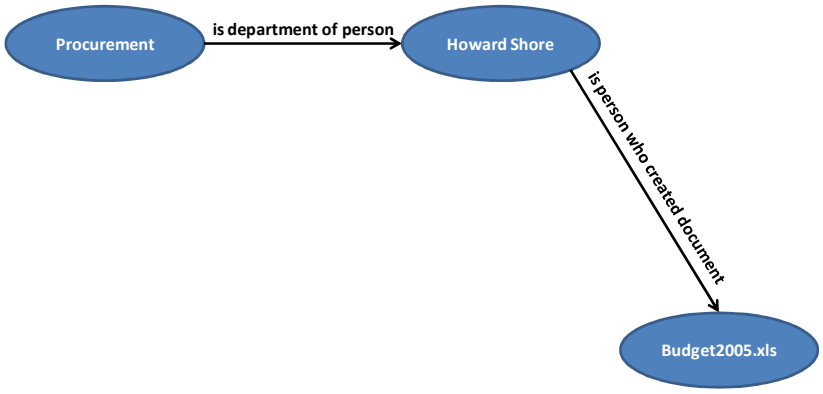


Figure 5.3: An example of a complex fact

- ***Types***

Conceptual framework *types* can be seen as units of classification within an information domain, and is formally defined as follows:

Multiple entities may be grouped together through the relation of each entity to a classification entity – known as a type – by means of an "is a"-relation. Relations can also be specified between types, and are propagated to the children of those types, although the use of inherited relations is not mandatory.

Although not mandatory, types are usually employed to group entities together based on some common characteristic shared between them. As indicated in the definition, a type is identifiable in a conceptual framework by being connected to its children entities through an "is a" relation. For example, a type "Person" can be defined for all entities signifying person-like characteristics, a type "Department" for all department-like entities, and so on (see Figure 5.4).

Entity	Relation	Entity
Procurement	is a	Department
Engineering	is a	Department
Howard Shore	is a	Person
Peter Jackson	is a	Person
Peter Jackson	is a	Manager

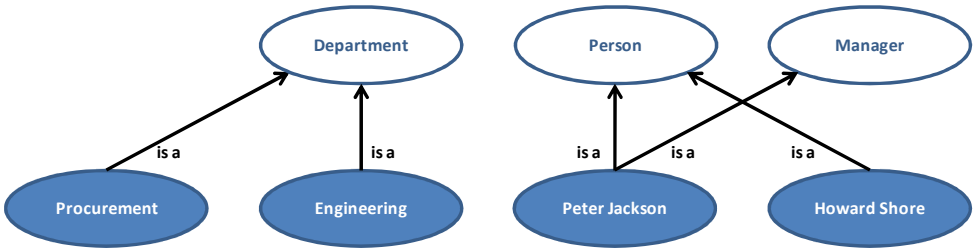


Figure 5.4: An example of types

Types are not mutually exclusive, which means that an entity may belong to more than one type. Relations can be defined between types, which give an indication of associations that can be assumed between different groups of entities. These relations are then also propagated to the children of the types, although the entities are not restricted to the relations of the types they are associated with. This is very apparent when one considers the conceptual framework represented in Figure 5.5.

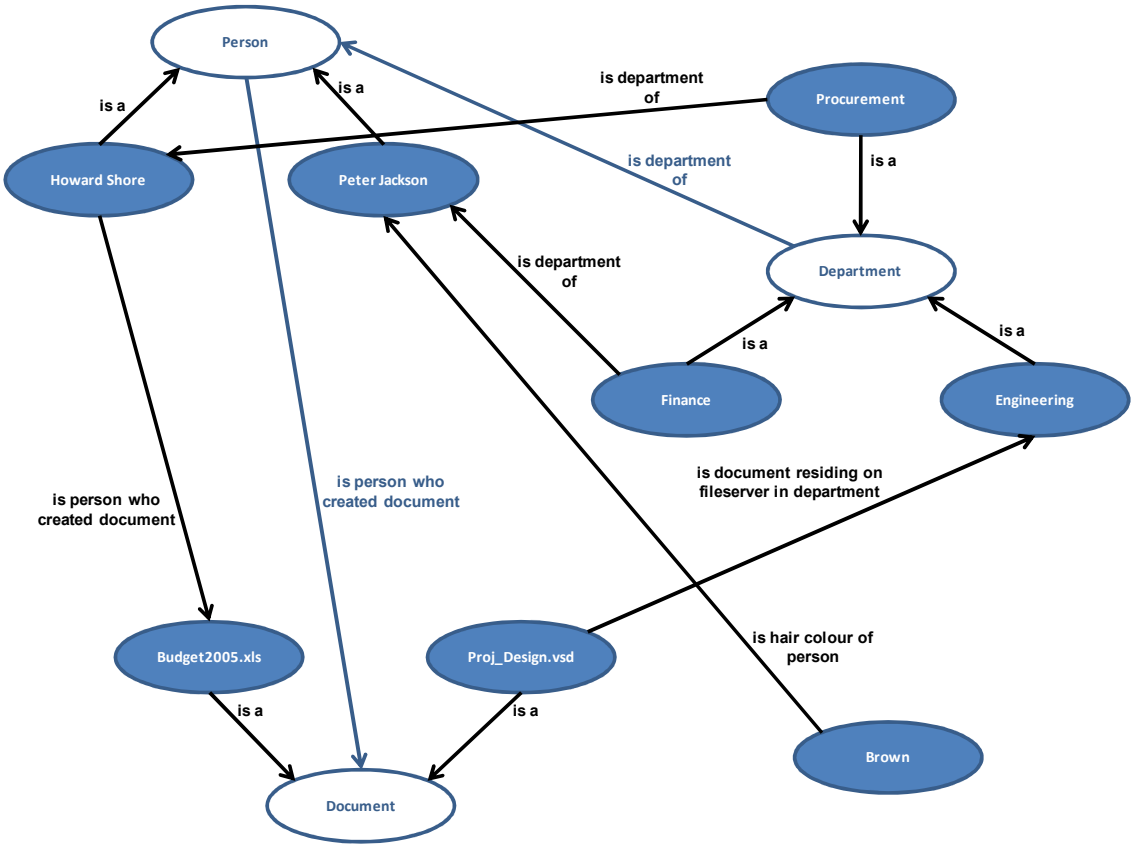


Figure 5.5: An example of a conceptual framework

Although relations are defined to communicate that “Department”-“is department of”-“Person” and “Person”-“is person who created document”-“Document”, there are no documents related to person “Peter Jackson” and no persons related to document “Proj_Design.vsd” and department “Engineering”. Furthermore the relation “is document residing on fileserver in department” is not a relation specified to describe the association between any types and “Brown” is an entity not associated with any type.

5.2.2 Comparison to other information modelling approaches

Conceptual frameworks share some similarities with the information modelling approaches discussed in the previous chapter (Chapter 4), but also differ from each of these in specific ways. When one considers the mechanism and purpose of semantic networks, thesauri, taxonomies, ontologies, folksonomies, concept maps and topic maps, it becomes apparent that conceptual frameworks can almost be seen as a super class of all of these approaches; that each of these approaches are in actual fact a very specific application of the conceptual framework concept, with additional rules or features imposed to allow the approach to better service the specific cause it was developed for.

- In the case of semantic networks, entities (or concepts) and (semantic) relations are employed solely to represent the associations that exist within a specific language, and serve to ensure that a language is understood and utilised in the correct manner. Entity-relation-entity constructs are not used to communicate facts about a certain environment, but rather for the explaining of concepts. Also, even though “is a” relations are used, it is not used to signify grouping by means of a type, but again purely as a means to explain a certain concept.
- A thesaurus is also an approach that is aimed solely for the clarification of a language, but unlike a semantic network, which aims to provide unambiguous explanations of the concepts used in the language, thesauri communicate the associations between the concepts (or words) themselves. A very small set of predefined relations is used (e.g. “broader term”, “narrower term”, “related term”), and again types (i.e. “is a” relations) are not employed, as grouping of concepts does not take place. The entity-relation-entity constructs only communicate linguistic “facts” of the concepts contained in a thesaurus.
- Another approach which makes use of a restricted, predefined relation set is the taxonomy. In the case of a taxonomy, only “is-a” relations are used, aimed solely at communicating the “natural” hierarchy that exists within a group of concepts, with entity-relation-entity constructs only serving to communicate parent-child relations. This differs from the manner in which

conceptual frameworks address this aspect, in that the inheritance of relations defined in a conceptual framework are not mandatory, and a much larger variety of relations are allowed.

- An ontology is one of the approaches that corresponds very closely to the definitions constituting the conceptual framework concept. Although ontology concepts are initially organised in a hierarchical structure, this by no means limits the manner in which different concepts can be related to each other. It therefore utilises the entity-relation-entity construct in much the same way that conceptual frameworks do. An aspect that ontologies employ, which conceptual frameworks do not, is the use of constraints and axioms that restricts the manner in which information is represented in the ontology. This is however a control feature employed to enhance the user-experience of ontologies, and from the point of view of this discussion not seen as part of the core features of ontologies. Ontologies also add an additional dimension to its objects by allowing the definition of attributes that describe the objects. This is not allowed by conceptual frameworks, as the impression is created that attributes are information of lesser value than the concepts they describe, which may result in a biased view on the domain being presented to the user, or that important relations that exist in the domain may be overlooked.
- The folksonomy approach typically makes use of only three fixed types – users, tags and (information) objects. On closer inspection it becomes apparent that these types and entities of these types are related in a fixed manner to any of the other two types and entities of these types (see Figure 5.6).

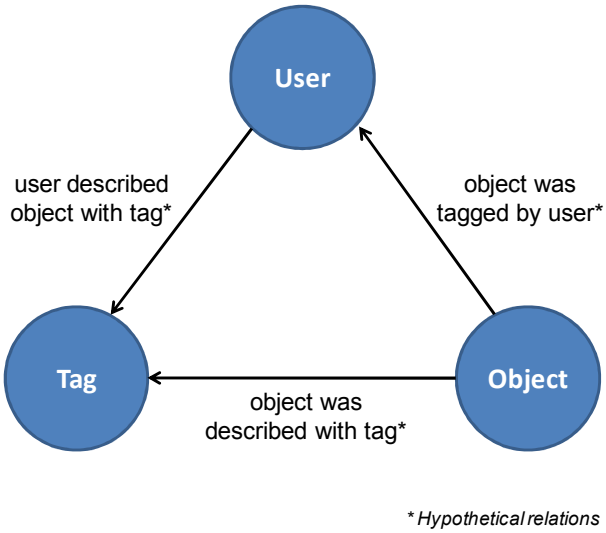


Figure 5.6: Relations that exist between different types in a folksonomy

There therefore exists no freedom in the definition of relationships between types. Also, to express the association that exists between entities of the same type (e.g. between two objects, two tags or two users) in a sensible manner, entities of the remaining types are required to indicate the relation. For example, to express the relation between two objects, one should

consider the tags that are shared between the two objects, as well as the users that have tagged or accessed both of the objects (see Figure 5.7).

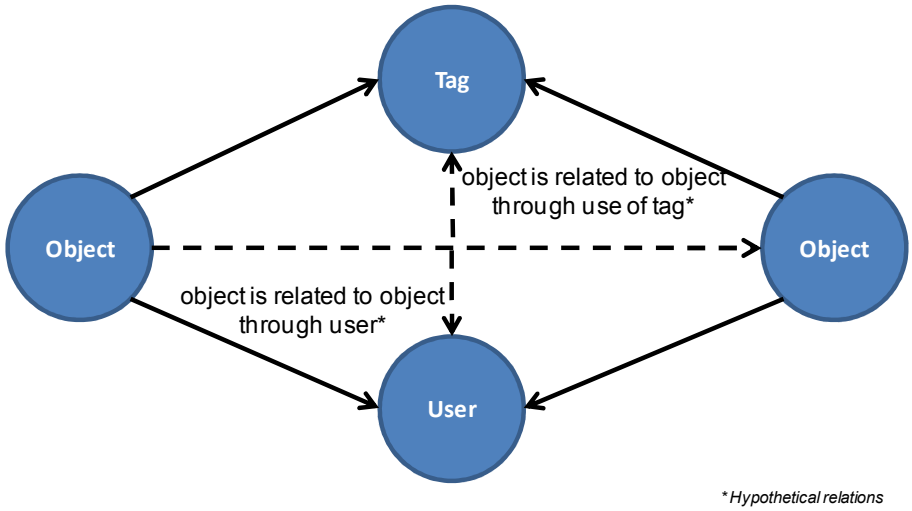


Figure 5.7: Example of relation between types of the same kind in folksonomies

These may be seen as complex facts in the conceptual framework context. Other than that, due to the fact that both the types and relations that exist within folksonomies are fixed, a folksonomy is a very restricted version of the conceptual framework approach.

- At first glance, concept maps seem to be an exact manifestation of the conceptual framework concept, as it seemingly also makes use of the four concepts that forms the core of conceptual frameworks, namely concepts (i.e. entities, see section Entities), relations, propositions (i.e. facts, for both see section Relations) and general concepts (i.e. types, see section Types). There are however subtle differences: Even though concept maps are hierarchical in structure, this does not imply that concept maps make use of types – concept map concepts that are placed higher up in the hierarchy are simply events or objects that represent more general concepts that those on the lower levels. Furthermore, concepts in concept maps do not necessarily represent actual aspects found in an information domain – it merely serves to improve the human understanding of the domain under consideration.

- Of all the approaches discussed in this document, topic maps are possibly the most similar to conceptual frameworks. Entities (or topics), relations (or associations), types and facts are all supported by topic maps, and in some cases the topic map approach even provides more functionality than the conceptual framework approach (e.g. the typing of associations and occurrences). Strictly speaking, this may mean that topic maps can actually be considered a sister class of conceptual frameworks rather than a subclass. It has however been determined that, due to the large variety of functionality offered by the topic map approach, a topic map is much less simple and intuitive to construct and interact with than a conceptual framework, which is a slight advantage that the conceptual framework approach have over that of the topic

map (and especially relevant in the case of users with a limited understanding of information modelling and representation).

For a summarised comparison between the conceptual framework approach and the other information modelling approaches mentioned in this section, see Table 5.1.

Table 5.1: A summarised comparison between conceptual frameworks and other information modelling approaches

Information Modelling Approach	Conceptual Framework Concepts				Differences from Conceptual Frameworks
	Entities	Relations	Facts	Types	
Semantic Network	Yes	Yes	Yes	No	Employed singularly for the clarifying of a specific language, i.e. does not attempt to describe a domain in its entirety.
Thesaurus	Yes	Restricted	Yes	No	Employed singularly to indicate the manner in which terms that exist within a specific language are related, i.e. does not attempt to describe a domain in its entirety.
Taxonomy	Yes	Restricted	Yes	Yes	Employed singularly to indicate the "natural" hierarchy that exists within a group of concepts. Mandatory inheritance of relations.
Ontology	Yes	Yes	Yes	Yes	"Soft" hierarchical structure. Constraints and axioms restrict the information represented by ontology. Allows definition of attributes that describe ontology objects.
Folksonomy	Yes	Restricted	Yes	Restricted	Fixed types and relations.
Concept Map	Yes	Yes	Yes	No	Does not make use of types – only distinguishes between more general and more specific terms. Is not necessarily aimed at describing a domain, but rather at improving understanding.
Topic Map	Yes	Yes	Yes	Yes	Great complexity which impacts on ease of construction and interaction.
Conceptual Framework	Yes	Yes	Yes	Yes	

5.2.3 Application

Through application in different scenarios at both the GCC and Indutech, it has become apparent that conceptual frameworks are very useful in representing information in a manner that reduces the time required to gain a full understanding of certain information.

When considering the modelling of information as a conceptual framework, there are two aspects to be considered that influences the manner in which conceptual framework construction takes place. The

first is the amount of structure that is inherent in the information to be modelled. This could range from extremely structured, like transactional databases, to extremely unstructured, like the content of text documents and the knowledge that resides in people's minds. Based on the structure of the information, the second aspect to be considered is the potential that exists for automating the conceptual framework construction process. Conceptual framework construction is typically performed in the following ways:

- The modelling of unstructured information in a conceptual framework is most often performed manually, through which entities, relations and types are created and added to the conceptual framework one at a time. This is especially the case when the main aim of the conceptual framework is the capturing of people's knowledge and experiences. There do however exist potential for the automation of the process through which unstructured information that reside in text documents specifically can be extracted and represented as a conceptual framework. Through research currently being conducted by the GCC and Indutech, software techniques are being developed by which entities, relations, facts and types that exist within text documents can be automatically identified and extracted, with which a conceptual framework can be constructed.
- The modelling of structured information, on the other hand, is usually performed in (but not restricted to) an automated manner. Through inspection, entity types that exist within information sources (e.g. person, document, etc.) are identified, and relations that can be inferred between these types are defined. Through techniques such as SQL (Structured Query Language) querying, information in structured information sources can be transformed into conceptual framework facts, and added in this format to the conceptual framework. Because conceptual frameworks can be created and updated in this manner, the concept can furthermore be extended to perform business intelligence functions, enabling a person to see the state of an organisation in a single structured view, which should improve understanding and provide the means for accurate decision making.

5.2.4 Visualisation

As is the case with all the other information modelling approaches discussed in the previous chapter (Chapter 4), an aspect that greatly contributes to the conceptual frameworks' ability to successfully communicate information is the fact that it lends itself to the graphical representation of the information. The value of visualisation techniques in assisting persons in their comprehension of information has been advocated for some time, and although there are some differing points of view on its efficacy in general (Chan 2001), great value has been found in representing information as graphs, utilising nodes and arcs to convey information (Celko 2006, Cogito White Paper n.d.). The

representation of information as graphs formed the basis of the development of Organon, a software application that was developed by Indutech as a tool for creating, navigating and editing conceptual frameworks. Organon visualises conceptual frameworks as network graphs, in which entities and types are represented as nodes and relations are represented as arcs. Currently, it is the only known software application through which conceptual frameworks can be created, updated and navigated, and is employed by the GCC and Indutech for conceptual framework applications and research. A screenshot of a conceptual framework visualised by Organon can be seen in Figure 5.8.

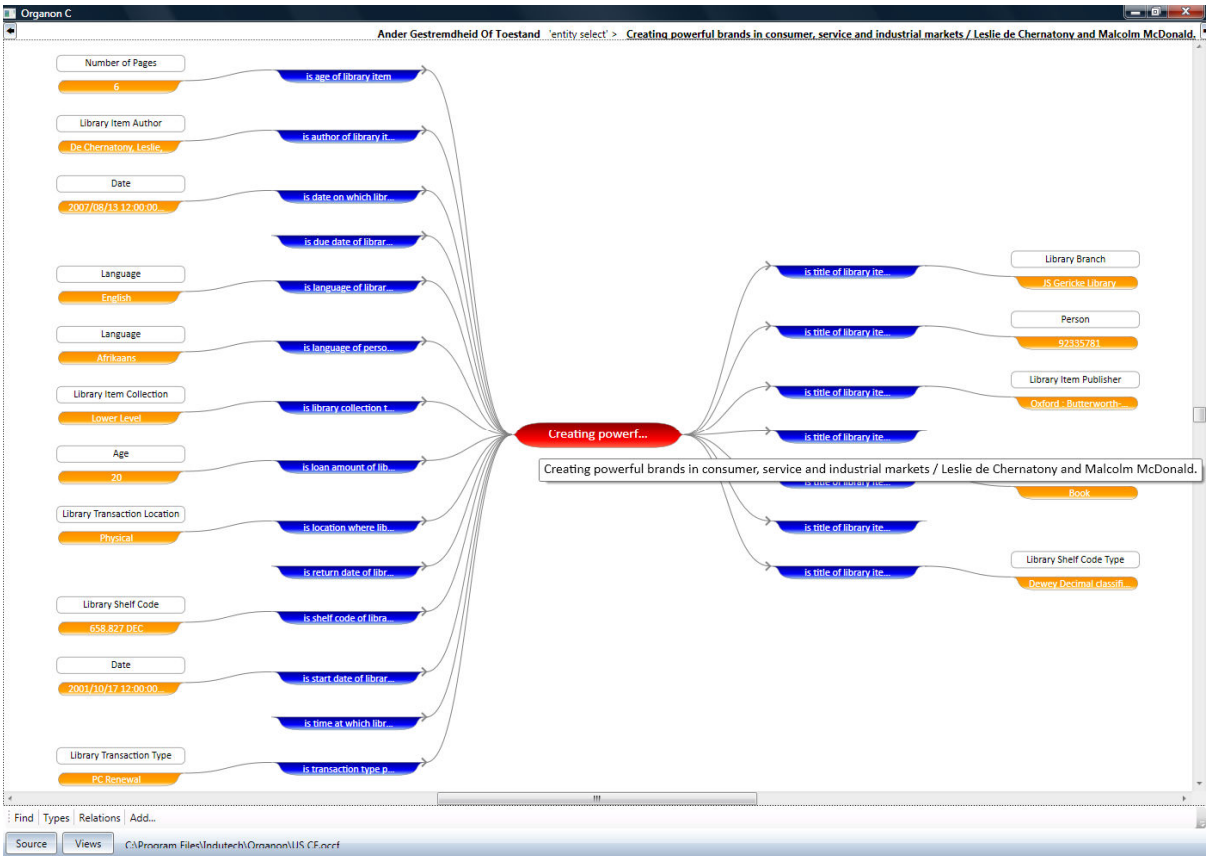


Figure 5.8: Screenshot of Organon

Organon was used in the research that is discussed by this document. Information on the manner in which it was applied can be found in Chapter 7, and other information on the software itself can be found in Appendix A.

5.2.5 Conceptual framework research

The concept of conceptual frameworks introduced in this chapter is currently still a largely experimental approach for modelling information in a very simple and unconstrained manner. A number of research efforts are currently underway that investigate the different manners and scenarios in which conceptual frameworks can be applied, the value that will be experienced by users that interact with conceptual

frameworks, as well as its overall value to the fields of information management and knowledge management. Current conceptual framework research efforts are effectively divided into two main streams, namely the representation of structured information and the representation of unstructured information for improved understanding (see Figure 5.9).

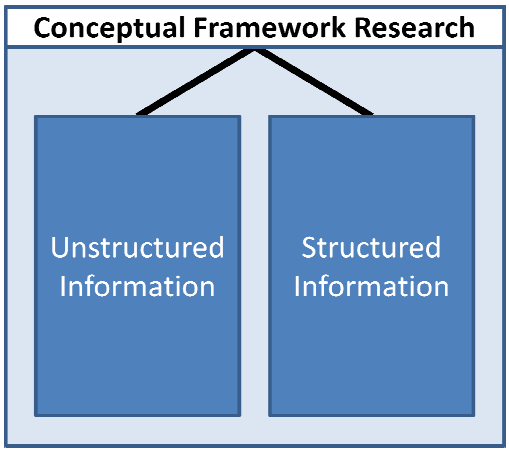


Figure 5.9: Conceptual framework research streams

The research presented in this document focuses on the application of conceptual frameworks to *structured information*, and contributes specifically to research aimed at establishing the worth of conceptual frameworks as decision-making tools within organisations. The hypothesis and method of research for this research is provided in the remainder of this chapter, which is followed by the research outputs (Chapter 6), information on case studies that were performed as part of the research (Chapter 7) and conclusions reached based on the research (Chapter 8).

5.3 Hypothesis

Research on conceptual frameworks has progressed to a stage where its applicability in organisations should be investigated. As a starting point for this investigation, it was deemed useful to develop a methodology through which a conceptual framework can be implemented in an organisation. Such a methodology would be required to address the following aspects of conceptual framework application:

- Development of a conceptual framework specification based on stakeholder inputs and analyses of organisational information sources;
- Creation and population of the conceptual framework with organisational data;
- Management of the information system issues when implementing the conceptual framework as an operational business tool (e.g. bandwidth, network privileges, storage, etc.);

- Management of user issues, such as security and training; and
- Establishment of a strategy for growing the conceptual framework in the organisation.

Once developed, the completeness and relevance of the methodology would be assessed through testing of the following hypothesis:

Hypothesis:	**Information represented by a conceptual framework implemented through the use of the developed conceptual framework implementation methodology, offer a truthful view on information found in an organisation.**

5.4 Research method

The main outcome of the research was a tried-and-tested methodology developed for the implementation of a conceptual framework in an organisation. Indutech (in conjunction with which the research was conducted) furthermore experienced additional benefit from the project, in that it was provided with opportunities through which Organon, the conceptual framework software tool developed by Indutech for the viewing and editing of conceptual frameworks (see section 5.2.4, Chapter 7 and Appendix A), could be tested and improved. This is however not a formal outcome of the research, and will therefore not be discussed.

The process that was followed in the development of the conceptual framework implementation methodology took place in the following manner (see Figure 5.10):

1. An initial version of the methodology was developed, based on knowledge gathered on conceptual frameworks through research conducted by the GCC and Indutech.
2. Once there was agreement on the completeness and usability of the methodology, a case study was conducted in order to ascertain the effectiveness of the methodology in implementing a conceptual framework at an organisation. Indutech was selected as the case study site, due to the fact that it is already heavily involved in conceptual framework research, and therefore provided a benign testing environment where most of the methodology's creases could be smoothed out.
3. After concluding the case study, all information gathered during the course of the implementation was used to refine the methodology and make improvements to the Organon software.
4. Once the modifications were made, a second case study was conducted to further refine the methodology. The site selected for the second case study was the Information Technology division of Stellenbosch University.

5. Upon completion of the second case study, observations and lessons learned resulted in another round of improvements made to the methodology and to Organon.

6. The second round of improvements ultimately yielded a conceptual framework implementation methodology that was deemed fit to be utilised as a tried-and-tested tool for the implementation of conceptual frameworks in organisations. This was done based on feedback from the GCC and Indutech research personnel who were involved in the case studies, and the organisational users of the Organon software at each of the case study sites (who also had to decide whether the hypothesis on which this research effort was focused was true or not).

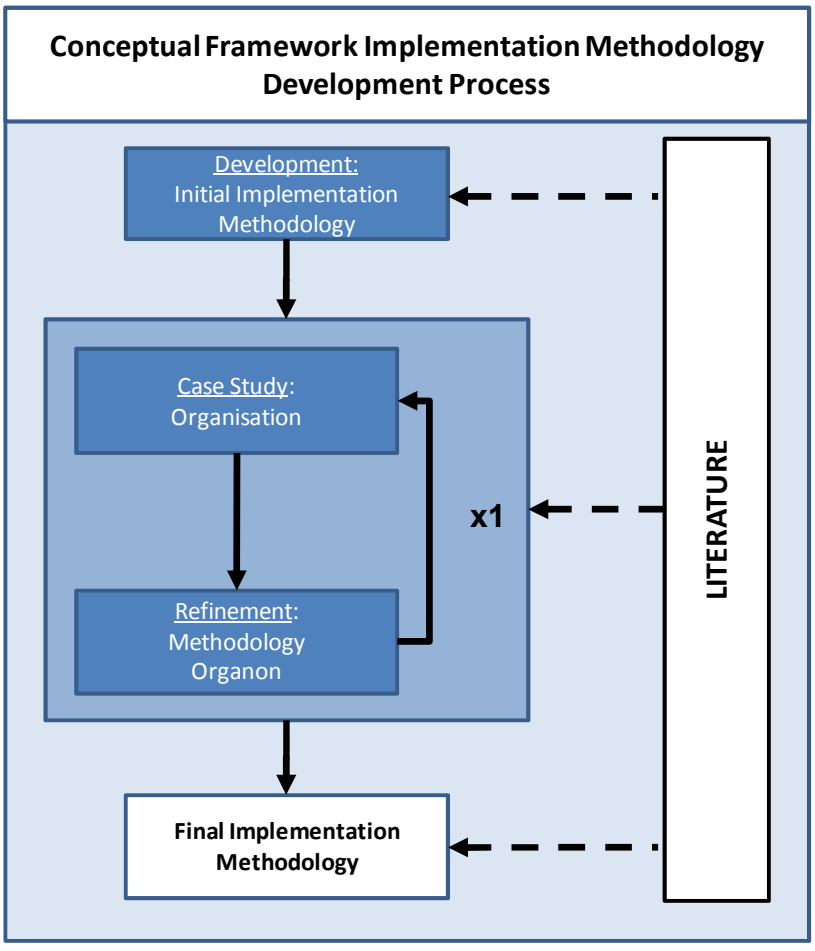


Figure 5.10: Conceptual framework implementation methodology development process

The use of literature served a supporting role throughout the entire conceptual framework methodology development process, and was specifically employed in the conducting of a literature study on information overload, information use in organisations, and other existing information modelling approaches (see Chapters 1 - 4), as well as in the validation of the methodology through comparison with other similar information modelling approach implementation methodologies (see Chapter 7).

Information on the two case studies that were performed, together with lessons learned through the case studies and the subsequent changes made to the methodology and software are discussed in more detail in Chapter 7.

5.5 Conclusion

This chapter introduced the concept of information modelling by means of conceptual frameworks, an approach which is seemingly able to provide a combination of flexibility and simplicity that cannot currently be matched by any of the more well-known information modelling approaches that already exist. Various research efforts are currently underway through which the value and usefulness of the conceptual framework concept are investigated, and the extent of its abilities are established. This document reports on a project that was conducted as part of the stream within conceptual framework research that focus on the representation of structured organisational information as a conceptual framework, and produced as an output a methodology through which a conceptual framework can be implemented in an organisation. The methodology and the development thereof are discussed in the following two chapters.

Chapter 6 – Conceptual Framework Implementation Methodology

6.1 Introduction

The previous chapter (Chapter 5) introduced the concept of a conceptual framework based approach for the representation of information. It further discussed a research project that was conducted in the field of conceptual frameworks (focused specifically on the representation of *structured* information) through which a methodology for the implementation and maintenance of a conceptual framework in an organisation was developed. This chapter introduces and discusses this methodology. An image depicting the structure of this chapter is displayed in Figure 6.1.

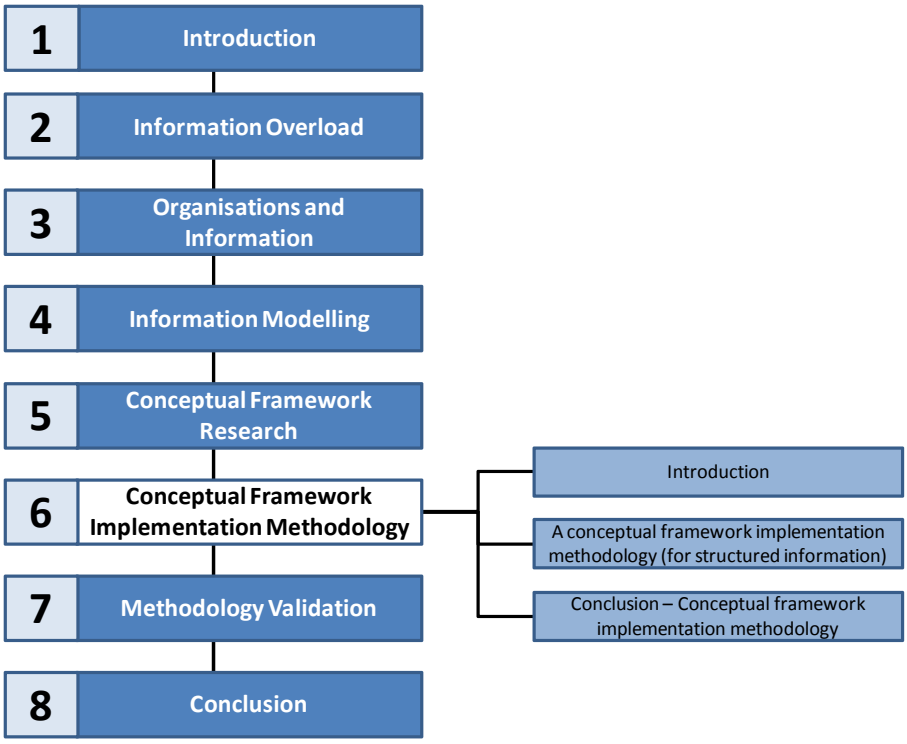


Figure 6.1: Image for navigating thesis and chapter 6

6.2 The proposed methodology[17]

The conceptual framework implementation methodology for structured information (see Figure 6.2) is a top-down process that is comprised of five phases, namely Planning, Analysis, Development, Implementation and Maintenance. Each of the phases consists of a number of activities that should be executed in order to ensure the quality of the conceptual framework and the eventual success of the implementation. For each of these phases and activities, the following aspects will be addressed in this discussion of the implementation methodology:

- The rationale behind the phase or activity;
- Prerequisites or inputs of the phase or activity;
- The execution of the phase or activity; and
- Outputs of the phase or activity.

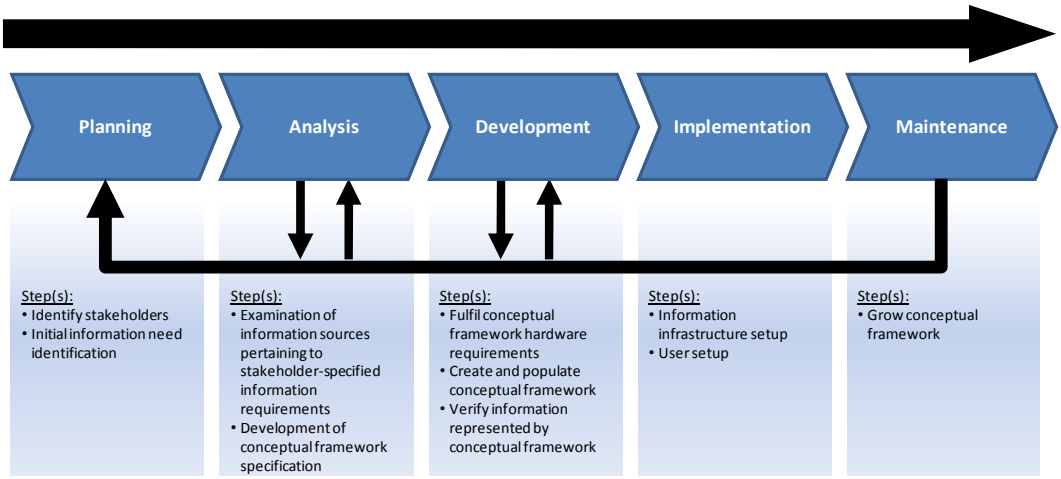


Figure 6.2: The Conceptual Framework Implementation Methodology for Structured Information

The phases of the methodology will forthwith be discussed.[18]

[17] *Or more formally: "A conceptual framework implementation methodology (for structured information)".*

[18] *The numbering system that will be used in the methodology discussion will differ from the numbering system employed in the rest of this thesis. Each phase will simply be designated by a letter based on its position in the sequence of phases that constitute the implementation methodology (e.g. "A" = Planning, "C" = Development, etc.). Each activity will have a designator composed of the indicator of the phase it forms part of, and a number that identifies the activity within the particular phase, based on its order of execution. For example, "C2" indicates the "Create and populate conceptual framework" activity of the Development phase.*

A – Planning phase

Designation:	**A**	*Title:*	**Planning Phase**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

The first phase of the implementation methodology is aimed at the identification and nomination of stakeholders within the organisation who will be involved in the implementation. The establishing of the implementation scope should also take place, and is mainly determined by the organisational information domain from which the conceptual framework is to be created.

Rationale:

The planning phase is required to ensure that the conceptual framework implementation efforts are focused on the right areas within the organisation, and to establish who the implementation's decision-makers in the organisation will be.

Prerequisites:

- A need for an integrated view on organisational information[19]

Execution:

- **Activity A1 – Identify stakeholders**
- **Activity A2 – Initial information need identification**

[19] *This is not a "formal" methodology prerequisite, but alludes to the fact that a need identified in the organisation for an integrated view on its information gave rise to the conceptual framework implementation.*

Outputs:

- A **"Conceptual Framework Implementation – Planning phase"** document which should be utilised as an information source and control measure in the management of the implementation. The document should contain the following information (which originated from the activities that constitute the Planning phase):
    - The names and contact details of the stakeholders in the conceptual framework implementation (see Activity A1);
    - The rationale behind the conceptual framework implementation (see Activity A1);
    - The information domains within the organisation for which the conceptual framework will be constructed (see Activity A1);
    - Any agreements reached with the information infrastructure stakeholders regarding the organisational information source access, network usage and data extraction that will be required for the conceptual framework implementation to succeed (see Activity A1); and
    - All information requirements provided by stakeholders, which are the main needs that the conceptual framework implementation should aim to address (see Activity A2).

- ***Activities***

Designation:	**A1**	*Title:*	**Identify stakeholders**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

This activity facilitates the identification of stakeholders in the conceptual framework implementation methodology, who are responsible for providing input into the design, development and implementation of the conceptual framework, and possess the authority and the means for effectively removing any obstacles that may be encountered during the implementation process.

Rationale:

For effective decision making during the conceptual framework implementation, it is essential to identify persons in the organisation at whom the implementation is aimed, and to stay in contact with them throughout the entire process. This will ensure that the implemented conceptual framework effectively addresses the specific needs of its users.

Prerequisites:

- Project initiation[20]

Execution:

Two distinct types of stakeholders are required for the successful implementation of a conceptual framework; managerial stakeholders and information infrastructure (system) stakeholders.

[20] *This simply alludes to the fact that the implementation had to already have been kicked off for this activity to be executed.*

Identify management stakeholders

The management component of the implementation stakeholders should be made up of the following role-players within the context of conceptual framework implementation:

- Project champion: The project champion is a person within the organisation tasked with promoting the benefits of pursuing the project, while also justifying the investment made by the organisation. This should be a person who is truly confident in the benefits that would result from the project, and have the necessary skill and mandate for effective advocacy of the project, as well as for effective roadblock removal (Withrow 2004). Once instated, the project champion should identify (alone or with the assistance of others) possible areas within the organisation that will provide a suitable domain for a conceptual framework implementation[21]. It would furthermore be beneficial if these areas required an integrated view of organisational information, as this will have a positive effect on the buy-in from other stakeholders in the implementation and the end users of the conceptual framework.
- Managers and decision makers within the implementation domain: Once the areas within the organisation that have been deemed most suitable for the implementation of a conceptual framework have been identified, managers and other decision makers within these areas together with other experts within the organisation who can make a useful contribution to the implementation, should be identified and included as stakeholders of the implementation process. These persons are responsible for determining the scope and deliverables of the implementation, and act as verifiers when sign-off of the information represented by the conceptual framework is required. A requirement of these stakeholders is that they have an adequate understanding of the interrelatedness of information inside the organisation, and are also convinced (or have been convinced by the project champion) of the need for an integrated view of information within the organisation.

Identify information infrastructure stakeholders

As a conceptual framework implementation requires the utilisation of an organisation's information infrastructure, it is imperative that representatives of this function within the organisation are also included as stakeholders in the implementation. These persons should naturally then also have sufficient influence in the organisation to authorise the interaction with information that will be required by those persons performing the conceptual framework implementation. Buy-in from these representatives is of the utmost importance, and if this cannot be obtained, the project champion may be required to escalate this issue to higher levels of management in order to ensure the cooperation of the organisation's information infrastructure department(s). At this stage however, preliminary backing from these representatives is sufficient, as the total

[21] *As this methodology advocates an iterative implementation process, it is not essential that the entire organisation is initially represented by the conceptual framework. An incremental implementation further allows for teething problems and difficulties with regards to internal organisational complexity to be experienced on a smaller scale, which makes it simpler and speedier to correct. This will be discussed in more detail in section E – Maintenance phase.*

extent of the implementation have not yet been determined, and will only be known after the Analysis phase (see section B – Analysis phase). Nevertheless, in backing the implementation, representatives must be made aware that the establishing of an organisational conceptual framework will require information to be extracted from different data sources in the organisation, and that allocation of organisational network bandwidth and scheduling of information extraction from organisational information sources will be required for the development and operation of the organisational conceptual framework.

Outputs:

- Information to be included in the **"Conceptual Framework Implementation – Planning phase"** document (indicated as the output for phase A – Planning phase). The information required is:
    - The names and contact details of the stakeholders in the conceptual framework implementation;
    - The rationale behind the conceptual framework implementation;
    - The information domains within the organisation for which the conceptual framework will be constructed; and
    - Any agreements reached with the information infrastructure stakeholders regarding the organisational information source access, network usage and data extraction that will be required for the conceptual framework implementation to succeed.

Designation:	**A2**	*Title:*	**Initial information need identification**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

This activity entails the collection of high-level stakeholder information needs, which acts as a scope definition for the conceptual framework implementation.

Rationale:

The specification of the implementation scope provides a control measure for the process, and will help to ensure that all efforts are focused on the needs of the stakeholders, and not on aspects of the organisation that are irrelevant.

Prerequisites:

- Conceptual framework stakeholder contact information (Activity A1)
- Information on agreements reached (Activity A1)

Execution:

Once the stakeholders of the conceptual framework implementation have been identified (Activity A1), they should specify the information to be included in the conceptual framework. Although this activity can be seen as the scope definition of the conceptual framework implementation, it should be understood that any definition of scope that is attempted at this stage would be done amidst great uncertainty about the amount of effort and time that would be required to perform the implementation. The aspects of the implementation for which sufficient information is available should therefore be defined in an acceptably rigid manner, while aspects which are more vague may be defined in a more general manner and remain so until later in the implementation process when the uncertainty surrounding these aspects has been reduced.

The one aspect that must be extremely clear and on which the stakeholders must decide first, is the domains

of information within the organisation that will form part of the implementation. This is simply a safety precaution and control measure to prevent the implementation from unnecessarily increasing in time and effort (i.e. so-called “scope-creep”). Other than that, it is sufficient (and also preferable) if stakeholders give their information requirements in the form of “fields of interest” rather than as fixed sets of data (e.g. "I want to see account information of clients" as opposed to "Show me a client's name, address, birth date, monthly income and risk to the business”). This generalisation of organisational information requirements will increase the objectivity of the conceptual framework, and prevent it from only showing specific persons’ views of the organisational information. Once the information requests from each of the stakeholders has been obtained, requests that are similar in nature should be grouped together, which will result in a more consistently and coherently defined information need.

Outputs:

- Similar to Activity A1, the output of this activity is information to be included in the **“Conceptual Framework Implementation – Planning phase”** document (indicated as the output for phase A). The information required is:
  - All information requirements provided by stakeholders, which are the main needs that the conceptual framework implementation should aim to address.

B – Analysis phase

Designation:	**B**	*Title:*	**Analysis Phase**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

The main objective of the second phase of the implementation methodology is the analysis of information sources pertaining to the implementation domain that was specified in the Analysis phase (see section A – Planning phase), and from which a specification will be developed for the organisational conceptual framework.

Rationale:

Through the Analysis phase, an organisational conceptual framework specification is developed which is used as the basis for the ensuing conceptual framework development and implementation.

Prerequisites:

- Initial information needs as identified by stakeholders (Activity A2)
- Permission of information infrastructure stakeholders to conduct analyses of organisational information and information structures (Activity A1)

Execution:

- **Activity B1 – Examination of information sources pertaining to stakeholder-specified information requirements**
- **Activity B2 – Development of conceptual framework specification**

Outputs:

- The following information that has been generated through the activities that constitute this phase, should be captured in a **"Conceptual Framework Implementation – Analysis phase"** document, which will function as the design of the implementation:
    - The organisational information sources that were identified to provide the source information for the conceptual framework implementation, together with information on the following aspects of these information sources:
        - The content and structure of the information sources;
        - The manner in which the information sources are related;
        - The manner in which the information sources satisfy stake holder information requirements; and
        - The frequency of change in the information sources. (see Activity B1);
    - The conceptual framework types that were identified from the information sources, together with the location of the information represented by these types (preferably down to table-level in the case of a database) and their frequency of change (see Activity B2);
    - The relations that were identified between the conceptual framework types (see Activity B2); and
    - The final type-relation structure of the conceptual framework after refining, also indicating the manner in which types were grouped together to create an integrated view on the organisational information (see Activity B2).

- ***Activities***

Designation:	**B1**	*Title:*	**Examination of information sources pertaining to stakeholder-specified information requirements**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

This activity focuses on evaluating the content and structure of the organisation's information sources and determining in what manner the stakeholders' information needs can be satisfied by this information.

Rationale:

Information on the structure and content of the organisation's information sources will be used during the Development phase (see section C – Development phase) of the implementation, specifically when the organisational conceptual framework is populated with information from the specific information sources.

Prerequisites:

- Initial information needs as identified by stakeholders (Activity A2)
- Permission of information infrastructure stakeholders to conduct analysis of organisational information and information structures (Activity A1)

Execution:

Identification and obtaining of access to relevant information sources

The domains identified by the stakeholders for the conceptual framework implementation (see Activity A2) contain certain information sources which could sufficiently address the information requirements they have provided. Discussions with the database and other information infrastructure administrators of these domains are required in order to identify the organisational information sources that would best address the needs

defined by the stakeholders.

Once the applicable information sources have been identified, it is essential that the required security privileges for these sources are obtained in order to gain access to the information that is required for the analysis. Normally, the information infrastructure stakeholders of the implementation would have sufficient authority to grant these privileges, but if not, the project champion should be employed to expedite or even escalate this issue, as it will impact directly on the continuance of the conceptual framework implementation. Should it happen that no sufficient information sources can be identified for the analysis and subsequent construction of the conceptual framework, the Planning phase (see section A – Planning phase) may have to be revisited to adjust the information domain(s) for which the conceptual framework should be constructed, or else a methodology for conceptual framework construction focusing on unstructured information should rather be considered.

Analysis of relevant information sources

In this activity, the main focus is placed on investigating the content that is contained in each of the identified information sources, as well as the structure (if any) in which the information is stored. An approach that works effectively, particularly in the case of relational databases, is the utilisation of entity relationship diagrams (ERDs, see also section 4.2 and Chen 1976) and data dictionaries (which lists the data terms or entities that exist within the organisation's information sources and the meaning that these terms have for different people in the organisation (Kendall & Kendall 2002)). In most cases ERDs and data dictionaries will already exist for organisational information sources, but where this is not the case, ERDs and data dictionaries (or any other means which will allow a person to understand the internal structure of an information source and the meaning of its terms or entities) will have to be compiled to ensure that an adequate description of the information source is available for use in this and the next phase (see section C – Development phase) of the conceptual framework implementation process.

Besides understanding each of the information sources on its own, it is also essential to acquire an understanding of the manner in which the various information sources relate to one another. It is important to understand the manner in which changes in any information source influence another, and to what extent overlapping and duplication of information exists between the information sources in question. This information plays a crucial role in the generation of a unified view of organisational information, which is to be generated in the Development phase.

Apart from the content, structure and interrelations of the information sources in question, there are two more aspects that should be addressed as part of the preparation for the Development phase, as well as for the Implementation phase (see section D – Implementation phase) of the conceptual framework implementation process. The first is to identify the specific information within the analysed information domain that fulfils the information needs expressed by the stakeholders. The reason for this is, because the analysis and development of a conceptual framework entails the interaction with very large amounts of

information, it can very easily happen that the conceptual framework is implemented without ensuring that the stakeholders' information needs are met. It therefore serves as a control measure which prevents those persons involved in the implementation from straying off course. The second aspect to be addressed is the gathering of information on the typical frequency of change within the information sources in question. This information will become quite important in the Implementation phase of the conceptual framework implementation process, as it will influence how regularly the information contained in the developed conceptual framework have to be updated in order to provide an accurate and current view on the domain in question.

Outputs:

- Information to be included in the **"Conceptual Framework Implementation – Analysis phase"** document (indicated as the output for phase B – Analysis phase). The information required is:
    - The organisational information sources that were identified to provide the source information for the conceptual framework implementation, together with information on the following aspects of these information sources:
        - The content and structure of the information sources;
        - The manner in which the information sources are related;
        - The manner in which the information sources satisfy stake holder information requirements; and
        - The frequency of change in the information sources.

Designation:	**B2**	*Title:*	**Development of conceptual framework specification**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

Through this activity, a conceptual framework specification (in the form of a type-relation structure) is created from which the planned conceptual framework will be developed and implemented.

Rationale:

The specification acts as a blueprint of the conceptual framework and is required to ensure that the conceptual framework that is implemented conforms to the needs of the implementation stakeholders.

Prerequisites:

- Information on the structure and content of the information sources in the conceptual framework implementation domain that can sufficiently address the information needs specified by the implementation stakeholders, as well as information on the manner in which these information sources are related (Activity B1)

Execution:

Identify types from information sources

The concept of conceptual framework types (defined in section 5.2.1) is useful for grouping conceptual framework entities that share some similarity. It enables persons interacting with the information in the conceptual framework to locate sought-after information quite rapidly, as it provides them with the ability to locate information by (usually) focusing on certain general characteristics of the information. It also provides a very simple means of creating a conceptual framework from structured information, as such groupings of information normally also exist within this environment. If one considers databases for instance, it should be

quite clear that the field names in a database table typically denote a common characteristic of the values in the database table that are associated with those fields.[22] For conceptual framework development it is necessary to first identify all the types that exist within the information sources for which associated types can be created in the conceptual framework. This identification will typically take place by scrutinising the ERDs and data dictionaries gathered and/or developed during the analysis that was performed on the organisational information sources performed in the first part of the Analysis phase (Activity B1).

To illustrate, consider a database example where a hypothetical database table (shown in Table 6.1) contains miscellaneous employee information.

Table 6.1: Example of utilising database fields as conceptual framework types

Emp_ID	Emp_Name	Emp_Birthdate	Emp_Address_Street	Emp_Address_Number	Emp_Address_Town	Emp_Department_FK
1	Howard Shore	1946/10/18	First Street	24	Ontario	1
2	Peter Jackson	1961/10/31	Valhalla Avenue	13	Pukerua Bay	2

"Howard Shore" and "Peter Jackson" are both associated with the database field **"Emp_Name"**, which obviously denotes that the values are employee names, and that the field stores employee names. A type "Employee" (or "Employee Name") can therefore be created in the conceptual framework, under which the names of all persons in the system can be grouped if they are employees of the organisation. Furthermore, if it is decided that a person in the organisation is signified by his/her name, the type "Employee" could possibly not only represent employee names, but the actual employees themselves[23]. More types that can be identified from the database table in question are **"Emp_Birthdate"**, **"Emp_Address_Street"**, **"Emp_Adress_Number"**, and **"Emp_Address_Town"**. Therefore, if one takes a more general view on the manner in which the information contained in database tables are transferred to the conceptual framework, it can be seen that types are typically identified by placing specific focus on the vertical relations that already exist in the tables (see Figure 6.3).

	Field_1	Field_2	Field_3	…	Field_m
Record_1	Value_11	Value_12	Value_13	…	Value_1m
Record_2	Value_21	Value_22	Value_23	…	Value_2m
Record_3	Value_31	Value_32	Value_33	…	Value_3m
⋮	⋮	⋮	⋮		⋮
Record_n	Value_n1	Value_n2	Value_n3	…	Value_nm

Figure 6.3: Identification of information source types through the vertical investigation of a database table

A special case **"Emp_Department_FK"** (another field of the database table under consideration, Table 6.1)

[22] *This is obviously just one way to identify types in an organisational information domain. Various other techniques can also be applied, for instance, filtering on certain values in database fields, e.g. "Organisation Employees" = all people having "@organisationname" as part of their e-mail address.*

[23] *Such decisions should however be made during the optimisation of the type-relations structure which forms part of the conceptual framework specification (see **Refine type-relation structure**)*

indicates a link between the employee database and a database table containing information on the organisation's departments. In this case the actual values of **"Emp_Department_FK"** are not important, but rather the relation between the values in the two tables, as this will guide those persons performing the analysis in finding the information source that contains specific department information which should be associated with the employee information.

Develop default relations between types

Once the types that exist in the defined information sources have been defined (see previous section) default conceptual framework relations (defined in section 5.2.1) should be defined between these types, as this will provide users with the means to access information through connections already existing within the information sources. These relations will most probably correspond very closely with the manner in which the types are already related in the information source (and will possibly not be very intuitive for some users), but will provide a starting structure through which all information can be accessed by the conceptual framework users. Once the conceptual framework has been implemented and users start to interact with it, they can define their own relations which will enable them to interact with the information in a much more comfortable manner.

Returning to the example discussed in the previous section (focusing on the database table shown in Table 6.1), a relation ***"was born on date"*** may be defined between type **"Emp_Name"** and **"Emp_Birthdate"**, which will yield the conceptual framework fact (defined in section 5.2.1) **"Emp_Name"**-***"was born on date"***-***"Emp_Birthdate"***. Other possible facts that may be defined are:

- **"Emp_Name"**-***"lives in street"***-**"Emp_Address_Street";**
- **"Emp_Name"**-***"lives in town"***-**"Emp_Address_Town"**;
- **"Emp_Name"**-***"works in department"***-**"Dept_Name"** (remember the foreign key example in the previous section); and even
- **"Emp_Birthdate"**-***"is the birthdate of an employee living in town"***- **"Emp_Address_Town"**;

It can be postulated that the number of possible relations that can initially be defined between the different types that were extracted from the organisational information sources is expressible as the square of the number of types, as each entity is potentially related to itself as well as to all the other types that were defined, i.e.

$$\#\text{ of relations} = n^2, \text{where } n = \#\text{ of types}.$$

This is however a misconception, as it presupposes that relations are unidirectional, i.e. A → B ≠ B → A. As this is not the case (i.e. A → B = B → A), the maximum number of possible relations that can initially be defined between the different types that were extracted from the organisational information sources, are

actually expressible as follows:

$$\#\ \text{of relations} = \tfrac{1}{2}(n^2+n), \text{where } n = \#\ \text{of types}.$$

Although the growth in the maximum number of relations is much slower than would have been the case were the number of relations expressed as the square of the number of types (see Figure 6.4), it can be seen that even a moderately sized conceptual framework could contain a very large number of relations, which will have a significant impact on the amount of effort required to create a conceptual framework specification.

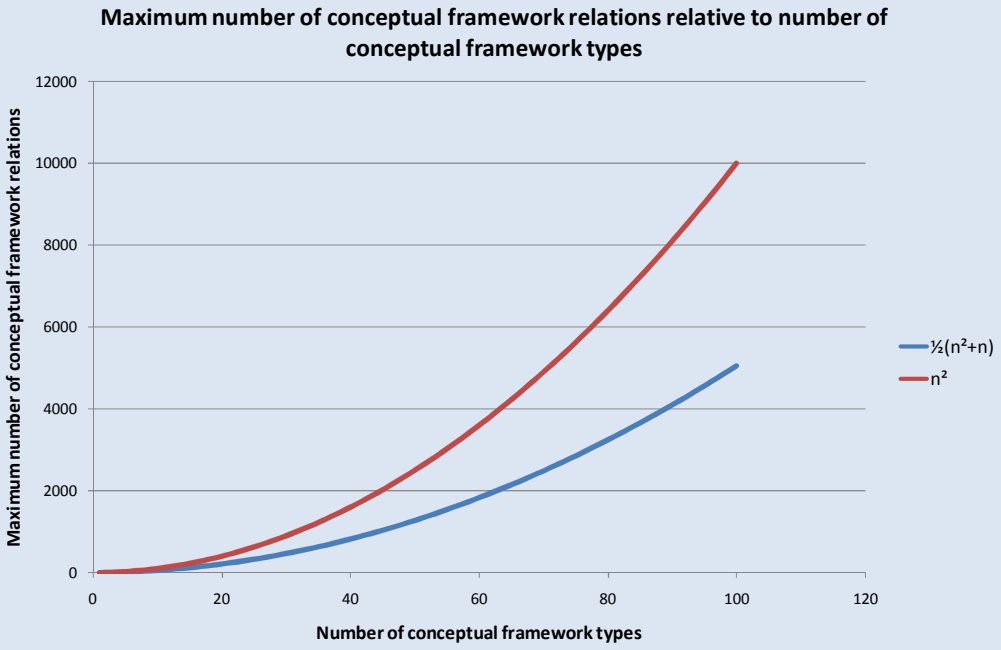


Figure 6.4: Maximum number of conceptual framework relations relative to the number of conceptual framework relations

The number of relations utilised in the conceptual framework can however be reduced through evaluation and subsequent exclusion from the conceptual framework, as some type combinations will exist that fail to represent any useful information (e.g. the relation between the telephone number of an employee and the number of pages printed by one of the organisation's printers). Type-pairs that do not warrant a relation should however be documented, for possible re-evaluation at a later stage. An evaluation approach that proved to be very effective (although unfortunately also very time-consuming) is the listing of all the identified types along the two axes of a matrix, by which every type combination can be evaluated (see Table 6.2).

Table 6.2: Table for evaluating conceptual framework type combinations

	Type_1	Type_2	Type_3	…	Type_n
Type_1	Relation_11	Relation_12	Relation_13	…	Relation_1n
Type_2		Relation_22	Relation_23	…	Relation_2n
Type_3			Relation_33	…	Relation_3n
⋮	⋮	⋮	⋮		⋮
Type_n				…	Relation_nn

Should it however be possible to automate the evaluation process, it will be possible to rapidly reduce the amount of time required to perform the type combination evaluation. The defining of relations between the types of a conceptual framework is a very labour-intensive task, but of the utmost importance to the conceptual framework construction process, as it is essential in providing the users of the conceptual framework with a view on organisational information that is both useful and familiar to them, and with which they can interact with confidence. When one then also takes a more general view on database tables and, in this case, the manner in which information regarding the relations that exist within the database is transferred to the conceptual framework, it can be seen that relations are typically identified by placing specific focus on the horizontal relations that already exist in the tables (see Figure 6.5).

	Field_1	Field_2	Field_3	…	Field_m
Record_1	Value_11	Value_12	Value_13	…	Value_1m
Record_2	Value_21	Value_22	Value_23	…	Value_2m
Record_3	Value_31	Value_32	Value_33	…	Value_3m
⋮	⋮	⋮	⋮		⋮
Record_n	Value_n1	Value_n2	Value_n3	…	Value_nm

Figure 6.5: Identification of information source relations through the horizontal investigation of a database table

Refine type-relation structure

The definition of types and relations from organisational information sources will probably have yielded a number of separate conceptual frameworks, which each represents a specific information source. These conceptual frameworks (and the information contained in them) remain separate from each other, as the overlapping of information contained within organisational information sources has not yet been considered. In order to create a consistent and "whole" conceptual framework which offers users a unified view of the information domain in question, overlapping types must be combined as they essentially represent the same entities within the organisational information domain. To illustrate, consider the following example: After the definition of types within the organisational information sources and relations between these types has been concluded, the organisational conceptual framework presents a fragmented view of the organisational information domain, consisting of three "sub"-frameworks which describe information from three information sources, namely an employee database, a product database and a department database (see Figure 6.6).

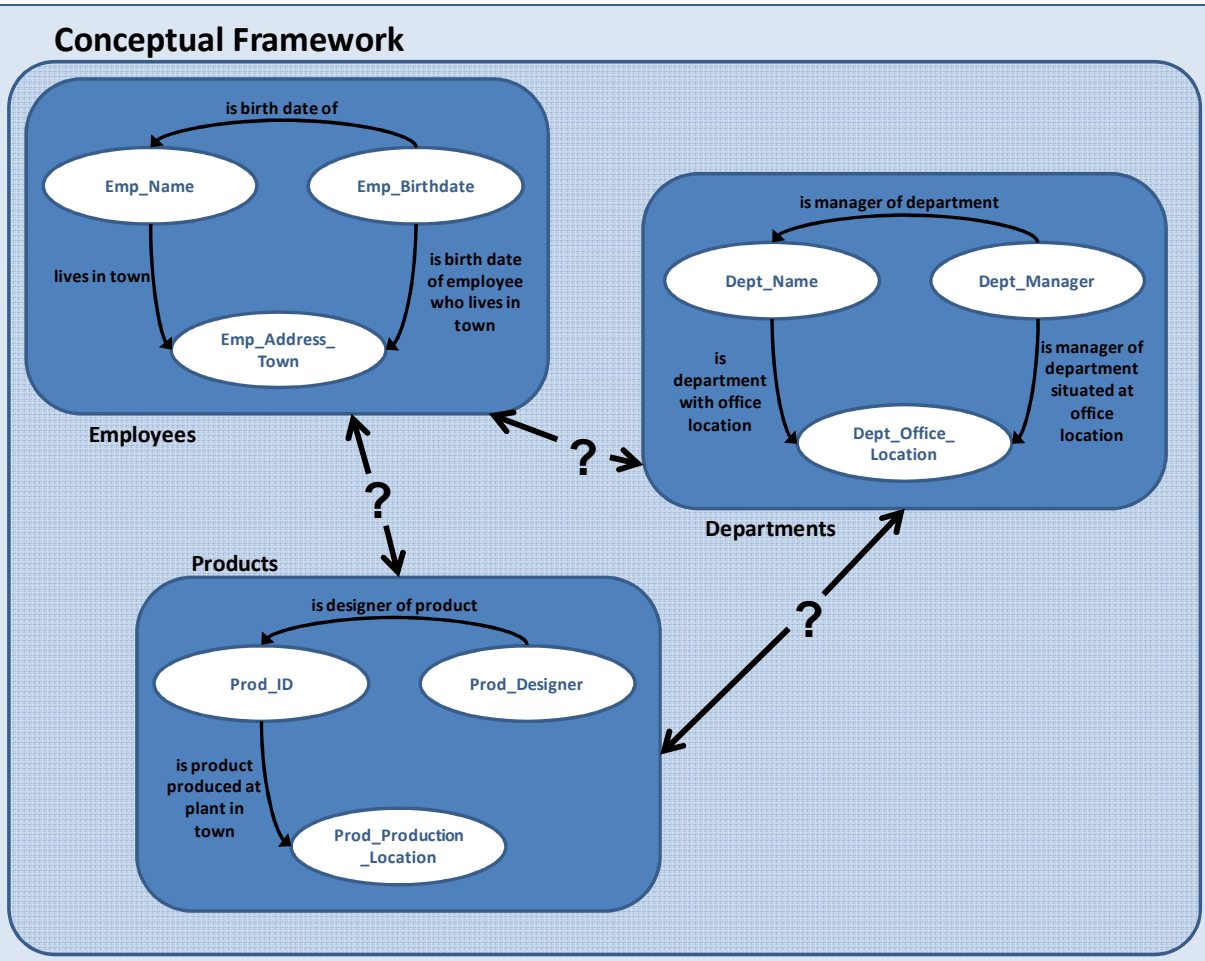


Figure 6.6: Example of a fragmented conceptual framework

It is obvious to see that **"Dept_Manager"** and **"Product_Designer"** are also employees of the organisation, which means that **"Emp_Name"**, **"Dept_Manager"** and **"Product_Designer"** can be combined to form a single type named **"Employee",** or even just **"Person"**. The same can be said of **"Emp_Address_Town"** and **"Prod_Production_Location"**, which may be combined to form a new type named **"Town"**. Once this combining of types has taken place (and some of the types have been given more user-friendly labels), the refined conceptual framework will look as shown by Figure 6.7.

The conceptual framework now provides an integrated view of the information contained in the three information sources discussed. It is, however, important to retain the information on the manner in which the combination of types took place, as this information will be crucial when the conceptual framework is populated in the Development phase of the conceptual framework implementation process (see C – Development phase).

(Note: It may happen that the conceptual framework remains fractured (even after refinement of the type-relation structure has been performed) due to the fact that certain information which is crucial to the conceptual framework implementation do not exist in an explicit manner in any of the organisation's

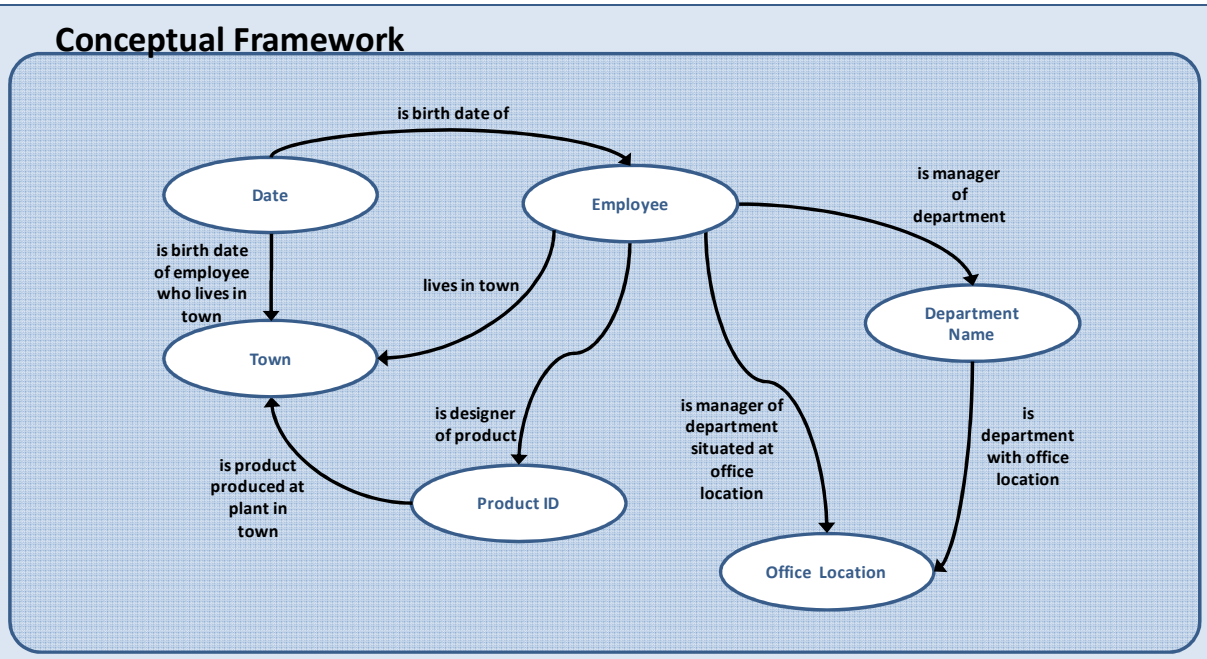


Figure 6.7: Example of a refined conceptual framework

information sources, but exists rather in an implicit form as a shared understanding between a number of the organisation's employees. The capturing of this information in an electronic form will be required in order for the conceptual framework to provide a complete and consistent view of the organisation's information. Once this information has been captured, its types and relations can be identified and also added to the type-relation structure which is developed in this phase of the implementation process.)

Verify type-relation structure

Once the refining of the conceptual framework has been concluded, the type-relation structure that was developed through the Analysis phase should be verified to ensure the correctness of the structure. The verification should take place on the following two levels:

- Information level: Discussions with the database and other information administrators of the information domain in question are required, in order to confirm the following aspects of the conceptual framework type-relation structure:
    - The correctness of the defined conceptual framework types, i.e. do the defined types correctly represent the information that constitutes the specified information domain?
    - The correctness of the defined conceptual framework relations, i.e. do the relations defined between the types correctly represent the manner in which information can be retrieved from the information domain?
    - The correctness of the type refining, i.e. were the information source types grouped together in the correct manner?

If any problems exist with any of these aspects, the relevant aspect should be readdressed, followed by additional verification until all persons involved are satisfied with the result.

- Logical level: Once the database and other information administrators are satisfied with the integrity of the type-relation structure, it should be presented to the stakeholders of the conceptual framework implementation (refer back to the Planning phase, see section A – Planning phase) for confirmation that it satisfies their information requirements (at least conceptually). Again, the three aspects discussed in the previous points should be evaluated (namely the correctness of types, relations and refining), but in this case the focus should be shifted to the perspective of a user of the information and the way in which the information domain under consideration is understood by the stakeholders. Also, because the total extent of the implementation is now known, complete cooperation is required from information infrastructure stakeholders to ensure that the implementation is accomplished as intended.

Once the stakeholders are satisfied that the type-relation structure mirrors their understanding of the types and the relations between them that constitute the specified information domain, the Analysis phase can be considered concluded and the Development phase can commence. Should the type-relation structure fail to satisfy the stakeholders' needs, it could be an indication that the stakeholders' initial information needs were incorrectly identified or that the definition of the type-relation structure was done in an incorrect manner, which would require that the Analysis phase of the conceptual framework implementation process be repeated, or even that the Planning phase of the conceptual framework implementation process be re-conducted.

Outputs:

- Similar to Activity B1, the output of this activity is information to be included in the **"Conceptual Framework Implementation – Analysis phase"** document (indicated as the output for phase B – Analysis phase). The information required is:
    - The conceptual framework types that were identified from the information sources, together with the location of the information represented by these types (preferably down to table-level in the case of a database) and their frequency of change;
    - The relations that were identified between the conceptual framework types; and
    - The final type-relation structure of the conceptual framework after refining, also indicating the manner in which types were grouped together to create an integrated view on the organisational information.

C – Development phase

Designation:	C	*Title:*	**Development Phase**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

During the Development phase of the conceptual framework implementation process the actual conceptual framework is created within the organisation, and populated with information extracted from its information sources.

Rationale:

Specific activities are required in order to successfully transform organisational information into conceptual framework format.

Prerequisites:

- Read-access to the relevant organisational information sources, as provided by the information infrastructure stakeholders (Activity B2)
- The organisational conceptual framework specification (refined type-relation structure) (Activity B2)

Execution:

- **Activity C1 – Fulfil conceptual framework hardware requirements**
- **Activity C2 – Create and populate conceptual framework**
- **Activity C3 – Verify information represented by conceptual framework**

Outputs:

- The following information that has been generated through the activities that constitute this phase, should be captured in a **"Conceptual Framework Implementation – Development phase"** document, and will function as a log of the conceptual framework generation, capturing information on the interaction with the organisational information infrastructure in order to generate the organisational conceptual framework:
    - General information on the conceptual framework software tool utilised in the implementation, especially noting its manner of information storage (see Activity C1);
    - The ETL processes (such as SQL queries, etc., associated with their specific information sources) that were developed and utilised in the creation of the conceptual framework (see Activity C2);
    - A declaration by the implementation stakeholders on the implementation, giving sign-off on the conceptual framework that was developed and releasing it for organisation-wide user implementation (see Activity C3).
- A fully functional and stake holder-validated conceptual framework, residing in the organisational conceptual framework software tool (and already part of the organisational information infrastructure).

- ***Activities***

Designation:	**C1**	*Title:*	**Fulfil conceptual framework hardware requirements**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

This activity focuses on the acquisition of the hardware on which the organisational conceptual framework (and the software which facilitates user interaction with the conceptual framework) will reside.

Rationale:

The conceptual framework and the software tool which is to facilitate user interaction with the organisational conceptual framework require specific hardware capabilities, which have a direct impact on the performance experienced by users when interacting with the implemented conceptual framework.

Prerequisites:

- Information on hardware capabilities required by the conceptual framework software

Execution:

The organisational conceptual framework will be created and made accessible to users through a conceptual framework software tool, designed to facilitate the interaction between users and conceptual frameworks. The hardware required by this software tool should be acquired to insure that the software tool can effectively provide the functionality required by users interacting with the implemented conceptual framework. With assistance from the information infrastructure stakeholders (see section A – Planning phase), the required equipment should be procured, installed and configured according to the specifications provided by the developers of the conceptual framework software, after which the software itself should be installed and configured in the prescribed manner.

Outputs:

- Information to be included in the **"Conceptual Framework Implementation – Development phase"** document (indicated as an output for phase C – Development phase). The information required is:
    - General information on the conceptual framework software tool utilised in the implementation, especially noting its manner of information storage.
- A fully functional and tested installation of the conceptual framework software tool residing on hardware that is part of the organisation's information infrastructure and is dedicated for conceptual framework purposes.

Designation:	**C2**	*Title:*	**Create and populate conceptual framework**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

Through this activity, organisational information is extracted from organisational information sources and, based on certain guidelines and requirements associated with the conceptual framework software tool being used (see Activity C1), and transformed into the organisational conceptual framework as described by the conceptual framework specification developed during the Analysis phase (see B – Analysis phase).

Rationale:

This activity creates the conceptual framework with which organisational users will interact in order to improve the manner in which information is located, viewed and utilised in the organisation.

Prerequisites:

- Read-access to the relevant organisational information sources, as provided by the information infrastructure stakeholders (Activity B2)
- The organisational conceptual framework specification (refined type-relation structure) (Activity B2)
- A fully functional and tested installation of the conceptual framework software tool residing on hardware that is part of the organisation's information infrastructure and is dedicated for conceptual framework purposes (Activity C1)

Execution:

Extraction, transformation and loading of data into conceptual framework

Once the necessary network and information source access have been acquired, the information specified by the type-relation structure developed in the Analysis phase (see section B – Analysis phase) must be extracted from the relevant information sources and transformed by the conceptual framework software tool into the

organisational conceptual framework with which users can interact. This is typically an ETL (Extraction-Transformation-Loading) process (Skoutas & Simitsis 2006) through which information is extracted from an information source, transformed into the storage format used by the conceptual framework software tool (should it be necessary), and then loaded into the storage of the conceptual framework software tool from where the tool will make the information available for user interaction.

The extraction of information from an information source and the subsequent transformation of said information is however not a trivial process, and may be performed in myriad ways, depending mostly on the manner in which information is stored within the specific information source (e.g. as relational database tables, spreadsheets, flat text files, etc), and on the structure in which the conceptual framework software tool requires the information to be stored in order to effectively interact with it (similar to Skoutas & Simitsis 2006). The success of this activity within the Development phase (see section C – Development phase) therefore depends on an exact understanding of the differences between the manner in which information is stored within the various information sources of the implementation domain, and the manner in which the conceptual framework software tool requires the information to be stored. For example, if an information source and the conceptual framework software tool both make use of relational databases for the storage of information, simple SQL (Structured Query Language) querying and other operations should be sufficiently able to extract, transform and store data in the format required by the conceptual framework software tool. Should the information source be a spreadsheet, the extraction, transformation and loading of the data should also be quite straightforward. In the event that the conceptual framework software tool utilises text-base data storage like XML (Extensible Markup Language) however, a SQL query output from a database will obviously have to undergo more complicated transformations in order to conform to the particular storage method.

Based on an understanding of the different manners in which information may be extracted from the organisational information sources, the conceptual framework should be constructed in an incremental fashion through the executing of the various associated ELT processes, which are customised to extract data from each of the conceptual framework implementation domain's information sources and provide it to the conceptual framework software tool in the required manner. Utilising this information, the conceptual framework software tool can now present the content of the information domain to the user and, through its various functionalities, provide the user with the means for navigation of the information, interaction with the information and information discovery.

Verification of conceptual framework rationality

To conclude the activity of creating and populating the organisational conceptual framework, verification of the rationality of the conceptual framework is required. This is purely a measure of control to ensure that the rationality contained in the conceptual framework specification was correctly transferred to the actual conceptual framework, and that the information that is presented by the conceptual framework makes sense. This is a simple exercise and should not require more effort than the evaluation of a few facts associated with

each relation between the various conceptual framework types. If nonsensical relations are found, the ETL process that generated the information should be investigated and corrected if found to be faulty. If the specific ETL process executed correctly, and the specification of the conceptual framework were actually found to be at fault (which means that the specific relation should therefore not exist within the conceptual framework), it should be corrected, and the relevant ETL processes be updated to reflect the change. Following any corrections, the conceptual framework should obviously be regenerated.

Outputs:

- Information to be included in the **"Conceptual Framework Implementation – Development phase"** document (indicated as an output for phase C – Development phase). The information required is:
    - The ETL processes (such as SQL queries, etc., associated with their specific information sources) which were developed and utilised in the creation of the conceptual framework.
- A fully functional and rationality-verified conceptual framework, residing in the organisational conceptual framework software tool (and already part of the organisational information infrastructure).

Designation:	**C3**	*Title:*	**Verify information represented by conceptual framework**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

At this stage, an operational conceptual framework has been developed with which users can interact for various information locating-type activities. To complete the Development phase of the conceptual framework implementation process, the implementation stakeholders should verify the correctness of the information represented by the conceptual framework, and determine whether the information and views represented by the conceptual framework conform to the requirements stipulated by them in the Planning phase.

Rationale:

This activity is aimed at ensuring that the conceptual framework that was created conforms to the requirements of the implementation stakeholders, as well as to their understanding of the organisation.

Prerequisites:

- A fully functional and rationality-verified conceptual framework, residing in the organisational conceptual framework software tool (and already part of the organisational information infrastructure) (Activity C2)

Execution:

Each of the stakeholders will be required to spend some time evaluating the conceptual framework and the information it contains. The duration of this evaluation should be determined by the stakeholders as a group, and will typically be proportionate to the magnitude and the complexity of the implementation domain. Should any problems be unearthed during this evaluation, the sources of these problems (which most probably would have occurred in one or both of the Analysis or Development phases) should be determined, and addressed. Once satisfied, the stakeholders should give sign-off on the conceptual framework, which will

act as validation for the conceptual framework and release it for organisation-wide user implementation.

Outputs:

- Information to be included in the **"Conceptual Framework Implementation – Development phase"** document (indicated as an output for phase C – Development phase). The information required is:
    - A declaration by the implementation stakeholders on the implementation, giving sign-off on the conceptual framework that was developed and releasing it for organisation-wide user implementation.
- A fully functional and stake holder-validated conceptual framework, residing in the organisational conceptual framework software tool (and already part of the organisational information infrastructure).

D – Implementation phase

Designation:	**D**	*Title:*	**Implementation Phase**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

Once a conceptual framework has been created and validated by its stakeholders, it should be configured and made available for organisation-wide use.

Rationale:

Up until this point in the implementation, access to the conceptual framework has been limited to those persons involved in the design and development of the conceptual framework. As the conceptual framework can benefit all knowledge workers in an organisation, it should now be made available to more persons within the organisation. This necessitates the reconfiguring of some aspects of the organisational information infrastructure in order to be able to support a changed flow of information because of an increase in user interaction with the conceptual framework.

Prerequisites:

- A fully functional and stake holder-validated conceptual framework, residing in the organisational conceptual framework software tool (and already part of the organisational information infrastructure) (Activity C3)

Execution:

- **Activity D1 – Information infrastructure setup**
- **Activity D2 – User setup**

Outputs:

- In order to preserve the context in which the Implementation phase of the conceptual framework implementation process took place, the following information should be captured in a **"Conceptual Framework Implementation – Implementation phase"** document:
    - The exact arrangements that were made with regards to the information infrastructure for the optimal functioning of the conceptual framework within the organisation, e.g. bandwidth allocation, the updating frequency of information, security, etc. (Activity D1)
    - The initial user base, with associated user account details and security privileges (Activity D2); and
    - The documentation used in the training of the conceptual framework users (Activity D2).
- A final **"Conceptual Framework Implementation"** document should be compiled by combining the four output documents listed below that were created through the course of the conceptual framework implementation process, namely:
    - The "Conceptual Framework Implementation – Planning phase" document;
    - The "Conceptual Framework Implementation – Analysis phase" document;
    - The "Conceptual Framework Implementation – Development phase" document; and
    - The "Conceptual Framework Implementation – Implementation phase" document.

  This document can function as the sign-off document for the conceptual framework implementation, as it contains all the main actions, decisions and designs that contributed to the implementation of the organisational conceptual framework.
- A fully functional operational conceptual framework available to knowledge workers for improved locating, viewing and utilising of organisational information, residing on organisational information infrastructure configured for optimal conceptual framework user interaction.
- Trained users for the organisational conceptual framework

- ***Activities***

Designation:	**D1**	*Title:*	**Information infrastructure setup**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

This activity entails the reconfiguring of various aspects of the organisational infrastructure because of changes in the flow of information within the organisation, resulting from the use of the conceptual framework by the organisation's employees.

Rationale:

This activity is performed to ensure that the benefit experienced through the use of the conceptual framework is maximised.

Prerequisites:

- A fully functional and stake holder-validated conceptual framework, residing in the organisational conceptual framework software tool (and already part of the organisational information infrastructure) (Activity C3).

Execution:

Relevant information infrastructure administrators and other decision makers should be approached and a strategy developed through which (at least) the following aspects are addressed:

- Bandwidth allocation: This aspect has a great impact on the overall efficiency of the organisational infrastructure and on the productivity of conceptual framework users. A large allocation of bandwidth to the conceptual framework implementation and its users will ensure fast response times for the software tool being used, but could have a detrimental effect on the other systems and the users of

these systems in the organisation. On the other hand, allocating too little bandwidth to the conceptual framework could obviously result in a negative experience by users of the software tool. It is therefore important to find a balance where as much bandwidth as can be spared be allocated to the use of the conceptual framework, whilst otherwise managing the expectations that users of the software tool will have.

- Updating of information: The organisational conceptual framework will be most useful if it could present a real-time view on the information domain it describes. Implementation stakeholders, as well as other relevant information infrastructure stakeholders and decision makers in the organisation should therefore be tasked to make a decision on the extent of relevancy that is expected of the conceptual framework. With this decision in mind, the various frequencies of change experienced within the information domain (as documented in the "Conceptual Framework Implementation – Analysis phase" document) must be inspected, leading to the creation of a conceptual framework update schedule, which specifies the frequency at which the various ETL processes used in the Implementation phase should be re-run in order to keep the conceptual framework updated (in any case to the extent that the stakeholders require). After the schedule has been compiled, the ETL processes should be configured to execute at the specified time intervals, based on the recommendation set out in the conceptual framework update schedule.
- Security: Within an organisation there are normally certain restrictions in place with regards to information access which is aimed at preventing the potential misuse of organisational information. These restrictions are usually structured around teams or departments and the roles of employees within that teams or departments. For the conceptual framework to be accepted as a tool for organisational use, it is crucial that security considerations also be taken into account and that the information residing in the conceptual framework has the same level of protection as the information residing in the organisational information sources. It would be ideal if the conceptual framework software tool has the capability to integrate with the security mechanisms already resident in the organisation's information infrastructure, as the existing permissions and restrictions associated with the organisational information can then be utilised by the software tool in managing access to specific pieces of information within the conceptual framework (i.e. acting as a layer on top of the conceptual framework which only allows user access based on existing organisational security settings). If the conceptual framework software tool does not offer that functionality, security information that forms part of the implementation domain's information sources should be identified and incorporated into the conceptual framework during the Development phase. The software tool should then be configured to interrogate the conceptual framework itself for security information, which it will then be utilised in managing access to specific information in the organisational conceptual framework.

Bandwidth allocation, updating of information and security are some of the more typical system aspects that will have to be addressed during the implementation of a conceptual framework in an organisation. This is

however by no means an exhaustive list and attention to other system aspects may be necessary, depending on the manner and extent of the conceptual framework implementation and its role in the organisation.

Outputs:

- Information to be included in the **"Conceptual Framework Implementation – Implementation phase"** document (indicated as an output for phase D – Implementation phase). The information required is:
    - The exact arrangements that were made with regards to the information infrastructure for the optimal functioning of the conceptual framework within the organisation, e.g. bandwidth allocation, the updating frequency of information, security, etc.
- A fully functional operational conceptual framework available to knowledge workers for improved locating, viewing and utilising of organisational information, residing on organisational information infrastructure configured for optimal conceptual framework user interaction.

Designation:	**D2**	*Title:*	**User setup**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description:

This activity entails the creation of user accounts through which organisation employees can access the information contained in the conceptual framework, as well as training of these employees in its use.

Rationale:

Controlled access to the information contained in the organisational conceptual framework is required in order to prevent the unsolicited viewing of sensitive information. User training is required as the approach of viewing information as a collection of entities and the relations between them are in its infancy, and not well known to most organisational employees.

Prerequisites:

- A fully functional and stake holder-validated conceptual framework, residing in the organisational conceptual framework software tool (and already part of the organisational information infrastructure) (Activity C3)
- Organisational information infrastructure configured for optimal conceptual framework user interaction (Activity D1)

Execution:

When the relevant information infrastructure aspects of the Implementation phase have been concluded and the conceptual framework is sufficiently configured for use in the organisation (see Activity D1), the user component should be introduced by performing the following activities:

- Creation of user accounts and allocation of security privileges: User accounts should be created in the conceptual framework software tool through which the persons in the organisation that have been

identified as users of the conceptual framework could gain access and utilise the information it contains. Each user account will obviously fit into the conceptual framework's security structures (configured in the information infrastructure setup which was discussed in the previous section), ensuring its associated user has access to the information that pertains to him/her and is prevented from interacting with restricted information. In the spirit of allowing users to perform information discovery through interaction with the conceptual framework (which may lead to the generation of new ideas, the improvement of business processes, etc.), it is however advised that those persons who manage user access to information within the organisational conceptual framework refrain from enforcing too many access restrictions on a user and only prevent user interaction with specific information if absolutely crucial.

- Training: Once the conceptual framework user accounts have been set up, training should be provided to these users, imparting both the theoretical and practical knowledge that they will require to interact with the organisational conceptual framework in the most beneficial manner. The theoretical component of the conceptual framework user training should explain the conceptual framework approach in general, the manner in which it utilises entities and relations to create a navigable structure for information gathering and discovery, as well as the ways in which the organisation (and the users themselves) will benefit from interaction with the conceptual framework. The practical component should focus on familiarising the users with the various features and functionalities offered by the specific conceptual framework software tool employed by the organisation (hands-on, if possible), and how it allows them to interact with the conceptual framework. A final understanding which should be impressed on the users of the conceptual framework is that it does not only serve as a mechanism for finding information; it enables them to integrate their own knowledge into the greater whole of the organisational knowledge and by interacting and contributing to the conceptual framework on a regular basis, they themselves (as well as the organisation as a whole) will benefit.
- Launch: Finally, to conclude the Implementation phase and to initiate organisation-wide use of the conceptual framework, relevant communiqués should be distributed within the organisation, announcing the commencement of use of the conceptual framework, its purpose and its intended value to the organisation.

Outputs:

- Information to be included in the **"Conceptual Framework Implementation – Implementation phase"** document (indicated as an output for phase D – Implementation phase). The information required is:
  - The initial user base, with associated user account details and security privileges; and
  - The documentation used in the training of the conceptual framework users.
- Trained users for the organisational conceptual framework.

E – Maintenance phase

Designation:	**E / E1**	*Title:*	**Maintenance Phase / Grow Conceptual Framework**

Navigation:

A - Planning	B - Analysis	C - Development	D - Implementation	E - Maintenance
A1 – Identify stakeholders	B1 – Examination of information sources pertaining to stakeholder-specified information requirements	C1 – Fulfil conceptual framework hardware requirements	D1 – Information infrastructure setup	E1 – Grow conceptual framework
A2 – Initial information need identification	B2 – Development of conceptual framework specification	C2 – Create and populate conceptual framework	D2 – User setup	
		C3 – Verify information represented by conceptual framework		

Description and Rationale:

A conceptual framework implementation typically focuses only on a subset of organisational information, delivering a conceptual framework that only describes the information found within a specific information domain within the organisation. In order to represent the organisation's information in its entirety, more iterations of the conceptual framework implementation process should therefore be performed. Such iterations can vary in magnitude and scope, being (for example) the correction of user-identified errors that occurred during the creation of the conceptual framework in the Development phase of an implementation, the adding of entities or relations that were excluded or overlooked in the design of the conceptual framework during the Analysis phase, or even another full implementation with (possibly) new stakeholders and new information requirements (which means adding the information of an entirely new domain to the existing conceptual framework). It is however imperative that, irrespective of the extent of such iteration, any changes to the organisational conceptual framework be captured in the relevant documentation as specified in this methodology, in order to retain the knowledge that was acquired on the process that was followed, as this could be beneficial in future conceptual framework maintenance efforts.

Finally: It is important that the conceptual framework is continuously grown through the execution and re-execution of the conceptual framework implementation process, as this will eventually yield a conceptual framework which can provide users with access to and a view on all the information that an organisation possesses, enabling them to interact with and utilise organisational information in an efficient manner.

Prerequisites:
• An existing conceptual framework
Execution:
• See *Description and Rationale* above
Outputs:
• Improved/expanded organisational conceptual framework

6.3 Conclusion – Conceptual framework implementation methodology

This chapter has introduced a methodology for implementing a conceptual framework in an organisation with the objective of improving the utilisation of organisational information. Through the execution of the five phases that constitute the methodology – Planning, Analysis, Development, Implementation, and Maintenance – a conceptual framework based approach for the improved viewing and utilisation of information can now be implemented in organisations. Additional to a fully functional conceptual framework, execution of the conceptual framework implementation methodology further yields documentation describing every phase of the methodology, users trained in interacting with a conceptual framework, and an organisation information infrastructure configured for optimal use of a conceptual framework by its employees.

The validation of the conceptual framework implementation methodology is discussed in the next chapter (Chapter 7).

Chapter 7 – Methodology Validation

7.1 Introduction

The previous chapter (Chapter 6) described a methodology that was developed for guiding the implementation of a conceptual framework in an organisation, which in turn is aimed at improving the viewing and utilisation of organisational information. Validation[24] of this methodology was done in the following manner:

- First, a very brief, high level comparison was made between the implementation methodology developed through this research project, and other existing methodologies that support the implementation of similar information modelling approaches (like ontologies) in organisations. The purpose of this "theoretical" validation was to ensure that the aspects on which the conceptual framework implementation methodology focuses are relevant, and that the methodology as a whole is conceptually sound.
- Secondly, two case studies were conducted in partnership with Indutech (see section 5.1), in which the methodology was used for implementing a conceptual framework in two different organisations. In each of these case studies the purpose of the implementation was first and foremost to validate the methodology through practical experience, and secondly (which is also less important) to test a conceptual framework software tool developed by Indutech, named Organon (see Appendix A).

This chapter will discuss the results of these two methods of validation (see also Figure 7.1). The section focusing on the theoretical validation of the methodology will discuss the purpose and manner of application of the other methodologies, together with the manner in which these methodologies correspond with and differ from the conceptual framework implementation methodology. The section on the practical validation of the methodology will discuss the case studies that were performed, focusing especially on the information that was transformed into the conceptual framework, the process that was followed to create the conceptual framework, the end result of the implementation, and implementation stakeholder impressions of the implemented conceptual framework and the implementation methodology.

[24] *Validation: to make valid; substantiate; confirm (Dictionary.com 2008g).*

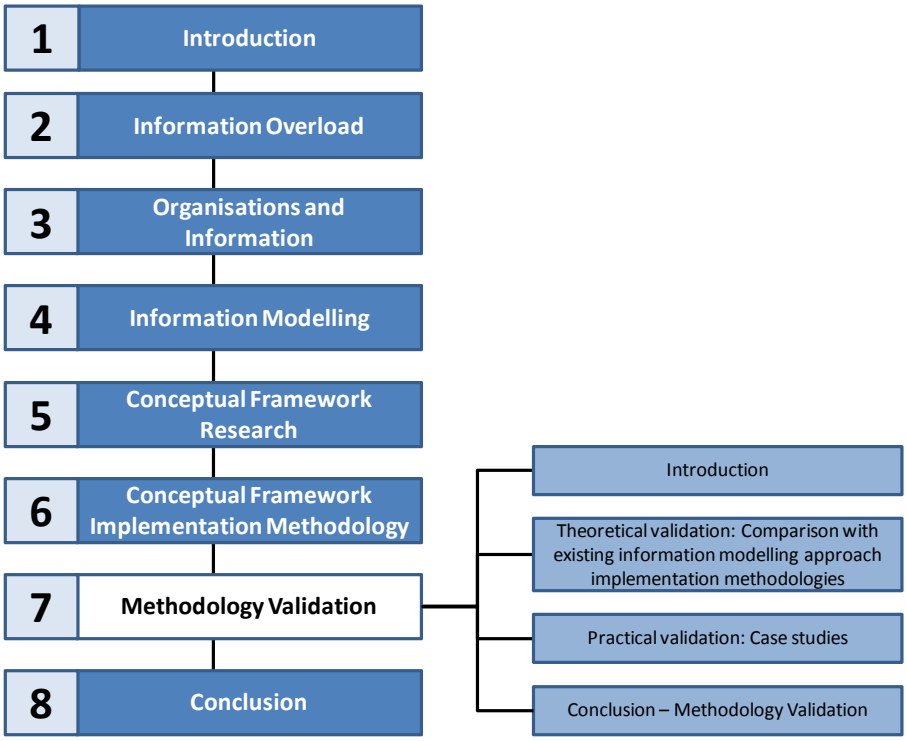


Figure 7.1: Image for navigating thesis and chapter 7

7.2 *Theoretical validation: Comparison with existing information modelling approach implementation methodologies*

Subsequent to the development of the methodology introduced in the previous chapter (Chapter 6), a brief investigation was performed through which a number of existing information modelling approach implementation methodologies were identified and compared to the conceptual framework implementation methodology for validation purposes. The focus and level of detail of the methodologies used for the comparison varies, but the collection of methodologies was deemed sufficient for validation of the conceptual framework implementation methodology. For the purpose of this discussion the methodologies investigated were divided into those that focus on the same level of detail as the conceptual framework implementation methodology, and those that do not.

7.2.1 Methodologies focusing on the same level of detail as the conceptual framework implementation methodology

This section discusses the information modelling approach implementation methodologies which were directly compared with the conceptual framework implementation methodology in order to ascertain its validity from a theoretical point of view.

- ***The Semantic Information Management (SIM) methodology***

The Semantic Information Management (SIM) methodology (Schreiber 2003) was developed in the earlier part of this decade by a company named Unicorn Solutions Inc. (which has since been acquired by IBM). Its function is to facilitate the establishing of the SIM approach in organisations (which is aimed at addressing the organisational problem of extracting relevant information from a knowledge base that contains heterogeneous types of information). This solution is based on the following core elements:

- Semantics, through which a common understanding of the different kinds of data that exist within the organisation is created;
- An Information Model, (typically an ontology, see also section 4.2.4) which is constructed to reflect the agreed-upon business view, business vocabulary and business rules of the organisation, and aimed at providing a common basis for understanding organisational data; and
- Metadata, which captures the formal meaning of organisational data in agreed-upon terms.

The SIM methodology facilitates the establishing of these core elements through execution of the following steps (or phases) (Schreiber 2003) (see also Figure 7.2):

- Gather Requirements - Establish the project scope, survey the relevant data sources and capture the organisation's information requirements.
- Catalogue Metadata - Catalogue data assets and collect metadata relevant to the organisation and its use of data.
- Construct Information Model - Capture the desired business world-view, a comprehensive vocabulary and business rules.

- Rationalize (Data Semantics) - Capture the meaning of data by mapping to the Information Model.
- Publish/Deploy - Share the Information Model, metadata and semantics with relevant stakeholders; customise it to their specialized needs.
- Utilize - Create processes to ensure utilisation of architecture in achieving data management, data integration and data quality.

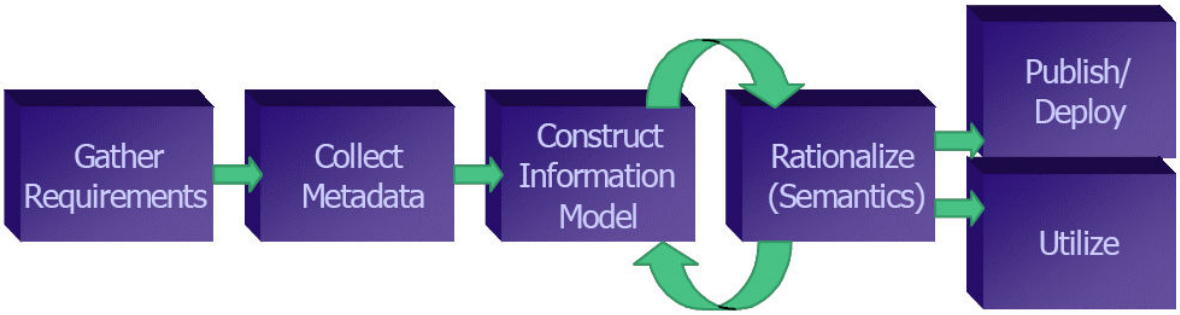


Figure 7.2: The Semantic Information Management methodology (Source: Schreiber 2003)

• *The Knowledge Lens methodology*

The Knowledge Lens methodology (Edgington, et al. 2004) was employed in a domain ontology development project at the Intel Corporation in the United States, in which a number of US academic institutions were also involved. The project arose from a need to better utilise past information in performing failure analysis and failure identification (FA/FI) on integrated chips at the large semiconductor manufacturing firm. A hybrid (top-down and bottom-up) modelling approach was used to develop a specific ontology as an enabler to integrating knowledge management practices and processes (Edgington, et al. 2004).

The aim of the project was the establishment of a "Knowledge Lens", which is an ontological focus on organisational knowledge units. 'A knowledge unit is a coarse set of information elements bound together by structure, assumptions, justifications, and process. These characteristics provide a perspective that typically does not exist with mere information or data elements' (Edgington, et al. 2004). The purpose of the Knowledge Lens was to assist the organisation in appropriately calibrating the coordination of its activities for countering innovation aggressiveness from competition, and provide a shared vocabulary, access and reuse, speed and relevancy, and definition of any knowledge unit.

The steps that were followed (and through which the ontology was implemented) are as follows (Edgington, et al. 2004) (see also Table 7.1):

Table 7.1: The iterative ontology development process utilised in the development of Knowledge Lens (Source: Edgington et al. 2004)

Processes	Tasks
Design	Formulate problem statement Define scope Audience/users, goals, resources, schedule, level of formality Develop success and acceptance criteria Investigate tasks and business area goals Analyze use cases
Develop	Extract control vocabulary Develop conceptual model(s) Incorporate vocabulary acquisition tools as appropriate
Integrate	Develop formal interview structure Review models with participants Initiate formal interviews and incorporate into models Expand models with expanded attributes and axioms Develop computational representation as appropriate RDF, XML, formal logic and/or other programmatic representation
Validate and feedback	
Iterate process	Analyze and refine after each process

- Design - The process of design includes formulating the problem statement, defining the project scope, developing success and acceptance criteria, investigating tasks and business area goals, and analysing use cases.
- Develop - Ontological development identifies and extracts the control vocabulary and, by incorporating relevant relationships, develops a number of conceptual models.
- Integrate – Ontology integration starts with the integration of each individual's "knowledge lens" or perspective, and expanding this to the organisation or inter-organisation in conformance with the project's scope. From this integrated model a computational representation is created. Additional extensions to the model in the form of attribute expansion and rules are also added to enrich the contents of the ontology. Any revisions are repeatedly reviewed with participants
- Validate and feedback – Validation of the ontology takes place in three phases: First, the sub-models are validated qualitatively among the participating users and among additional users as adequately representative of the project's focus. Secondly, the ontology is validated by applying the models within constructed queries and manually examined to see how the ontology subsumed the query. Finally, the project champion has face-to-face interaction with the various groups and locations for which the ontology and ontological process were deemed to have value, through which questions from the individual groups can be addressed, support for the approach solicited, and the management value of the approach be high-lighted.
- Iterate process - Iteration should be occurring within most, if not all, of the process steps as interviews and investigations reveal opportunities for improvement and clarity. Analysis of

the validation and feedback step should be expected to produce subsequent opportunities for refinement.

• *The Cue-N-Anchor guided ontology construction strategy*

In a general discussion on ontologies and ontology construction, Kishore, Sharman and Ramesh (2004a) discuss a generic methodology for constructing ontologies, named the Cue-N-Anchor guided ontology construction strategy. The strategy resulted from a synthesis of common themes and directions identified from methodologies proposed by a number of researchers in the field of ontologies (e.g. Borst et al.; Fernández et al.; Holsapple and Joshi; Kishore et al.; Noy and McGuinnss; Uschold; Uschold and Gruninger; Uschold et al., all cited in Kishore, Sharman & Ramesh 2004a). The strategy (which is fundamentally a number of guidelines for ontology construction) is as follows:

- *Guideline 1:* Define the area and scope of the ontology – To gain a clear understanding of the purposes for which the ontology is being built (in order to define its area and scope), the concept of competency questions should be employed. Competency questions are essentially queries about the scenarios for which an ontology is designed and which the ontology should be able to answer. These questions can be used to guide the scope definition, such as what ontology is needed, what should be its level of detail, and whether it will serve its presumed purpose and similar inquiries.
- *Guideline 2*: Perform a baseline analysis – The baseline analysis consists of two tasks, namely brainstorming and the review of existing ontologies and relevant literature. The purpose of these activities is to produce all concepts and relationships relevant within the scoped ontology area, eliminating redundancies and ambiguities, and building a tentative ontology.
- *Guideline 3:* Anchor ontology well and use cues to guide its development – Employing cues and anchors, is a way of protecting the developers construction the ontology from information that may be brought about from interacting with vast amounts of information. The strategy suggested by Kishore, Sharman and Ramesh (2004a) is as follows:

    1. *Identify a set of ideas as anchors.* These ideas could be domain-specific, context-specific or even literature-specific. The proposed ontology should be adequately grounded in these anchors so that the development effort is both guided and protected from loss of direction.
    2. *Identify a set of ideas as cues.* Again these ideas could come from different sources. The cues are ancillaries that could be used to both enrich the ontology as well as guide the development.

- *Guideline 4:* Develop a glossary of terms and refine the competency questions – The glossary should enumerate the concepts, relationships, behaviours and even rudimentary structures (if possible), of the ontology. The competency questions furthermore assume more definitive shapes and are specified formally at this stage. The glossary and the refined and formalised competency questions constitute the baseline ontology document.

- *Guideline 5:* Structure the baseline glossary into a specifiable ontology – At this stage the baseline ontology typically consists of concepts and basic relations between them, like *is-a*, *has-a*, and *member-of*. Extended relations should now be defined by overlapping the concepts and relations found in the glossary (developed in Guideline 4). The ontology specification can also be made more formal through the incorporation of class structures, and their properties. Axioms and constraints should be derived by applying logical and evidential reasoning to the ontology specification. Finally, the emerging structure should be tested for soundness using its constraints.

- *Guideline 6:* Decide on integrating existing ontologies with the one being built and evaluate formal representation mechanisms – While building the ontology, existing ontologies may be found that could be integrated with the current one in order to achieve its purpose. Integration could occur on two levels: the knowledge capture level, and the knowledge level. At the knowledge capture level, well-understood knowledge capture mechanisms used in existing ontologies could be used to extend, refine, and even simplify the current ontology. At the knowledge level, representational concepts, tools and other artefacts utilised by the existing ontologies could be used to enrich and simplify the current ontology.

- *Guideline 7:* Develop the formal representation of the ontology – The formal specification of the ontology can now be constructed. It should contain the following:

    - The foundational conceptual model of the universe being modelled at appropriate levels of granularity; and

    - The full schema of the ontology describing the concept-relation structures, behaviour models, assumptions, axioms and constraints, and proofs of bounded completeness and soundness of representation.

• *The On-To-Knowledge methodology*

The On-To-Knowledge methodology (Lau & Sure 2002) is yet another methodology that focuses on ontology implementation and was developed at the University of Karlsruhe in Germany. According to Lau and Sure (2002), it was developed in response to a need from industry at a time when organisations were starting to adopt the idea of ontologies, but had lacked the knowledge on how

to implement this information modelling approach. The On-To-Knowledge methodology consists of the following steps (or phases) (see also Figure 7.3):

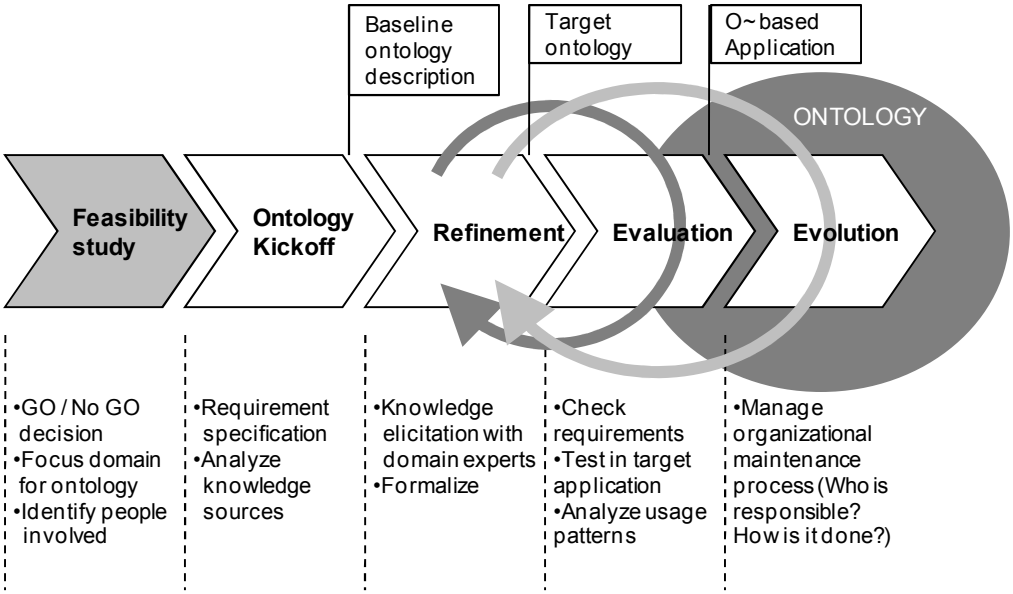


Figure 7.3: The steps of the On-To-Knowledge methodology (Source: Lau & Sure 2002)

- Feasibility study – A feasibility study is performed initially to first, identify problem/opportunity areas and potential solutions, and second, to put these in a larger perspective. By providing information on economical and technical feasibility, it serves as a decision support in order to select the most promising focus area, i.e. the domain for the ontology based system to be developed. Besides the implementation domain of the system it also helps to identify the people involved in setting up and using the system (i.e. the domain experts, users and supporters of the system).

- Ontology kickoff – The requirements of the ontology to be implemented should be captured in an Ontology Requirements Specification Document (ORSD), which describes what the ontology should support, and its area of application. It should guide an ontology engineer to decide about inclusion, exclusion and the hierarchical structure of concepts in the ontology. A "baseline ontology" is gathered through the analysis of available knowledge sources. This ontology typically contains only the most important concepts and relations that have been identified on an informal level.

- Refinement – The purpose of this phase is to produce a mature and application-oriented "target ontology" according to the specification given by the kick-off phase. This phase is divided into the following different sub-phases:

    - A knowledge elicitation process with domain experts based on the initial input from the kick-off phase. This serves as input for further expansion of the baseline ontology. Typically axioms are identified and modelled in this phase.

- A formalisation phase to transfer the ontology into the "target ontology" expressed in formal representation languages like DAML+OIL (see section 4.2.4). The representation language is chosen according to the specific requirements of the envisaged application.

This phase is closely linked with the evaluation phase, in that any gaps or misconceptions detected in the ontology during the evaluation phase act as input to another iteration of the refinement phase. It might therefore be necessary to perform several iterations of the refinement phase.

- Evaluation – Evaluation of the target ontology takes place in two steps: First, it is determined whether the ontology corresponds with the requirements set out in the ontology requirements specification document, and whether it sufficiently answers competency questions specified during the kick-off phase. Secondly, the ontology is tested in the target application environment, where the feedback of beta users may be useful for further refining the ontology. Patterns of ontology usage may also be used as a valuable input for configuring the ontology to more closely correspond with the needs of its users.
- Evolution – Like other software, ontologies also have to be maintained. This phase is aimed at ensuring that the maintenance of ontologies becomes an organisational process, and that clarity exists about who is responsible for maintenance and how it is performed. Changes to the ontology should furthermore be captured, and together with feedback from users, be utilised in additional cyclic refinement and evaluation phases.

- ***Pre-Built Information Space (PreBIS) implementation methodology***

Finally, the Pre-Built Information Space (PreBIS) implementation methodology (Böhm et al. 2005) was collaboratively developed by a number of German and Austrian research institutes to assist organisations in effectively utilising the information they possess. This is done through the structuring and filtering of organisational information and again the methodology is focused around the ontology information modelling approach. The PreBIS methodology is structured along two dimensions, namely the phases of the implementation, and action fields of the implementation. The phases of the implementation methodology are as follows:

- Demand analysis, which addresses the initiation of projects which are relevant to the organisation
- Mapping, which focuses on the manner in which the structure of the organisation is represented for analysis

- Analysis, which addresses the manner in which the organisational structure is analysed in order to ascertain the potential for optimisation
- Optimisation, which addresses the manner in which the structure of the organisation should be optimised
- Implementation, which addresses the manner in which the optimisation is implemented in the organisation

The action fields (together with their subunits) which play a role in each of the implementation phases are as follows:

- Context, which addresses Processes, Roles, Ontologies, and Search classification
- Content, which addresses Documents, and Information sources
- Information technology, which addresses User-interface, Reporting/System analysis, Software environment, Hardware environment, and Support
- Project control, which addresses Enterprise strategy, Change management, Training, Project management, and Cost consideration

A visual representation of the PreBIS implementation (focusing on the 'Context' action field) can be seen in Figure 7.4.

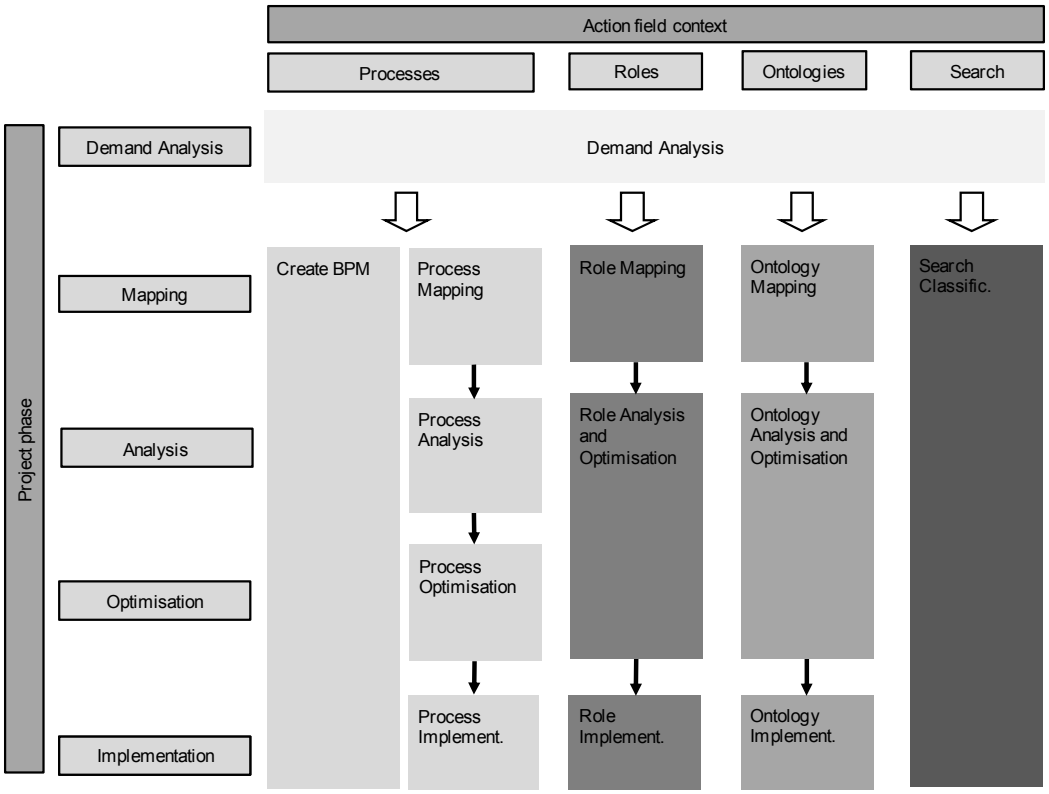


Figure 7.4: Service modules in the 'Context' action field (Source Böhm et al. 2005)

7.2.2 Methodologies focusing on different levels of detail as the conceptual framework implementation methodology

The methodologies that were found to focus on different levels of detail than that of the conceptual framework implementation methodology will only be listed below as a discussion of each are not really warranted within the scope of this thesis. In most of the cases the implementation of a specific information modelling approach are only one of the phases of these methodologies, and not enough information is provided to compare these methodologies with the conceptual framework implementation methodology in a fair manner. The methodologies are:

- The multi-facet taxonomy system (MTS) (Cheung, Lee & Wang 2005);
- The DILIGENT knowledge process (Vrandečić, Pinto, Tempich, Sure, 2005);
- The Inferential Modelling Technique (IMT) (Chan 2004); and
- A bottom-up strategy for enterprise ontology implementation (Lee et al. 2007)

7.2.3 Comparison of the conceptual framework implementation methodology with existing information modelling approach implementation methodologies

The conceptual framework implementation methodology that was developed through this thesis was compared with each of the methodologies discussed in section 7.2.1 to ascertain whether it is in line with similar, existing methodologies and whether it focuses on the same general aspects. A notable difference between the conceptual framework implementation methodology and the other methodologies used in the comparison is that the other methodologies all focus on the implementation of the *ontology* information modelling approach, whereas the conceptual framework implementation methodology obviously does not. This fact was however not deemed to influence the comparison, as it would only be relevant when comparing methodologies on a very detailed level.

It was decided to start comparing the methodologies on a phase level, and increase the level of detail until the similarity of the methodologies can be proven (or disproven). The phases constituting each of the methodologies (compared to those of the conceptual framework implementation methodology) can be seen in Figure 7.5.

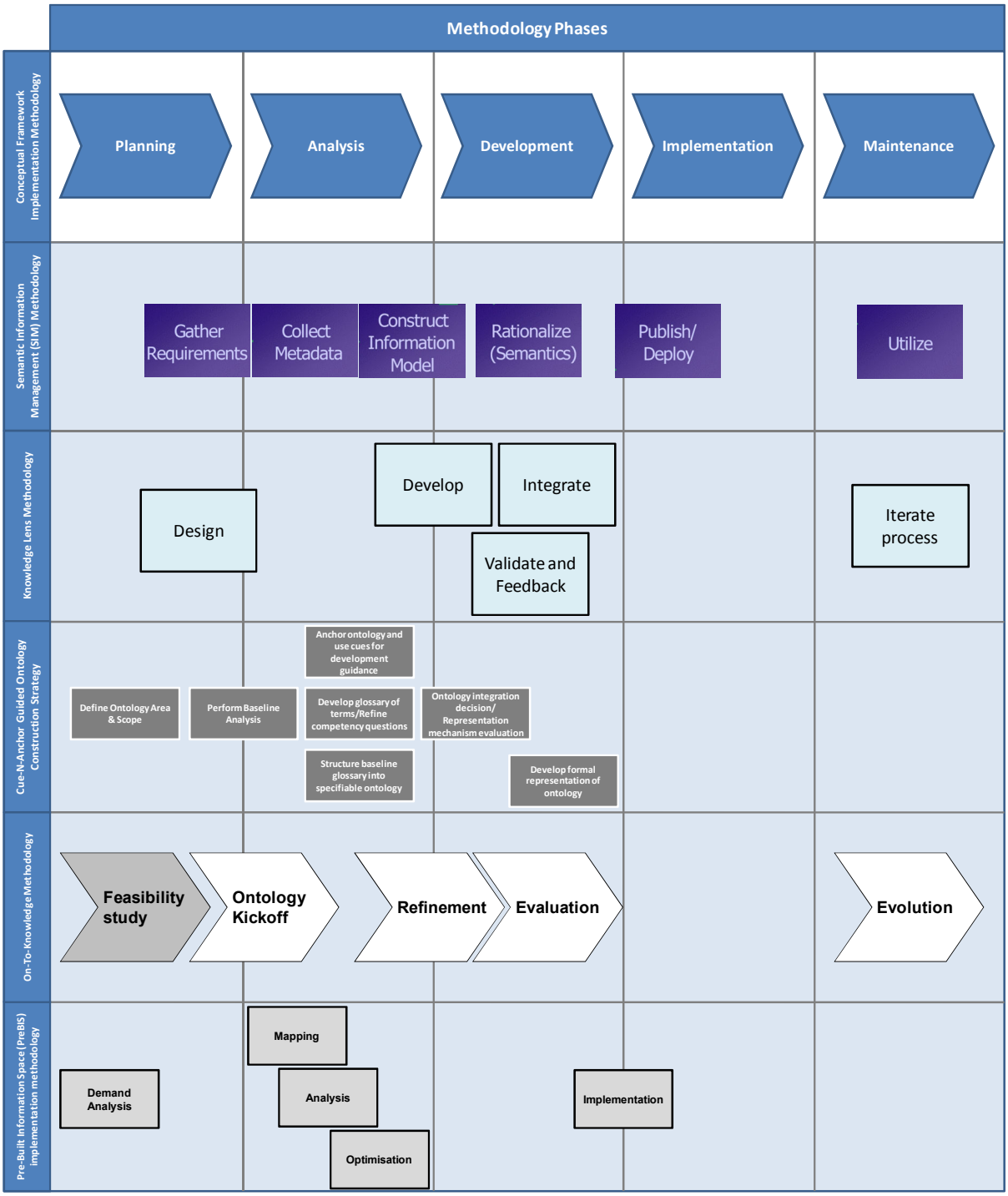


Figure 7.5: Comparison of conceptual framework implementation methodology with other similar information modelling approach implementation methodologies

The methodology phase comparison clearly indicates that, although there is a clear overlapping between the phases of the conceptual framework implementation methodology and the phases of the other methodologies, the manner in which the phases of the other methodologies have been defined prevents the overlapping from being very exact. When one however increases the level of detail when comparing the methodologies, and compare the required activities of each of these methodologies (see section 7.2.1) with those proposed by the conceptual framework implementation methodology, a

significant improvement in the overlapping will be observed. Table 7.2 shows that the activities of the other methodologies can very easily be restructured according to the phase-structure of the conceptual framework implementation methodology, and therefore largely address the same issues as the activities of the conceptual framework implementation methodology (except for the cases where some of the methodologies do not provide guidance regarding post-development and post-implementation strategies).

Table 7.2: Overlapping of activities from other information modelling approach implementation methodologies with the phases of the conceptual framework implementation methodology.

Information Modelling Approach	Conceptual Framework Implementation Methodology				
	Planning	Analysis	Development	Implementation	Maintenance
Semantic Information Management (SIM) Methodology	Yes	Yes	Yes	Yes	Yes
Knowledge Lens Methodology	Yes	Yes	Yes	*No*	Yes
Cue-N-Anchor Guided Ontology Construction Strategy	Yes	Yes	Yes	*No*	*No*
On-To-Knowledge Methodology	Yes	Yes	Yes	*No*	Yes
Pre-build Information Space (PreBIS) Implementation Methodology	Yes	Yes	Yes	Yes	*No*

The fact that the conceptual framework implementation methodology in general addresses the same issues as the other information modelling approach implementation methodologies discussed in this section, indicates that the reasoning behind it is conceptually sound, and that it compares well with tried-and-tested methods for implementing information modelling approaches in organisations. The methodology is thus deemed valid from a theoretical viewpoint.

An aspect of the conceptual framework implementation methodology which warrants special mention, however, is the fact that it places a specific focus on the deployment of the conceptual framework in the organisation with regards to information infrastructure configuration and the users of the approach – something that most of the other methodologies fail to do. These aspects were deemed crucial for the creation of an environment within the organisation that is conducive to the effective interaction with an organisational conceptual framework. It would therefore seem that the conceptual framework implementation methodology is, in this respect at least, more complete than the other methodologies that were used in the comparison. It does not, however, influence the validity of the methodology, and the conceptual framework implementation methodology therefore remains valid (from a theoretical point of view).

7.3 Practical Validation: Case Studies

This section discusses the execution of two case studies in which conceptual frameworks were implemented in two different organisations by utilising the implementation methodology discussed in this thesis.

7.3.1 Variations from the final implementation methodology

The process pursued throughout the case studies varies from the final version of the methodology because of the development and consequent refinement of the conceptual framework implementation methodology through the case studies discussed in this chapter. The two major changes that were made to the methodology will be briefly discussed in this section, and other minor changes will be discussed in the sections dedicated to each of the case studies.

The first major change that was made to the conceptual framework implementation methodology subsequent to the completion of the case studies, was to take the frequency of change of the information represented by an implemented conceptual framework into consideration. An activity was added to the Analysis phase of the implementation methodology, by which the frequency of change experienced by the information sources that form part of a conceptual framework's information domain is documented. In another new activity (added to the Implementation phase of the methodology), this information is utilised to develop a schedule that dictates how regularly the various processes aimed at updating information in the conceptual framework are executed. This improvement to the methodology therefore improved the ability of the implemented conceptual framework to consistently present updated information.

The second major change that was made to the conceptual framework implementation methodology was to introduce the utilisation of formal control documentation during an implementation (i.e. the Conceptual Framework Implementation – Planning phase, Conceptual Framework Implementation – Analysis phase, Conceptual Framework Implementation – Development phase and Conceptual Framework Implementation – Implementation phase documents). This change impacted each of the Planning, Analysis, Development and Implementation phases of the implementation methodology, and resulted in an improvement in the knowledge management of the implementation process.

7.3.2 Case study: Indutech (Pty) Ltd.

Indutech is a medium-sized industrial engineering company situated in Stellenbosch, South Africa which focuses on providing consulting services and solutions for enhancing client company projects through the incorporation of programme management, knowledge management and innovation management principles into the project processes. The company's core offering is in the form of methodologies and tools that assists its business engineers in enabling companies to perform Company Wide Innovation Management (What We Offer 2008).

Indutech was selected as the site for the first case study to test the conceptual framework implementation methodology as it was already heavily involved in conceptual framework research, and provided a benign testing environment where the worst of the methodology's creases could be smoothed out. Organon, a software tool for the creation, visual navigation and updating of conceptual frameworks, was also developed by Indutech, and was used in both of the case studies that are discussed in this chapter.

Because of the fact that this was the first utilisation of the developed conceptual framework implementation methodology, the conceptual framework that was constructed was quite small in size and was not based on very complex information requirements. The choice to first perform a smaller implementation was made purely to corroborate the relevancy of the aspects addressed by the methodology, and to easily correct any major issues that have not been foreseen during the development of the methodology.

- ***The implementation process***

The implementation process will be discussed in terms of the five phases of the conceptual framework implementation methodology.

Planning

Identify stakeholders

The first activities in the conceptual framework implementation methodology centre around the identification of stakeholders who will take part in the implementation. With the Indutech case study, the manager of Indutech's Integration Services and Support (ISS) department was identified and instated as the project champion and main stakeholder of the implementation (both in the managerial and information infrastructure sense), because of the following reasons:

- He was already heavily involved with conceptual framework research and therefore had an adequate understanding of the purpose of the implementation;
- He had a very good knowledge of the information landscape within the organisation; and
- He managed the department responsible for all information infrastructure administration of the organisation and therefore had the authority to allow access to any of the information sources that Indutech possess.

Initial information need identification

The domain of information within the organisation on which the implementation was to focus, was information on employee activity at Indutech, and the project champion requested information to be provided that was centred around answering the following questions:

- At what time do employees arrive at Indutech?
- Which documents have been worked on by employees on a particular day?
- What projects have been worked on by employees?
- What is the amount of time spent by employees working on a specific project?

It can therefore be seen that the information requirements provided have been stated in a fashion which allows sufficient flexibility in terms of the information that could be included in the conceptual framework. No grouping of information requirements were necessary, because of the small number of requirements provided.

With all activities that form part of the Planning phase concluded, the Analysis phase was undertaken.

Analysis

Examination of information sources pertaining to stakeholder-specified information requirements

With the assistance of the project champion, various information sources in Indutech that could sufficiently address the stated information requirements were identified, and access to these information sources obtained. These information sources were the following:

- Two EDEN™[25] databases (providing information on the documents Indutech employees work on);
- The database of the Indutech BioAccess biometric access control system (providing information on when Indutech employees arrive at work);
- The database of a previous Indutech time-keeping system (providing information on the projects which Indutech employees are involved with, as well as the amount of time that employees spend on these projects); and
- A Microsoft Excel spreadsheet, which is currently used for time-keeping at Indutech (also providing information on the projects in which Indutech employees are involved, and the amount of time they spend on these projects).

A study was made to gain an understanding of the content of these information sources, as well as the structure in which the information is stored. In the case of the EDEN™ databases and the database of the previous Indutech time-keeping system, the entity relationship diagrams were studied in order to gain a sufficient understanding of the information sources. The inter-relationship of the information sources were also studied, but the systems were found to function relatively independent of each other and that only very minor overlapping of information occurred between the information sources. Finally, it was determined that the information in the analysed information sources could sufficiently meet the project champion's information requirements. As mentioned in section 7.3.1, the frequency of change in information sources did not play a role in this case study.

Interestingly enough, it was found during the study of the content and structure of the information sources (and the subsequent development of the type-relation structure used as a specification for the conceptual framework) that the information needs identified by the project champion could not in all cases be addressed by the information sources found in the implementation domain, or anywhere else in the organisation. This state of affairs called to attention the existence of information that only exists as a shared understanding between a number of Indutech employees, and not explicitly within an organisational information source, as well as the need to capture and include this information in the conceptual framework specification in order to create a consistent and complete view of the organisation's information.

[25] *EDEN™ is another software tool developed by Indutech, and is a platform used by its business engineers, encapsulating (amongst others) various innovation management, project management, knowledge management and document management functionalities.*

Development of conceptual framework specification

From the analysis that took place earlier in this phase of the conceptual framework implementation methodology, fourteen (14) conceptual framework types were identified. These types, together with their descriptions, can be seen in Table 7.3.

Table 7.3: Names and descriptions of conceptual framework types identified in Indutech conceptual framework implementation case study

Type Name	Type Description
Client	Companies that are clients of Indutech
Date	Any dates that exist in Indutech information sources
Department	Departments of Indutech
Document	Documents that reside in EDEN™ databases
EDEN	Denoting the names of specific Indutech EDEN™ databases
EDENRights	User access rights with regards to documents found in EDEN™ databases (e.g. Read, Write, Delete, etc.)
EDENRole	Different groups of users in EDEN™ databases (e.g. Administrators, Consulting, Management, etc.)
E-mail	E-mail addresses of Indutech employees
Keyword	Keywords used by Indutech employees for describing documents in EDEN™ databases
Person	Indutech employees
Project	Projects conducted by Indutech internally or for client companies
Project_Role	Roles of Indutech employees in projects
Time	Any time-stamps that exist in Indutech information sources
Work_Duration	The amount of time spent by Indutech employees on projects

A total of forty-four (44) relations were defined between the types, by employing the matrix approach mentioned in section A – Planning phase. This is a great deal less than the $\frac{1}{2}(n^2+n) = \frac{1}{2}(14^2+14) = 105$ potential relations predicted by the conceptual framework implementation methodology, and makes for a very usable framework for interacting with the information. A type-relation structure was created and refined at the same time, resulting in the structure shown in Figure 7.6. The diagram however shows different relations for both A→B and B→A (with A and B being types and → being the relation between them). This does not insinuate that

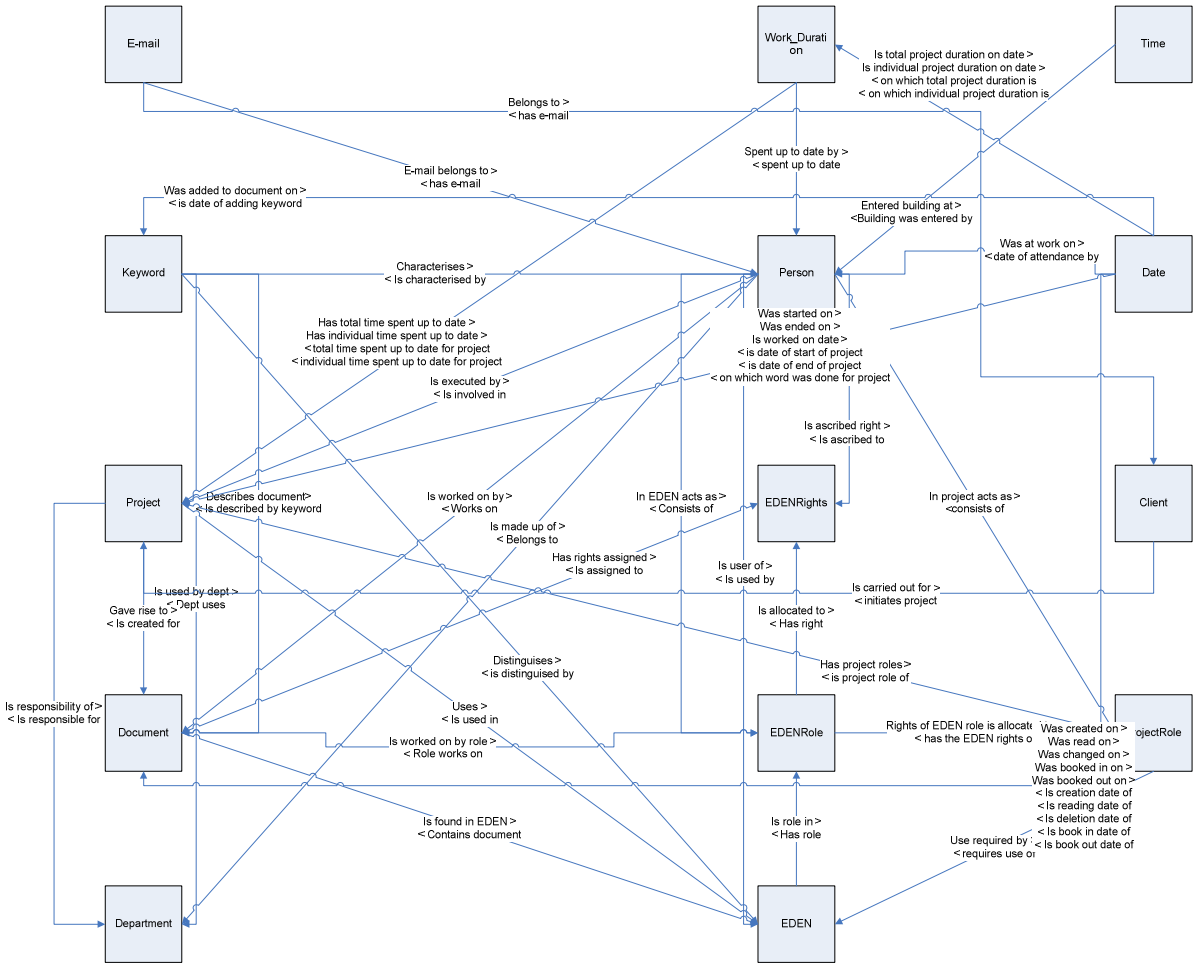


Figure 7.6: Refined type-relation structure of the Indutech conceptual framework implementation case study

the conceptual framework facts in this type-relation structure are not multi-directional – it only indicates a choice provided to the persons implementing the conceptual framework of text to use when constructing the conceptual framework in the conceptual framework software tool. The facts which ultimately made up the Indutech conceptual framework are listed in Appendix B.

Upon completion of the type-relation structure, it was verified by the project champion on both the information and logical level, which marked the end of the Analysis phase of the implementation.

Development

Fulfil conceptual framework hardware requirements

Once the size and complexity of the planned conceptual framework was known, sufficient data processing capabilities and storage space were obtained with the help of the project champion to facilitate the development of the Indutech conceptual framework.

Create and populate conceptual framework

In the case of the Indutech conceptual framework implementation, it was found that the conceptual framework could very easily be constructed by transferring the data contained in the respective information sources of the implementation domain to a Microsoft (MS) SQL Server database implementation, from where a mapping layer extension of the Organon conceptual framework software tool could obtain the data and transform it into the conceptual framework denoted by the type-relation structure shown in Figure 7.6.

Information from the implementation information sources were transferred to such a MS SQL Server database implementation at Indutech, with only a few complications:

- A small set of information was identified which did not exist explicitly within the implementation domain (see section B – Analysis phase), but rather as a shared understanding between Indutech employees. This information (the names of Indutech employees together with the corresponding departments they belong to) was however deemed crucial for ensuring the consistency and completeness of the Indutech conceptual framework, and was therefore collected formally in a MS Excel spreadsheet, which was subsequently transferred to the MS SQL Server database implementation.
- It was found that, in some cases, references to certain entities may differ in spelling from one information source to another. This clearly had severe consequences for accommodating the overlapping of information between these information sources, which necessitated cleaning of this data, ensuring that references to entities residing in different information sources are all spelled in the same manner. This problem of the existence of fractured information within the organisation's information sources is however not unique to Indutech, and will probably occur in most organisations. For this reason, this aspect is something to be prepared for when performing conceptual framework implementations, as it may require a large amount of time to address.
- Finally, it proved very labour intensive to transfer information from the MS Excel spreadsheet utilised by Indutech for timekeeping to the MS SQL Server database, due to the manner in which the data was captured in this spreadsheet.

Once all the relevant information resided in the MS SQL Server database, the mapping layer extension of the Organon conceptual framework software tool was used to compile this information into a conceptual framework (which will be discussed in more detail in the Final Result section). Finally, the creation and population of the conceptual framework was concluded through the investigation and verification of the conceptual framework's rationality, which was found to be intact.

Verify information represented by conceptual framework

A multi-disciplinary group of persons from Indutech and the GCC (see also section 5.1) was selected to verify the information represented by the Indutech conceptual framework, as well as to share their impressions of the conceptual framework approach and the implementation methodology. This group consisted of three researchers from the GCC, two software developers of Indutech, and the project champion.

The verification was done by providing each of the participants with a copy of the Indutech conceptual framework and a copy of the Organon software tool, and giving them the opportunity to first-hand experiment and interact with the conceptual framework. A questionnaire was also provided (see Appendix C), through which the impressions and suggestions of each person were captured. The contents of all these questionnaires were consolidated, and discussed during a final verification session held with the group. The issues discussed and the conclusions of these discussions will be discussed in more detail in the Observations section. The various participants of the verification activity agreed that the information represented by the conceptual framework was correct, and thereby verified the Indutech conceptual framework.

Implementation and Maintenance

After the conclusion of the Development phase of the conceptual framework implementation process, a decision was made by the project champion not to continue rolling out the approach to the rest of Indutech. This was due to the fact that there lacked a need for such a tool at the time of the implementation, and therefore the amount of effort required to perform the final two phases of the conceptual framework implementation process was not warranted. The case study concerning the implementation of a conceptual framework at Indutech was thereby concluded with the completion of the Development phase.

From the point of view of the persons performing the implementation of the Indutech conceptual framework, the implementation was a success. It was also deemed that the methodology functioned very effectively and that it addressed relevant issues during the different phases of the conceptual framework implementation.

• *Final Result*

The final result of the Indutech conceptual framework implementation case study was a conceptual framework consisting of more than 240,000 facts (see section 5.2.1), providing a view on the employees of Indutech, and the manner in which they work. Types of information that are

represented by the conceptual framework include the documents people are working with, the projects people are involved with, the time spent by persons on certain projects, the time people arrive at work, etc. In the rest of this section a number of Organon screenshots which display some of the information contained in the Indutech conceptual framework are briefly discussed. Views such as these were utilised by participants in the validation activities of the Analysis phase to check the information represented by the conceptual framework for correctness:

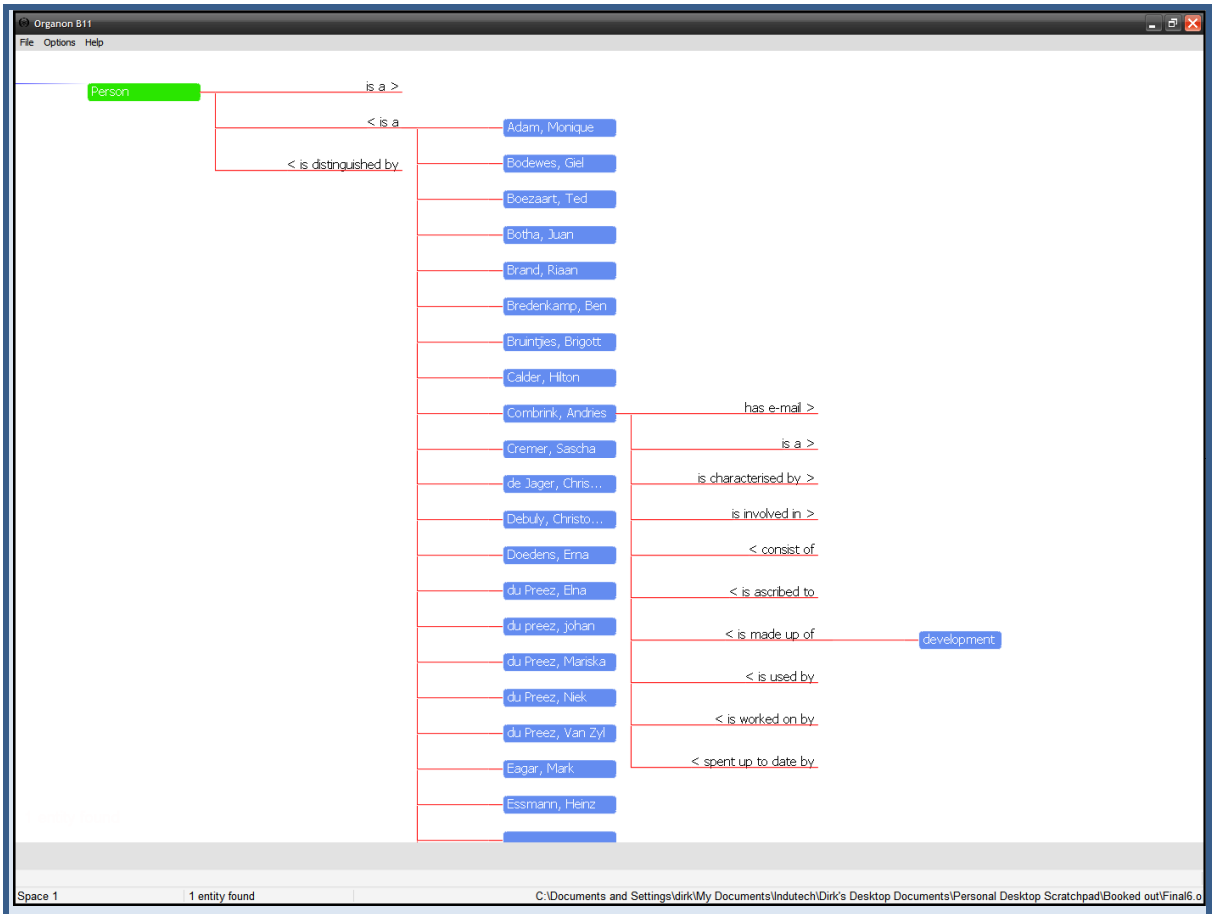


Figure 7.7: Screenshot 1 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.7 shows the employees of Indutech (or **"Person"** entities), and some of the relations that pertain to entity "Combrink, Andries". It specifically indicates that this person belongs to the "Development" department[26], and also provides access to the following information on this person:

- The person's e-mail address
- The entity type of the person
- The keywords added by this person to Indutech documents (and which thereby characterises this person)
- The projects in which this person participated
- The roles this person assumed within the EDEN™ software environment
- The document access rights ascribed to the person within the EDEN™ software environment
- The EDEN™ databases that this person has accessed
- The documents on which this person has worked
- The time this person has spend up to a certain date on different projects

[26] *The "<" and ">" characters indicate the direction in which the text of a relation is applied, e.g. "Adam, Monique"* "is a" *"Person".*

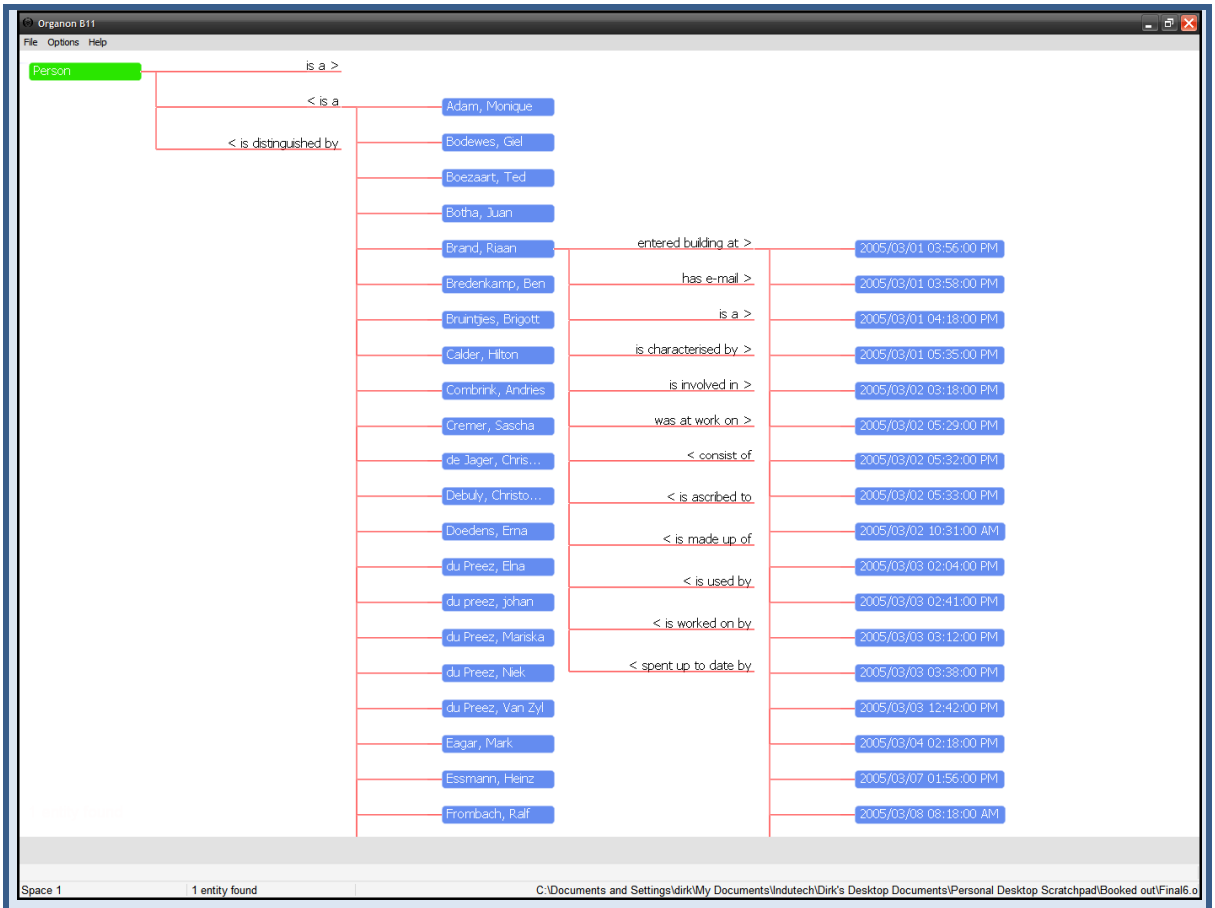


Figure 7.8: Screenshot 2 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.8 shows the relations of **"Person"** entity "Brand, Riaan" and focuses on dates and times which he arrived at Indutech, and entered the premises through use of the BioAccess biometric access control system. Access to the following information on this person is also provided:

- The person's e-mail address
- The entity type of the person
- The keywords added by this person to Indutech documents (and which thereby characterises this person)
- The projects in which this person participated
- The dates on which this person was at work
- The roles this person assumed within the EDEN™ software environment
- The document access rights ascribed to the person within the EDEN™ software environment
- The department to which this person belongs
- The EDEN™ databases that this person has accessed
- The documents on which this person has worked
- The time this person has spend up to a certain date on different projects

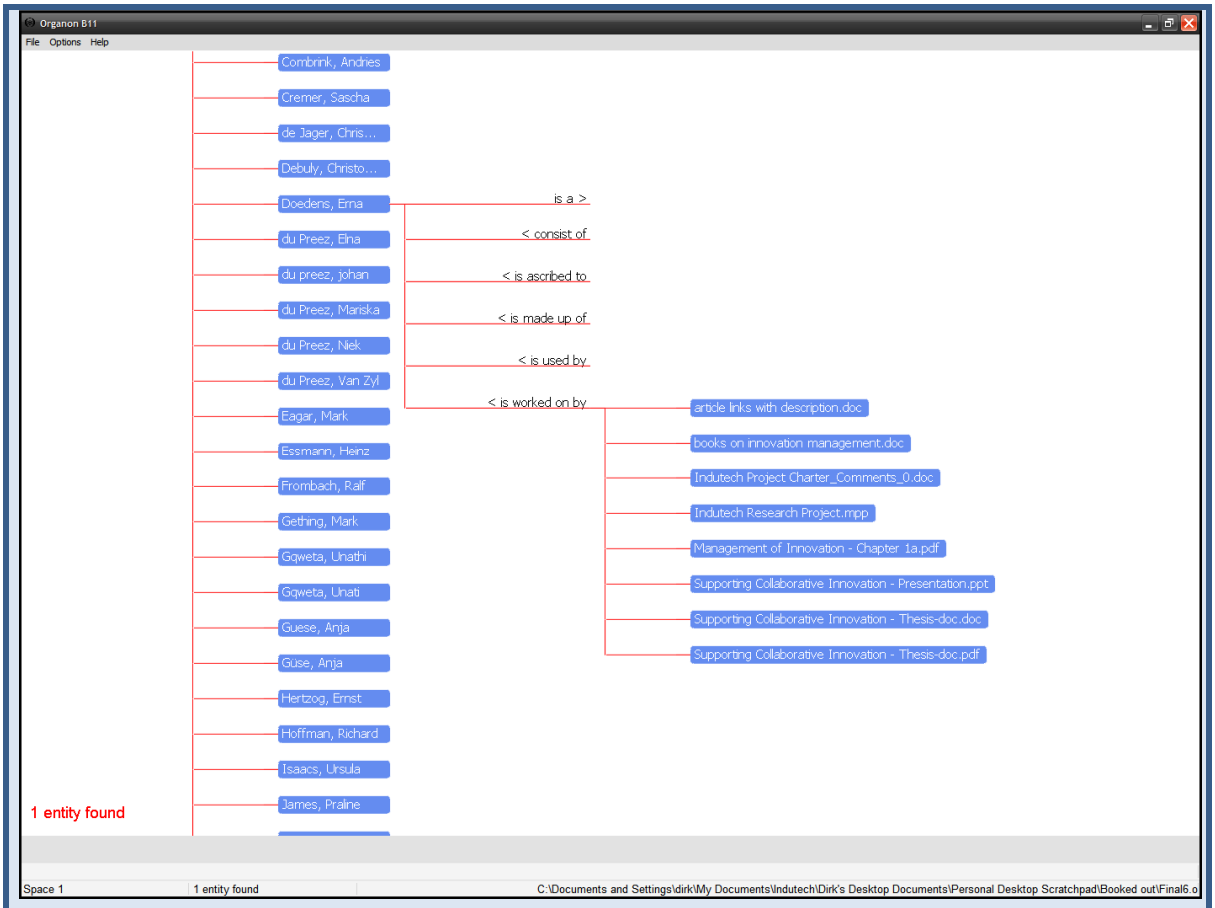


Figure 7.9: Screenshot 3 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.9 shows the **"Person"** "Doedens, Erna" and the documents residing in the EDEN™ databases which she had worked on. Access to the following information on this person is also provided:

- The entity type of the person
- The roles this person assumed within the EDEN™ software environment
- The document access rights ascribed to the person within the EDEN™ software environment
- The department to which this person belongs
- The EDEN™ databases that this person has accessed

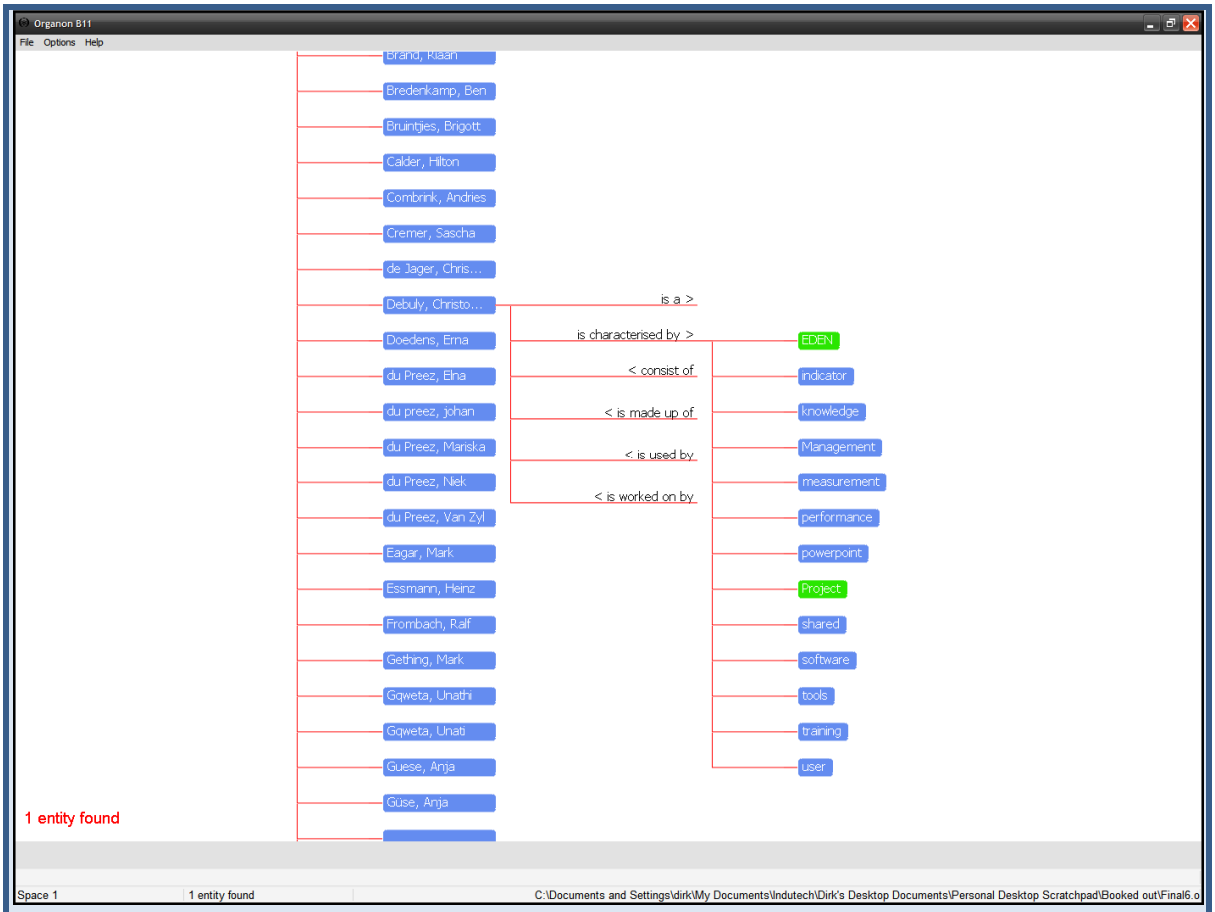


Figure 7.10: Screenshot 4 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.10 shows the **"Person"** "Debuly, Christophe" and the keywords that he has attached to documents in the EDEN™ databases that he has worked on. This view can be useful for a person interested in knowing what type of work another person is involved with. Access to the following information on this person is also provided:

- The entity type of the person
- The roles this person assumed within the EDEN™ software environment
- The department to which this person belongs
- The EDEN™ databases that this person has accessed
- The documents on which this person has worked

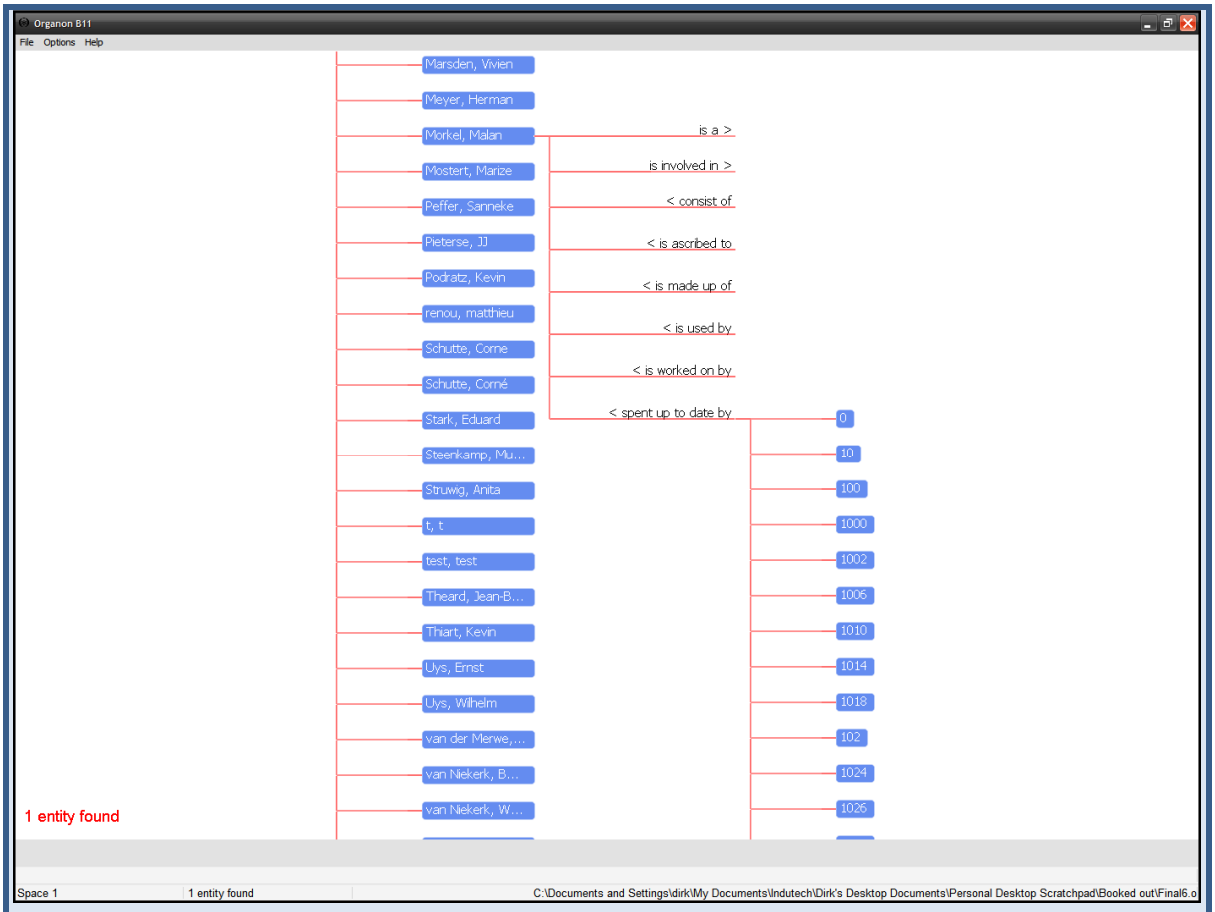


Figure 7.11: Screenshot 5 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.11 shows the **"Person"** "Morkel, Malan" and the amount of hours he has spent on different projects (both daily and up to date totals), as well as access provided to the following information on this person:

- The entity type of the person
- The projects in which this person participated
- The roles this person assumed within the EDEN™ software environment
- The document access rights ascribed to the person within the EDEN™ software environment
- The department to which this person belongs
- The EDEN™ databases that this person has accessed
- The documents on which this person has worked

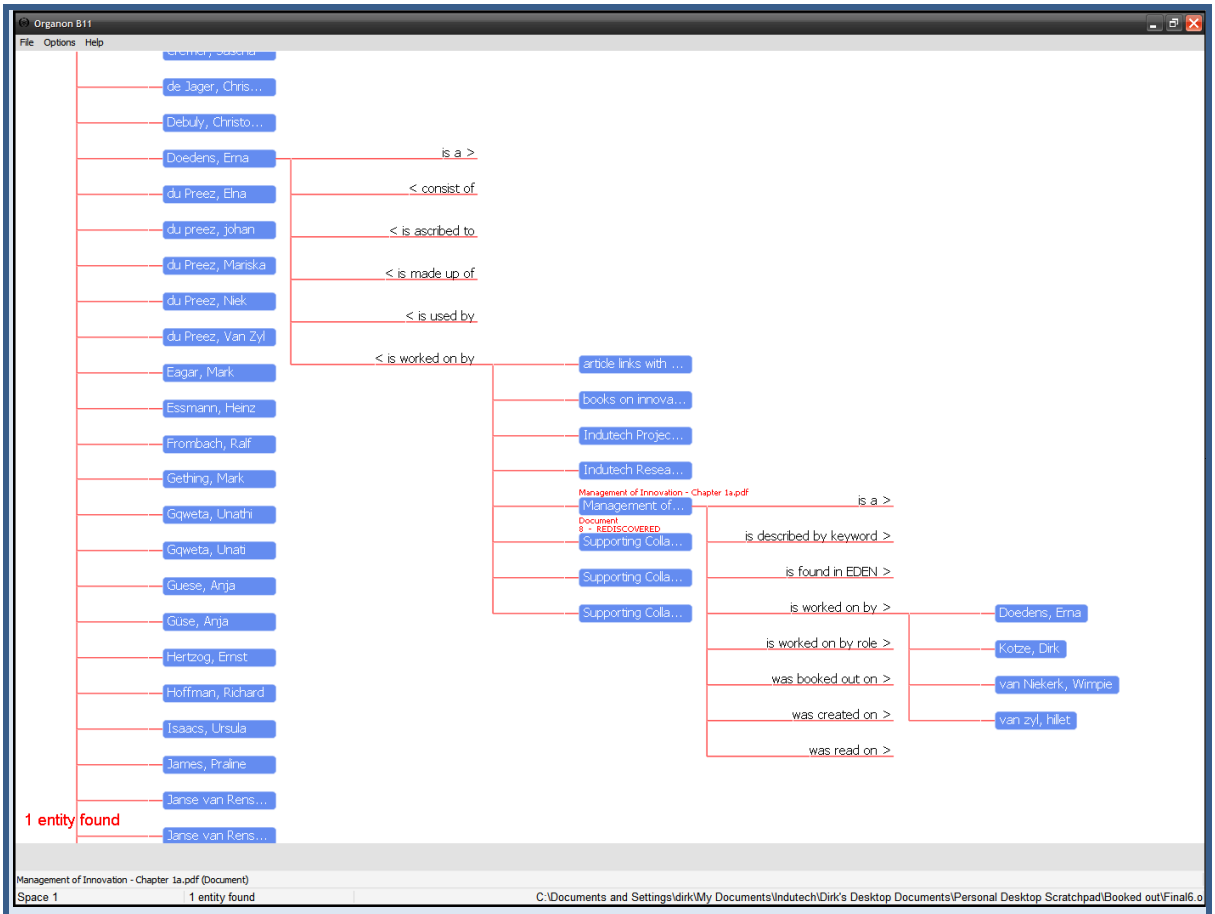


Figure 7.12: Screenshot 6 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.12 shows the **"Document"** "Management of Innovation – Chapter 1a.pdf" which was worked on by the **"Person"** "Doedens, Erna", as well as the names of all persons that has worked on the document, namely "Doedens, Erna", "Kotze, Dirk", "van Niekerk, Wimpie" and "van zyl, hillet". Access to the following information on this document is also provided:

- The entity type of the document
- The keywords that have been used to describe this document
- The EDEN™ database in which this document resides
- The EDEN™ roles of the persons that have worked on this document
- The dates on which a user have booked this document out of EDEN™
- The date on which this document was created in EDEN™
- The dates on which users have accessed this document in EDEN™

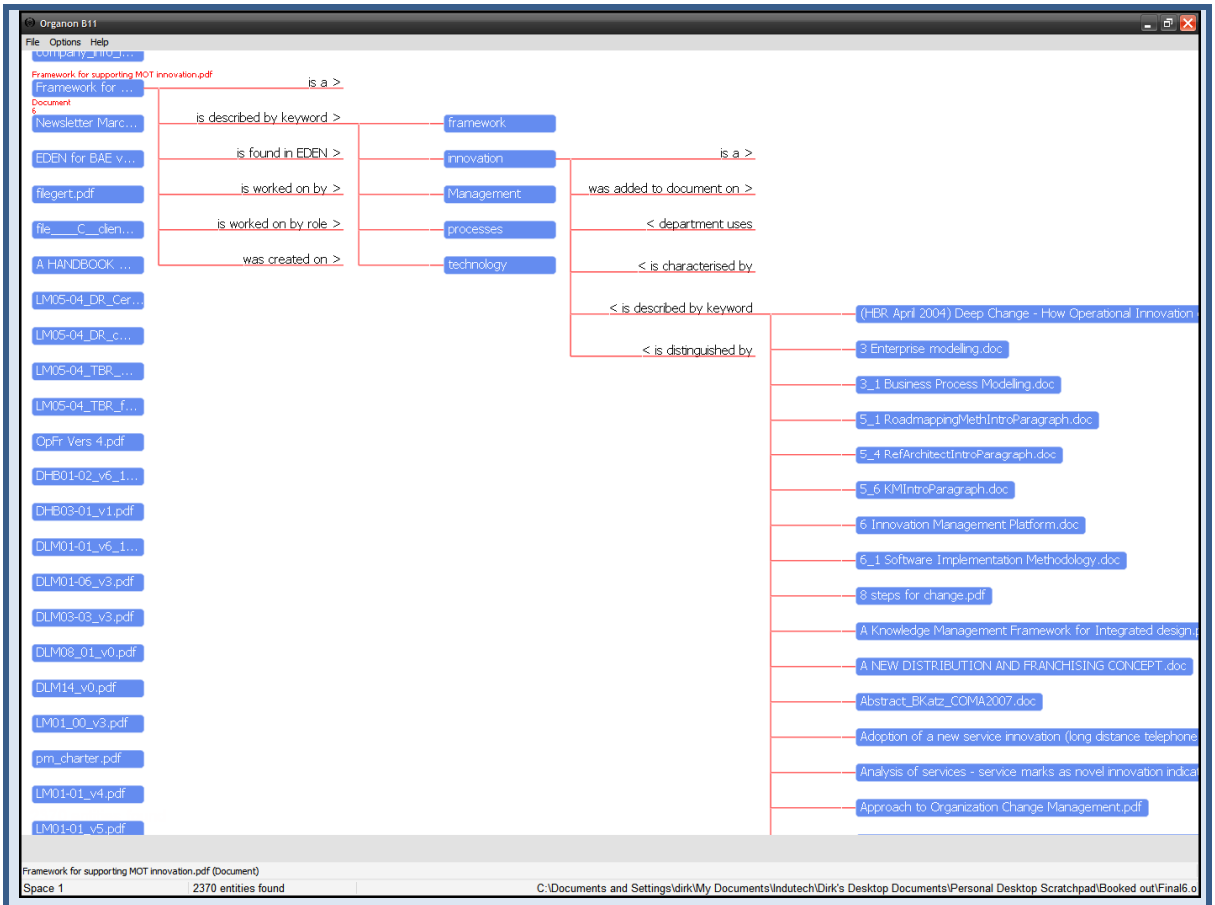


Figure 7.13: Screenshot 7 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.13 shows the **"Document"** "Framework for supporting MOT innovation.pdf" together with the keywords that was used to describe it in the EDEN™ database in which it resides. Other documents are also shown which share the use of the keyword "innovation" with "Framework for supporting MOT innovation.pdf". Access is also provided to the following information on this keyword:

- The entity type of the keyword
- The date this keyword was associated with a document
- The departments that make use of this keyword
- The persons that have used this keyword
- The EDEN™ database in which this keyword is used

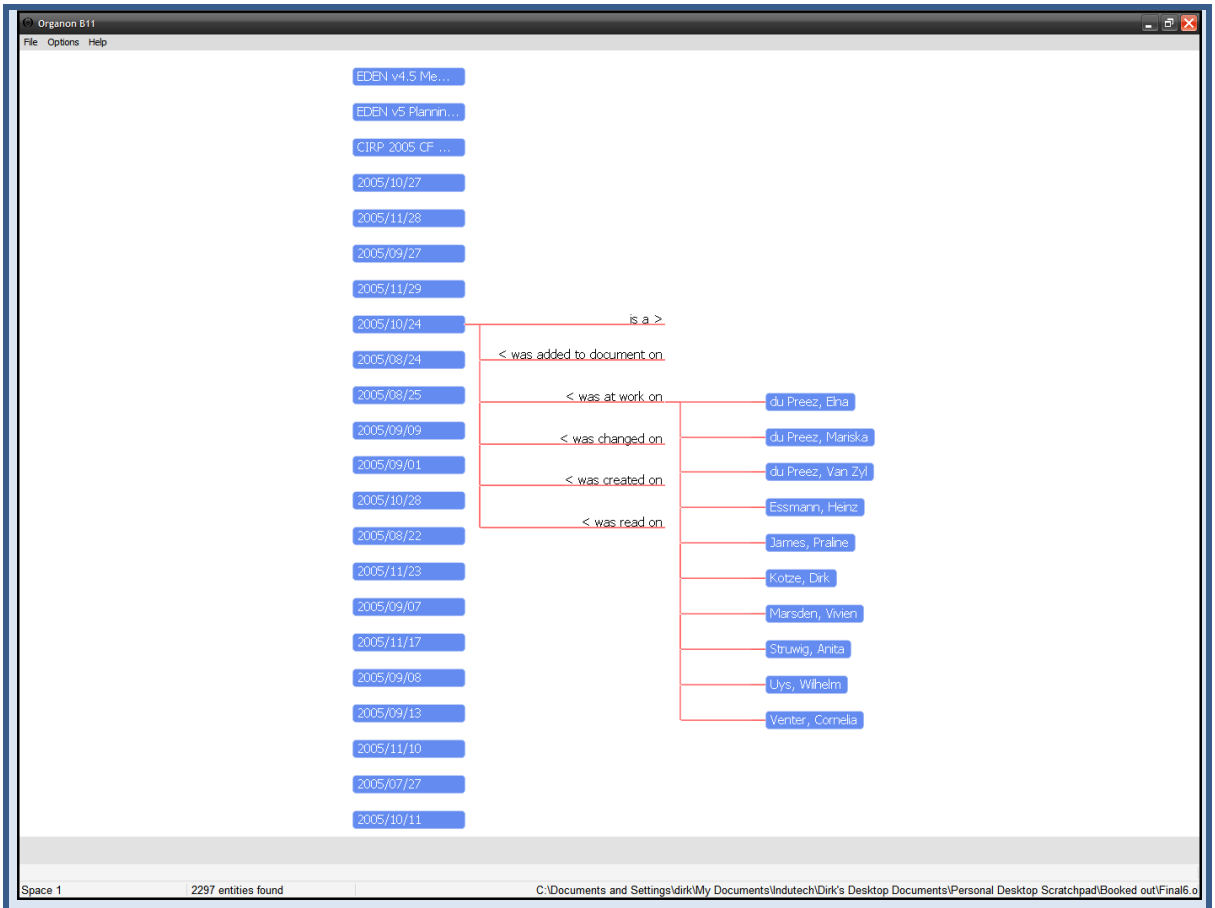


Figure 7.14: Screenshot 8 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.14 shows the **"Date"** "2005/10/24", as well as the persons who was at work on that day (which is strictly speaking not entirely correct, as employees often arrive at Indutech in groups, and usually only one person from such a group makes use of the BioAccess biometric access control system to ensure access to the premises for all). Access is also provided to the following information associated with this date:

- The entity type of the date
- The keywords added to documents on this date
- The documents that changed in EDEN™ on this date
- The documents that was created in EDEN™ on this date
- The documents that was accessed in EDEN™ on this date

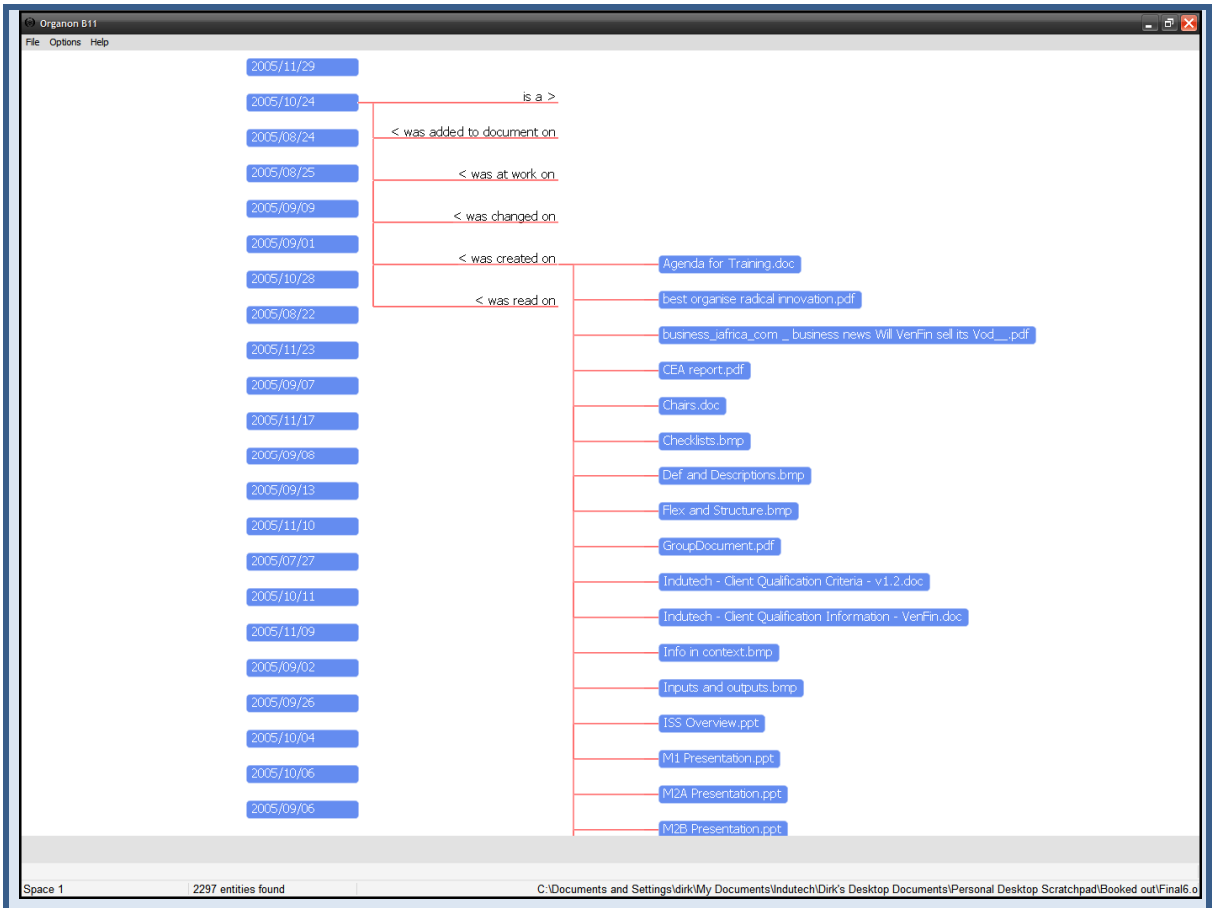


Figure 7.15: Screenshot 9 of Organon showing information of the Indutech conceptual framework implementation case study

Figure 7.15 shows the **"Date"** "2005/10/24", as well as the documents that were created on this date, together with access provided to the following information associated with this date:

- The entity type of the date
- The keywords added to documents on this date
- The persons who were at work on this date
- The documents that changed in EDEN™ on this date
- The documents that was accessed in EDEN™ on this date

- ***Observations***

This section will discuss some important aspects of the conceptual framework implementation and the final conceptual framework that was identified and discussed by participants involved in the verification of the conceptual framework. The discussion is structured around the main focus areas of the questionnaire that the participants had to complete.

Benefits of the Indutech conceptual framework

There was an overall consensus from the participants in the verification of the conceptual framework that it provides a useful and unique view on information that resides within Indutech. In general it was felt that the conceptual framework succeeds in presenting a sufficient overview of the Indutech operations, as well as providing useful insight into the relations that exist between entities. It was furthermore felt that, by taking the overlapping of information into consideration, the searching possibilities of an organisation are drastically increased, as a person is able to easily locate information by starting with a piece of information that is known and, through relations that exist within the organisation, navigate over departmental boundaries to the sought-after information. Specific facts that participants found to be most useful are the following:

- **"Person"** *"is characterised by"* **"Keyword"** (which can give an indication of the interests of a person, or the work he/she is involved with);
- **"Document"** *"is worked on by role"* **"EDEN Role"** (which can give an indication of the relative importance of a specific document);
- **"Document"** *"was created on"* **"Date"** (indicating the date on which a document was created);
- **"Document"** *"is described by keyword"* **"Keyword"** (which can give an indication of the contents of a document according to its author(s)); and
- **"Person"** *"is involved in"* **"Project"** (which can give an indication of the projects that a person have been involved in).

Shortcomings of the Indutech conceptual framework

The only shortcomings that were identified by the participants in the verification of the conceptual framework, focused on specific information that was represented by the conceptual framework, and not on the fundamental manner in which the conceptual framework was constructed. In most of the cases the shortcomings were caused by relations not being specified clearly enough for all participants to understand, or certain groups of data from the organisations' information sources not being complete, which were then reflected as such in the Indutech conceptual framework. A

specific aspect of the Indutech conceptual framework which failed to deliver the expected results was the relating of different entities within the conceptual framework to the **"Work Duration"** group of entities. This group consisted of the time durations that Indutech employees spent working on projects, and contained both aggregate and singular values which caused confusion on the part of the verification participants. Such a value may for instance have been the time a single person spent working on a project on a specific day, a total number of hours spent by a number of persons on a project, the number of hours spent up to a certain date by all persons involved on a project, etc. Participants struggled to use this information in useful ways, and suggested that the work durations be separated into more detailed entity type groupings, to increase its potential usefulness. Another aspect which was deemed a shortcoming of the conceptual framework by the participants is the fact that they were unable to view and interact with complex facts, which forms part of the conceptual framework approach (e.g. Person X worked for Hours Y on date Date T on project Project G) (also see section 5.2.1). Although observed as a shortcoming of the conceptual framework, this was however actually a shortcoming of the Organon software at the time, and was not due to the conceptual framework being constructed in an incorrect manner.

Growing the Indutech conceptual framework

Even though the Maintenance phase of the conceptual framework implementation process was not executed in the case of the Indutech case study, the participants in the verification of the conceptual framework were asked to provide suggestions regarding the manner in which the Indutech conceptual framework could be grown[27]. Most of the suggestions however focused on improving the information that the Indutech conceptual framework currently possesses, and not on the adding of new domains of information to the existing conceptual framework. Some of the suggestions provided are listed below:

- Data from some of the Indutech information sources has to be cleaned in order to improve the usefulness and consistency of the information divulged by the conceptual framework. This cleaning has to be done either through additional mappings or through improved operational procedures.
- Some of the facts lack usefulness due to their generality, and should be split into more detailed facts to address this issue. As an example, a relation *"is worked on by"* between entity types **"Document"** and **"Person"** (i.e. **"Document"** *"is worked on by"* **"Person"**) could be replaced by *"was edited by"*, *"was created by" and "was deleted by"* to enrich the conceptual framework (should this information be available and accessible).

[27] *All of these suggestions focused on the information contained in the conceptual framework, and no suggestions were made that required any changes to the conceptual framework implementation methodology.*

- It can be seen that the BioAccess biometric access control system database is not a very authoritative source of information, as it does not record that persons are at work if they did not physically use the system to gain access to the Indutech premises. A better source of information should therefore be located, or operational procedures adjusted to address this issue.

Conceptual framework implementation methodology

The participants in the verification of the conceptual framework were requested to comment on the manner in which the conceptual framework was implemented and to provide their impressions of the conceptual framework implementation methodology that was used.

It was mentioned that the Indutech conceptual framework implementation involved a great deal of data cleaning, which raised the question whether data cleaning should not be added as a phase or activity in the implementation methodology. Conceptual framework implementation however requires a certain level of maturity in the manner which an organisation understands and interacts with its information. It should be understood that the implementing of a conceptual framework is not a mechanism to directly address operational issues which should in actual fact not even exist. If the cleaning of organisational data is a necessity, it should be performed in parallel with the conceptual framework implementation, or even more preferably, should be driven by requirements provided by users of an already-implemented organisational conceptual framework. Through their interaction with the conceptual framework, users should encounter discrepancies in the organisation's information and their input could help to focus any data cleaning efforts, ensuring that information with higher importance to the organisation is corrected more rapidly.

The importance of the shelf-life of information was also discussed, and how to determine when a piece of information has out-lived its usefulness. The discussion focused especially around information pertaining to previous Indutech employees, and whether (and when) the details of these persons should be removed from the Indutech conceptual framework. This is however a decision that has to be made at a higher level of the organisation. Ideally, one would like to retain such information in the organisation, and have it available in the organisational conceptual framework should it be required. Special relations may furthermore be applied to conceptual framework entities which have out-lived their usefulness, and hidden from users by means of filters that focus specifically on these relations. In this way, the information is conserved, but do not clutter a user's view of the information, thereby risking information overload. In this discussion it was also observed that the right updating frequency of an organisational conceptual framework is important to ensure that the conceptual framework presents current information without consuming exorbitant information infrastructure resources. This discussion led to the inclusion of tasks in the conceptual framework implementation methodology which addresses the

aspect of scheduling the update of a conceptual framework in a manner that achieve a balance between information currency and the load on information infrastructure resources.

It was acknowledged that the existence of both explicitly documented information and implicit information complicates the conceptual framework implementation process. Such information should be captured in an explicit manner for easier incorporation in a conceptual framework, or some other strategy should be devised through which implicit information in the organisation is also represented and regularly updated in the organisational conceptual framework.

It was suggested that design and/or control objectives be added to the conceptual framework implementation methodology in order to ensure the success of such an implementation. For this reason, the inclusion of formal conceptual framework implementation documentation was added to the methodology. It requires the implementers of a conceptual framework to report on the various activities that take place during the implementation, and if certain activities are not performed, to provide the decision that gave rise to their omission.

In conclusion, the participants in the verification of the Indutech conceptual framework agreed that the implementation was done successfully and in an acceptable manner. They also expressed their satisfaction with the methodology, and agreed that it is entirely applicable and useful for conceptual framework implementation in organisations.

Other observations

Other observations that were made by the participants in the verification of the Indutech conceptual framework were centred more around their experience of using the Organon conceptual framework software tool to interact with the Indutech conceptual framework. Both positive and negative experiences were discussed, but the general feeling was that the software performed as expected. Based on suggestions and bugs found in the Organon tool, changes were suggested which were ultimately included in the specification for the development of the next version of the software.

• *Conclusion – Indutech (Pty) Ltd. case study*

The Indutech conceptual framework implementation was seen as a success by all participants involved, and through the case study a number of important conclusions could be made:

- By employing the conceptual framework implementation methodology discussed in this thesis, a conceptual framework of organisational information can successfully be created and implemented in the organisation.

- The manner in which a conceptual framework presents information is beneficial to an organisation and provides a means to more rapidly locate sought-after information. This may result in improving the organisation's ability to adjust to its environment and innovate more rapidly, and gain a competitive edge over its competitors.
- The consistency of the data in an organisation's information sources has a significant impact on the time and effort that is required to implement a conceptual framework in an organisation. The methodology assumes that an organisation that wishes to implement a conceptual framework will be mature enough as to ensure the consistency of its information, and therefore the activity of data cleaning is not included as part of the conceptual framework implementation methodology.
- The shelf-life and updating frequency of information plays a major part in ensuring the relevancy and currency of a conceptual framework, and should be addressed through the methodology.[28] Organisational preferences regarding the shelf-life and updating frequency of information do however influence the load placed by the conceptual framework on the organisational information infrastructure, and should therefore be managed carefully to ensure optimal performance.

Following this case study various changes were made to both the conceptual framework implementation methodology and the Organon conceptual framework software tool, which were employed later on in a second case study performed as part of the conceptual framework implementation methodology validation. This second case study will be discussed in the next section.

7.3.3 Case study: Stellenbosch University – Information Technology

Stellenbosch University is situated in the Western Cape town of Stellenbosch, South Africa, and was founded in 1866 when it first opened its doors with about 500 students and 39 lecturers. Since then it has grown to one of the four top research universities in South Africa with more than 22 500 students, 800 lecturers and some 50 research and service bodies.

"The University has ten faculties, of which eight – AgriSciences, Arts and Social Sciences, Education, Engineering, Law, Science, Theology and the larger part of Economic and Management Sciences - are located on the main campus in Stellenbosch with the Faculty of Health Sciences situated on the Tygerberg campus. The Bellville Park campus is home to the Business School and Business School -

[28] *At this stage the methodology did not yet address this aspect.*

Executive Development and the coastal town of Saldanha serves as the base for the Faculty of Military Sciences" (About Stellenbosch University 2007).

The Information Technology (IT) division of Stellenbosch University (SU) (which will henceforth be referred to as the *organisation* for consistency's sake) was approached to be the site for the second case study for testing the conceptual framework implementation methodology. This site was chosen because the department could provide access to a wide variety of different types of information (much like in a typical organisation) through which the implementation methodology could be thoroughly tested, whilst also standing sympathetic towards research conducted into information management by a student of the university, therefore providing more assistance and access to information than could otherwise have been expected from a more typical organisation. Therefore, similar to the Indutech case study, the SU IT conceptual framework implementation also took place in a somewhat benign environment, but the extent of the implementation was much larger than the one conducted previously, and was aimed to properly test the implementation methodology and software tool that was developed.

• *The implementation process*

As with the first case study, the implementation process of the SU IT conceptual framework will be discussed in terms of the five phases of the conceptual framework implementation methodology.

Planning

Identify stakeholders

In the case of the SU IT case study, a project champion was not selected, as the persons involved from the organisation's side preferred to play a less prominent role in the driving of the implementation process, and wished rather to be involved only where information identification, extraction and verification was required. A mixed team from the GCC and Indutech therefore performed the largest portion of the implementation with the organisational stakeholders only contributing to certain activities of the SU IT conceptual framework implementation process. Based on their respective roles in the process, the participants from the organisation were divided into *principal*, *primary* and *secondary* stakeholders. The designations of these persons are indicated below (see also Appendix D):

- **Principal Stakeholder**
    - Senior Director: Information Technology

- **Primary Stakeholders**
    - Director: Administrative Information
    - Director: User Services
    - Director: e-Business and Card Facility Manager
- **Secondary Stakeholders**
    - Manager: Student Information Systems
    - Head Engineer: Systems Integration
    - Contact person – WebCT[29] Reporting Information
    - Manager: Access Control
    - Contact person – Printing Transaction Information
    - Senior Software Specialist: Library Services

Initial information need identification

The stakeholders of the SU IT conceptual framework implementation decided to let the case study take the form of a pilot study for the ascertaining of the benefits and shortcomings of the conceptual framework based approach for viewing and utilising organisational information. The domain of organisational information on which the implementation was to focus, was information on student interaction with the organisational information infrastructure. As sources of information for the SU IT conceptual framework, six domains of information within the organisational environment were identified (listed below):

- Card reader transaction information, which is information captured whenever a student swipes his/her student card through an organisation card reader. Five different types of card reader transactions were identified to be included in the implementation, namely access control transactions, photocopying transactions, washing transactions, meal transactions and vehicle pool transactions.
- Demographic and study programme information, which is student demographical information and information on the modules, programme, etc. that a student is currently enrolled for.[30]
- Internet usage cost information, which focuses on the amounts of internet bandwidth utilised by students over/during certain periods.[31]

[29] *An e-learning system, now owned by Blackboard Inc. (http://www.blackboard.com)*

[30] *Due to obvious privacy issues, this information did not include the marks that students achieved for any modules or the respective programmes they are enrolled for.*

- Library transaction and item information, which is information on transactions conducted by students at any of the libraries of the organisation, as well as information on the library items found within the organisation's library system.
- Printing transaction information is information captured whenever a student performs a printing transaction using an organisation printer.
- WebCT usage information, which is information focusing on student interaction with the organisation's e-learning environment, WebCT.

The stakeholders did not pose any specific questions which the conceptual framework constructed from the information listed above were meant to address. A decision was made to rather construct a conceptual framework that is based solely on the types of information that exist within the six organisational information domains, and on the interpretation of the persons that perform the conceptual framework implementation.

Analysis

Examination of information sources pertaining to stakeholder-specified information requirements

The examination of the information found in each of the domains listed in the previous section was done in a less formal manner than is prescribed by the conceptual framework implementation methodology. Meetings were held with each of the secondary stakeholders mentioned in the previous section, with the focus being only to determine the different types of information found in each particular domain. The reason why a more in-depth analysis of the information was not performed, is because of the fact that the implementation was only seen as a pilot study through which the organisation was to acquire a new view on its own information. For security reasons it was therefore preferred that the persons who perform the implementation have as little direct interaction with the organisational information infrastructure as possible until such a time that the approach has been accepted, and a decision made to formally integrate the conceptual framework approach with the organisation's information infrastructure.

Development of conceptual framework specification

From information gathered through meetings held with the secondary stakeholders of the SU IT conceptual framework implementation (discussed in the previous section), a total of sixty-five

[31] *Again, due to obvious privacy issues, information relating to the internet addresses that are visited by students when accessing the internet was omitted.*

(65) conceptual framework types were identified. These types, together with their descriptions can be seen in Table 7.4.

Table 7.4: Names and descriptions of conceptual framework types identified in SU IT conceptual framework implementation case study

Entity Name	Description
Age	Age of a SU student or a library item
Date	Any references made to dates in the SU IT environment being mapped. This include the following: • Dates of card reader transactions; • Dates of internet transactions; • Dates of library transactions; • Dates of printing transactions; and • Dates of WebCT actions.
Document Type	Types of documents, e.g. MS Word documents, MS Excel documents, PDF documents, etc. (mainly related to printing operations)
Education department	Denoting a South African education department, from which a scholar can receive a senior certificate used in applying for study at the SU
File Size	Byte size of a computer file (mainly related to printing operations)
Gender	Male/Female
Handicap	Manner of handicap experienced by a SU student (this includes wheelchair handicap)
High school	High school attended by a SU student
Language	Language spoken by a SU student, as well as language a library item was written in
Library Branch	Branch of the SU library services where items can be loaned to students
Library Database Name	Name of an on-line information database managed by the SU library, and which can be accessed by students through a web-portal
Library Dewey Classification	Number of a library item corresponding to a numbering system utilised by the SU library to categorise its items
Library Item Author	Author of a library item

Entity Name	Description
Library Item Collection	Indication of the type of SU library material (e.g. reference material)
Library Item Loan Amount	Amount of times library items have been loaned
Library Item Material	Media format of a SU library item
Library Item Publisher	Publisher of a SU library item
Library Item Status	Status of a SU library item
Library Item Subject Topic	Broad topic that includes a specific SU library item
Library Item Title	Title of a SU library item
Library Item Vendor	Vendor of a SU library item
Library Shelf Code	Shelf code where a SU library item is stored
Library Shelf Code Type	Distinguishes between different shelf code systems used by the SU library
Library Transaction Location	Indicates whether a SU library transaction was conducted at the counter in the library, or through a web portal
Library Transaction Patron Type	Type of person performing a certain transaction at the SU library
Library Transaction Type	Type of transaction performed at the SU library
Marital status	Marital status of a SU student
Nationality	Nationality of a SU student
Network traffic	Network usage experienced by SU IT systems (mainly indicates internet usage and usage of SU library on-line information sources in bytes per hour)
Number of pages	Number of pages used through printing and photo-copying activities
Page Cost	Page cost of printing and photocopying activities
Person	User of the SU IT system, accessing the system through a student number or computer username. This includes the usage of card readers, internet access, printing and library services and WebCT. It further provides access to academic and demographic information stored by the SU on its students.
Population group	Population group of a SU student
Postal code	Postal code of SU student

Entity Name	Description
Time band	Any references to time made in the SU IT system, grouped in time bands. This include the following: • Times of card reader transactions; • Times of internet transactions; • Times of library transactions; • Times of printing transactions; and • Times of WebCT actions.
Town	Home town of SU students, as well as cities where certain publishers of SU library items are situated
US Campus	Indicates one of the four main SU campuses: Stellenbosch Main Campus, Bellville Campus (SU Business School), Tygerberg Campus (Health Sciences) and the Saldanha Military Academy.
US Card Reader	Name of a US card reader
US Card Reader Type	Indicating type of SU card reader (e.g. door access, photo copier, washing equipment, etc.)
US Department	Department of the SU
US Faculty	Faculty of the SU
US Module	Module (or subject) taken by SU student and presented by a SU department
US Photocopier	Name of a SU photocopier
US Printer	Name of a SU printer
US Programme	Name of a programme presented by the SU
US Programme Continuity	Indicates whether a programme is studied full-time or part-time
US Qualification	Type of SU qualification that can be attained by a student, e.g. degree, diploma, certificate, etc.
US Residence	Name of a SU residence. Only applicable to the usage of card readers with regards to washing equipment
US Residence Type	Indicates a type of SU residence, e.g. hostel, student house, etc.
US Vehicle License Plate	License plate number of a SU vehicle in the SU motor pool
US Washing Equipment	Type of SU washing equipment that can be used by SU students, e.g. washing machine, tumble drier, etc.

Entity Name	Description
US Workstation/PC	Name of a SU workstation or PC used for printing of documents and internet access
WebCT Action	Action performed by a user in the SU WebCT environment
WebCT Dwell time	Time spent by a user in the SU WebCT environment between actions
WebCT Event Role	Role of a user when performing an action in the WebCT environment
WebCT Learning Context	Learning context in the WebCT environment
WebCT Learning Context Administrative Period	Period for which a specific WebCT learning context is accessible
WebCT Learning Context Role	The role of a user in a specific learning context
WebCT Learning Context Type	Indicates the level where a learning context can be found in the WebCT environment
WebCT Number	Any references to numbers in the WebCT environment, usually used in reporting the contents of WebCT templates and WebCT learning contexts
WebCT Organisation	Name of organisation using specific WebCT learning contexts
WebCT Organisation Type	Value distinguishing general categories of the organisation
WebCT Organisation Unit	Name of sponsoring/administering unit within the organisation
WebCT Size	Any references to sizes in the WebCT environment, usually used in reporting file sizes
WebCT Template	Component of WebCT learning context that can be utilised by users

A total of one thousand and twenty-two (1,022) conceptual framework relations were defined between the types (which were still less than half of the $\frac{1}{2}(n^2+n) = \frac{1}{2}(65^2+65) = 2145$ potential relations predicted by the conceptual framework implementation methodology), again by employing the matrix approach mentioned in section B – Analysis phase. In the case of the SU IT conceptual framework case study this approach proved to be extremely time-consuming as there was no information on the internal structure of the organisational information sources to guide this activity, which meant that all the possible combinations of type-pairs that could be identified had to be evaluated from a purely logical viewpoint and for each case the decision made if a relation exists between a particular pair of types. Because of time constraints, and the fact that the definition of relations were so time-consuming, not all of the relations that might possibly exist within the information constituting the SU IT conceptual framework were

evaluated, which means that the actual number of relations could probably be much closer to the maximum of 2145 relations than previously indicated.

During the time that relations were being defined for the SU IT conceptual framework, it also became apparent that the defining of relations between the types of a conceptual framework can be performed in a much more consistent manner if the types and relations that exist within each of the organisation domains are defined first in terms of each domain alone (i.e. in isolation from the other domains, resulting in the specification of several "mini"-conceptual frameworks), followed by consolidating these separate parts into a singular whole. During this consolidation, types essentially representing the same information are collapsed and grouped together in a type which best describe the collective property of these types (e.g. types "**Student"**, **"Library Patron", "Computer user",** etc. can all be grouped under type **"Person"** or **"Student"**). This discovery eventually resulted in an adjustment being made to the activities required in the Analysis phase of the conceptual framework implementation methodology, through the adding of the activity named *Refine type-relation structure* (see section B – Analysis phase).

In an attempt to enrich the variety of organisational information that the SU IT conceptual framework could present, three different kinds of relations were defined between the various conceptual framework types listed in Table 7.4 namely *basic*, *inferred* and *aggregation* relations (which all still conform to the conceptual framework definition of relations, but were defined in different manners for the purpose of this particular implementation):

- Basic relations were the relations that are inherent to the various organisational domains being mapped and which already (at least conceptually) existed before the mapping was done. Basic relations were defined for all the organisational domains that formed part of the SU IT conceptual framework implementation case study.
- Inferred relations were relations identified through the overlapping of domains. Because an immense number of inferred relations can be developed from overlapping domains, only the relations that made the most sense at the time were included. Other relations could have been added to the conceptual framework at a later stage if so requested by the stakeholders during the conceptual framework verification. Inferred relations were defined as indicated below:
    - *Card reader transaction information*
        - ←→ Internet usage cost information
        - ←→ Library transaction and book information
        - ←→ Printing transaction information
        - ←→ WebCT usage information

- *Demographic and study programme information*
  - ←→ Card reader transaction information
  - ←→ Internet usage cost information
  - ←→ Library transaction and item information
  - ←→ Printing transaction information
  - ←→ WebCT usage information
- *Internet usage cost information*
  - ← → Library transaction and book information
  - ← → WebCT usage information
- *Library transaction and book information*
  - ← → Printing transaction information
  - ← → WebCT usage information
- *Printing transaction information*
  - ← → WebCT usage information

- Aggregate relations were relations providing links to aggregate values that were included in the conceptual framework in order to provide more useful information to users utilising information in the conceptual framework for decision making. Aggregate values such as averages, totals, minimums, and maximums are just some of the values that were provided and obviously focus more on entities of a numerical nature. Aggregate relations were defined for both basic and inferred relation scenarios.

A visual representation of the final type-relation structure of the SU IT conceptual framework is shown in Figure 7.16.

Upon completion of the conceptual framework specification (i.e. the type-relation structure), an SU IT conceptual framework 'Planning and Analysis' document (see Appendix D) was compiled, which provided the following information that was gathered and developed as part of the conceptual framework implementation process up until this point:

- A visual depiction of the conceptual framework implementation methodology (at that stage);
- The rationale behind the implementation;
- Contact information of the implementation stakeholders;

Figure 7.16: Refined type-relation structure of the SU IT conceptual framework implementation case study[32]

- The organisational information domain on which the implementation was focusing;
- The conceptual framework types identified within the information domain;
- The conceptual framework relations defined between the identified types;
- Decisions which had to be made before the implementation could continue; and
- The next steps of the implementation.

This document was presented to the principal stakeholder and discussed. The principal stakeholder verified the conceptual framework as specified in the document, but made the decision to reduce the scope of the implementation due to time constraints on the project. Once the implementation was completed, stakeholder impressions of this smaller conceptual framework would then determine whether information from the remaining domains should be added to the existing conceptual framework. A decision was made to select a group of only 1,000 students (as opposed to the entire student body) which to study, and focus on information relating to their presence in and interaction with only the following three domains:

- Card reader transaction information;
- Demographic and study programme information; and

[32] *Note that the purpose of this image is not to convey the detail of the SU IT conceptual framework in a human readable format, but rather to provide an indication of its magnitude and complexity.*

- Library transaction information.

The principal stakeholder furthermore proposed to have each of the domains in question prepare a set of data (which is representative of that domain) from which the conceptual framework could be created. This was done for two reasons: firstly, because the organisation had to encrypt sensitive student-specific information that form part of the domain information (like student numbers) in order to protect student privacy; and secondly, because it negated the need for direct interaction by external parties with the organisational information infrastructure, which was preferred from a security viewpoint. The delivery of this packaged information initiated the Development phase of the SU IT conceptual framework implementation.

Development

Fulfil conceptual framework hardware requirements

Due to the fact that the SU IT conceptual framework implementation was seen as a pilot study, hardware requirements were not fulfilled by the organisation. It was decided to construct the conceptual framework off-site at the Indutech premises using Indutech hardware and software, and only once the implementation stakeholders of Stellenbosch University have seen the result of the study and indicated their willingness to perform a proper conceptual framework implementation would the ensuing SU IT conceptual framework be integrated with the organisation's information infrastructure.

Create and populate conceptual framework

Similar to the Indutech conceptual framework implementation, the data from the various Stellenbosch University information domains were transferred to a MS SQL Server database implementation at Indutech from where the mapping layer extension of the Organon conceptual framework software tool could obtain the data and transform it into a conceptual framework. No real complications were experienced in creating a conceptual framework from the data residing in the MS SQL Server database. The only difficulty that was encountered was that the Organon software was unable to manage the masses of information that were generated through the creation of the SU IT conceptual framework. A decision was therefore made to further reduce the number of students around which the conceptual framework was to be constructed (to 424), as well as to reduce the timeframe of the conceptual framework to one month (August 2007). This resulted in a conceptual framework that fortunately was still deemed sufficiently complex to provide stakeholders with a clear view on the potential of the conceptual framework approach in their organisation. Once the creation and population of the SU IT conceptual framework (which will be discussed in more detail in the Final Result section) was

completed, the rationality of the conceptual framework was investigated and verified successfully.

Verify information represented by conceptual framework

Verification of the SU IT conceptual framework took place in two stages. First, the conceptual framework was demonstrated and discussed through one-on-one contact sessions with each of the secondary stakeholders who supplied information for inclusion in the SU IT conceptual framework. During these sessions emphasis were placed specifically on the domain that each person are responsible for, and whether the information from that domain is represented correctly by the SU IT conceptual framework. Other aspects that were also discussed in a more general fashion, were whether the particular secondary stakeholder could see the benefit of the conceptual framework information modelling approach in their own environment, and what other information within the organisational information environment could be added to the conceptual framework to make it even more useful.

The second stage of the SU IT conceptual framework verification was done as a demonstration of the conceptual framework to a group of persons consisting of the principal stake holder of the implementation, one of the primary stake holders of the implementation and two researchers from the GCC. During this session, the process that was followed to construct the conceptual framework was discussed, and the participants from the organisation asked to verify the correctness of the information as represented by the conceptual framework. Furthermore the benefits and shortcomings of the conceptual framework were discussed, as well as the manners in which the conceptual framework could be improved.

Unlike the first case study, the participants in the conceptual framework verification discussions were not each provided with a personal copy of the conceptual framework that was developed, and therefore could not verify the information in the SU IT conceptual framework based on first-hand experimentation and interaction. It was deemed sufficient however, to have the participants verify the information represented by the conceptual framework through their comments and impressions during the verification discussions. These comments and impressions were captured in questionnaires (see Appendix E) that each of the participants were asked to complete subsequent to the discussions.[33] Through specifically directed questions in the questionnaire, it was established that the information represented by the conceptual framework was correct, which therefore verified the SU IT conceptual framework.

[33] *The results of which will be discussed (together with other comments and impressions gathered during the verification discussion) in more detail in the Observations section.*

Implementation and Maintenance

Similar to the Indutech conceptual framework implementation case study, a decision was made not to continue rolling out the conceptual framework approach to the rest of the organisation following the Development phase. This decision was largely based on the fact that, for the conceptual framework approach to have actual value to the organisation, certain functionalities are required from the conceptual framework software tool which Organon is currently unable to provide. Furthermore, because the SU IT conceptual framework was constructed in a very general manner, and not according to some definite information needs, the conceptual framework did not provide definite benefits to the stake holders, which created an obvious barrier to acceptance and the continuation of the implementation. The case study concerning the implementation of a conceptual framework at the information department of Stellenbosch University was thereby concluded with the completion of the Development phase.

Again, similar to the Indutech conceptual framework implementation, the SU IT conceptual framework implementation was deemed a success from the viewpoint of the persons that performed the implementation. The methodology was found to function very effectively, and addressed relevant issues during the different phases of the conceptual framework implementation.

• *Final Result*

Through the SU IT conceptual framework implementation case study a conceptual framework was constructed containing more than 267,000 facts, which provides a view on the students of Stellenbosch University, their academic and demographic profiles, as well as their use of the university's library and card reader systems. Additional information that the conceptual framework could provide include the profiles of university library items, university photocopiers and university washing equipment. The rest of this section discusses a number of Organon[34] screenshots (which were improved since the last case study), which give an indication of the information contained in the SU IT conceptual framework:

[34] *A new version of Organon was used in the SU IT conceptual framework implementation case study, which is why the screenshots will look different from those describing the conceptual framework constructed during the first (i.e. Indutech) case study.*

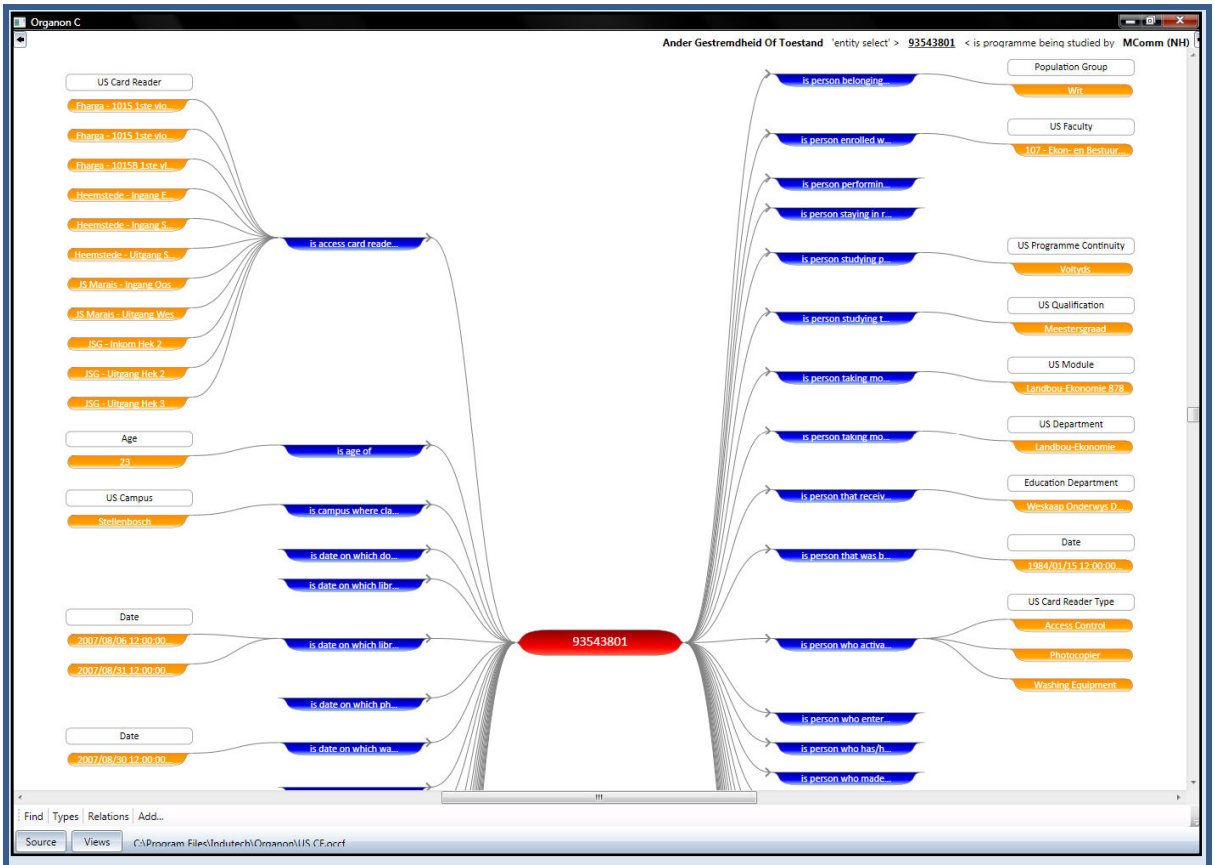


Figure 7.17: Screenshot 1 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.17 shows a **"Person"** entity named "93543801", and some of the relations that pertain to this entity. It furthermore indicates the following information about this person:

- The university card readers this person utilises to gain access to various buildings on the university campus
- The age of this person
- The campus where this person attends class
- Dates on which library transactions were performed by this person
- Dates on which washing equipment was used by this person
- The population group this person belongs to
- The faculty where this person is enrolled
- The manner in which this person studies his/her course, which is full-time
- The type of qualification this person is studying towards, which is a master's degree
- The module(s) this person is taking as part of his/her course
- The department where this person is taking his/her module(s)
- The education department where this person received his/her senior certificate
- The person's date of birth
- The types of university card readers this person uses

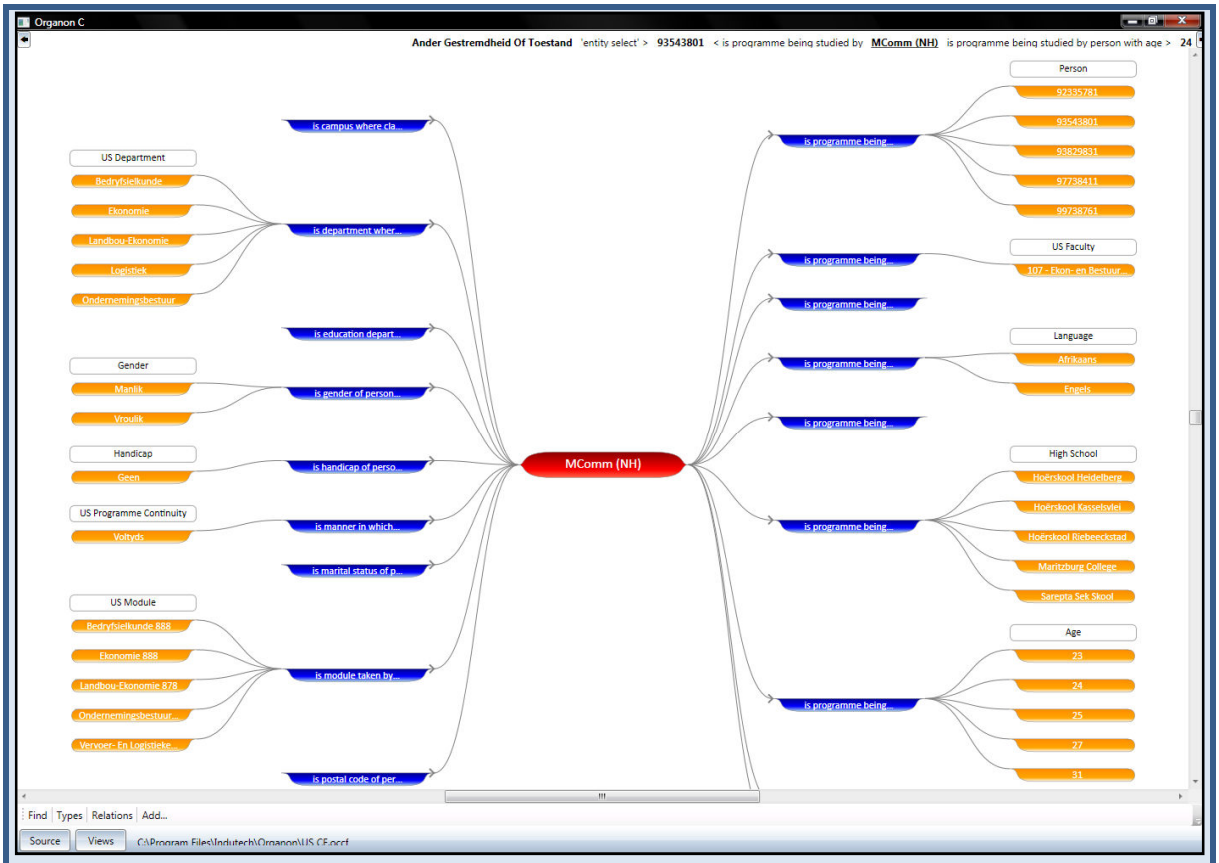


Figure 7.18: Screenshot 2 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.18 shows a **"US Programme"** entity named "MComm (NH)" (which is the course studied by person "93543801" of the previous screenshot), and some of the relations that pertain to this entity. It furthermore indicates the following information about this programme:

- The university departments where persons studying this course take their modules
- The genders of persons studying this course
- The manner in which persons studying this course aim to achieve their qualification, which is full-time
- The modules taken by persons studying this course
- The handicaps of persons studying this course
- The persons studying this course (which obviously include person "93543801")
- The faculty that at which persons are enrolled in order to study this particular course
- The home languages of persons studying this course
- The high schools attended by persons studying this course
- The current ages of persons studying this course

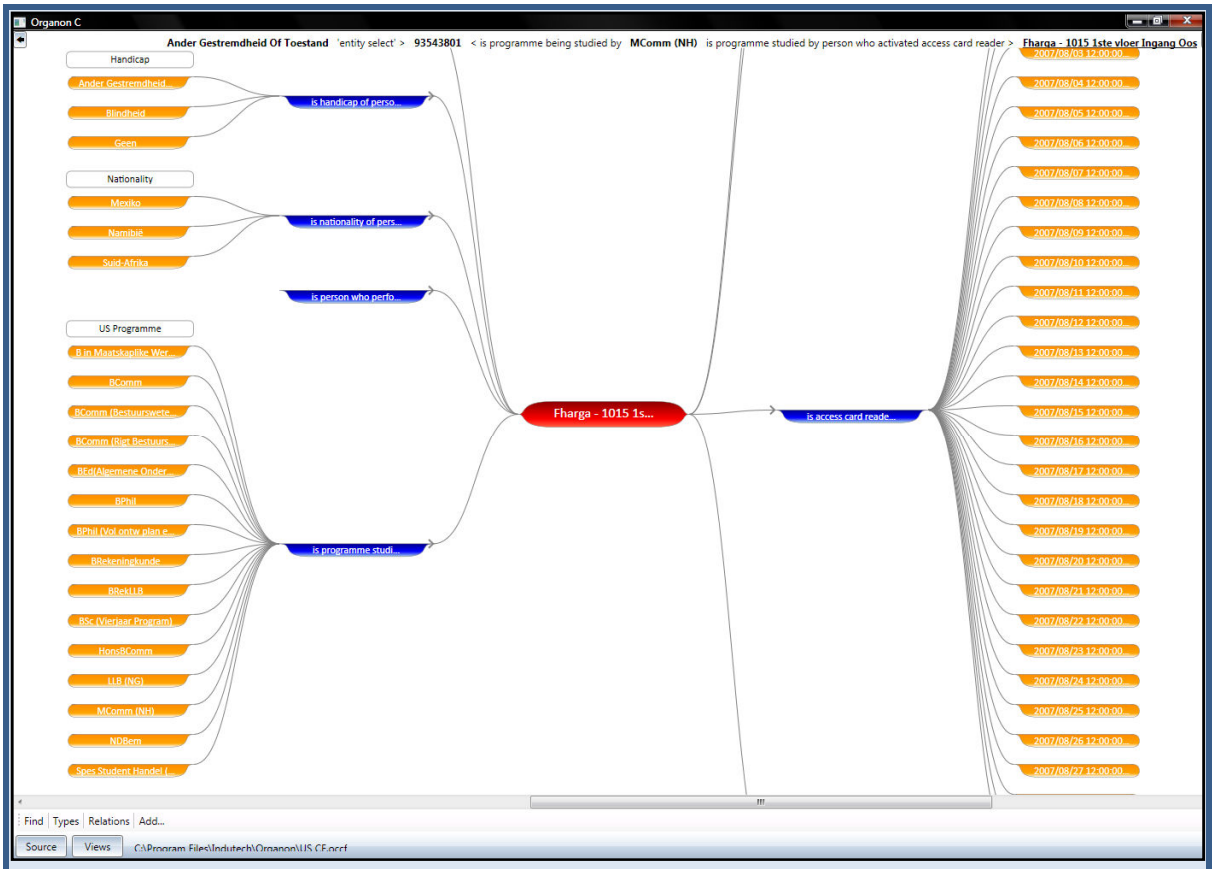


Figure 7.19: Screenshot 3 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.19 shows a **"US Card Reader"** entity named "Fharga 1015 1ste vloer Ingang Oos" (or Fharga 1015 1st floor Entrance East, which is one of the access card readers utilised by persons studying SU Programme "MComm (NH)", discussed in the previous screenshot), together with some of the relations that pertain to this entity. The following information on this particular card reader is also provided:

- The handicaps of persons utilising this card reader
- The nationalities of persons utilising this card reader
- The courses studied by persons utilising this card reader
- The dates on which this card reader was utilised

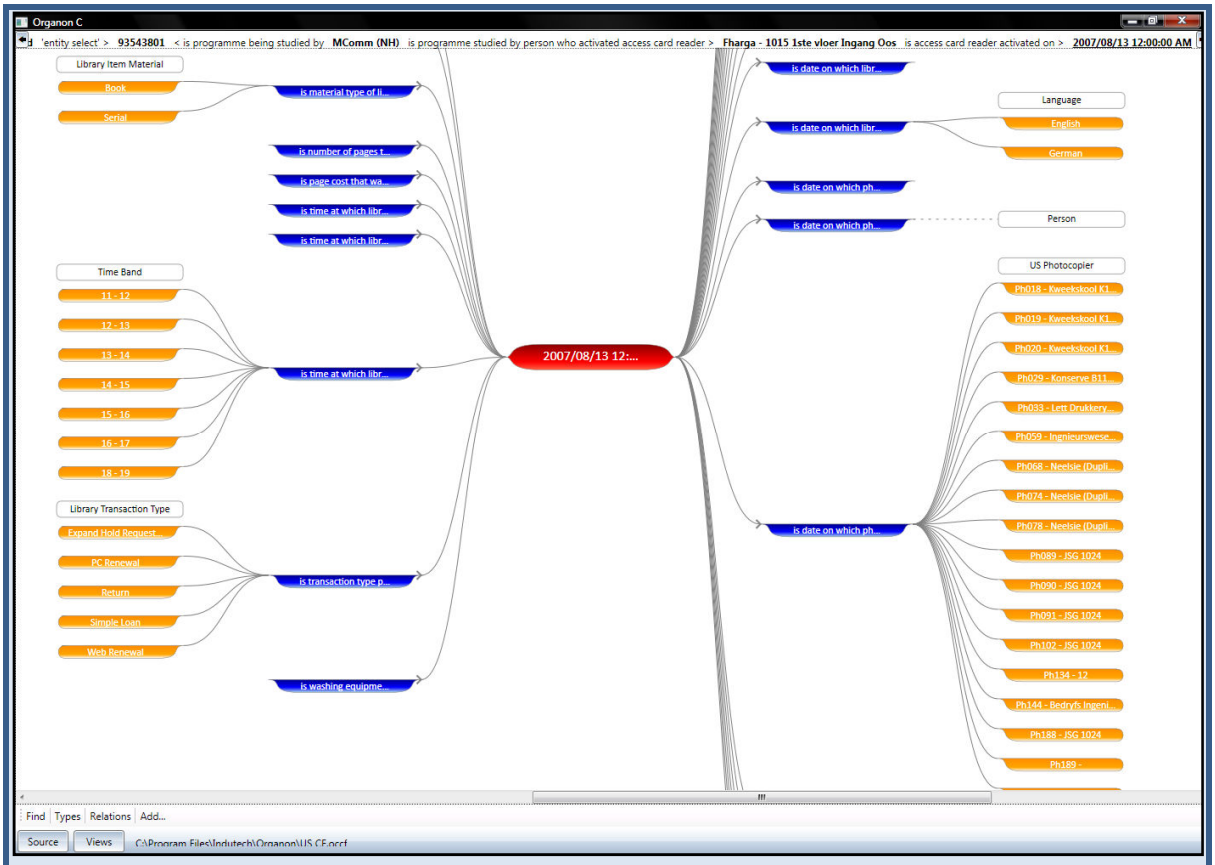


Figure 7.20: Screenshot 4 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.20 shows a **"Date"** entity named "2007/08/13 12:00:00 AM", which provides a view of activities that occurred throughout the information domains in question on a specific date. Some of the relations that pertain to this entity are displayed, together with the following information:

- The types of library material on which library transactions were performed on this date
- The hours in which library transactions were performed on this date
- The types of library transactions that were performed on this date
- The languages of the library items on which transactions were performed on this date
- The photocopiers that were used on this date

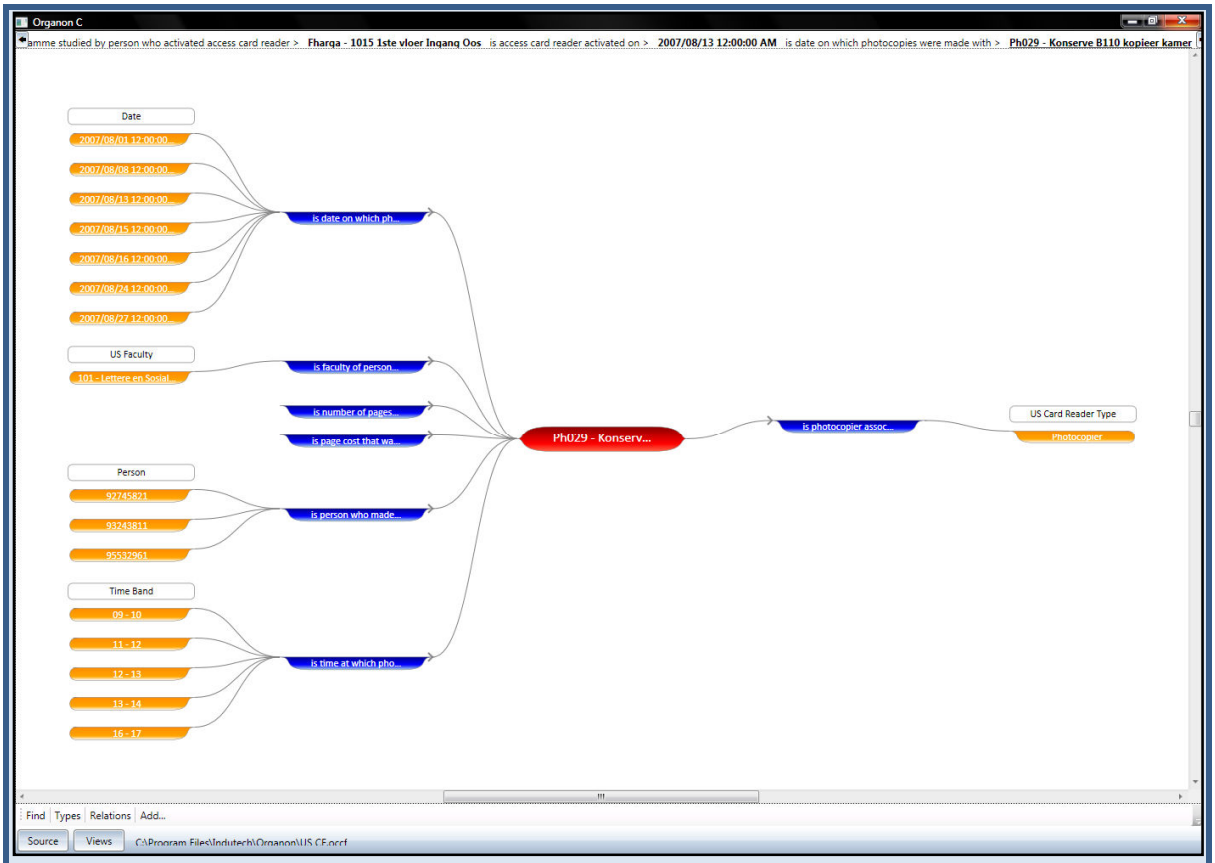


Figure 7.21: Screenshot 5 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.21 shows a **"US Photocopier"** entity named "Ph029 Konserv B110 kopieer kamer" (or Ph029 Conservatory B110 copying room, which is one of the photocopying machines used on the date 2007/08/13, discussed in the previous screenshot), together with some of the relations that pertain to this entity. The following information relating to this particular photocopier is also provided:

- The dates on which this photocopying machine was used
- The faculty with which the persons who used the photocopying machine are enrolled
- The persons who used the photocopying machine
- The times at which the photocopying machine was used
- The type of card reader associated with the photocopying machine

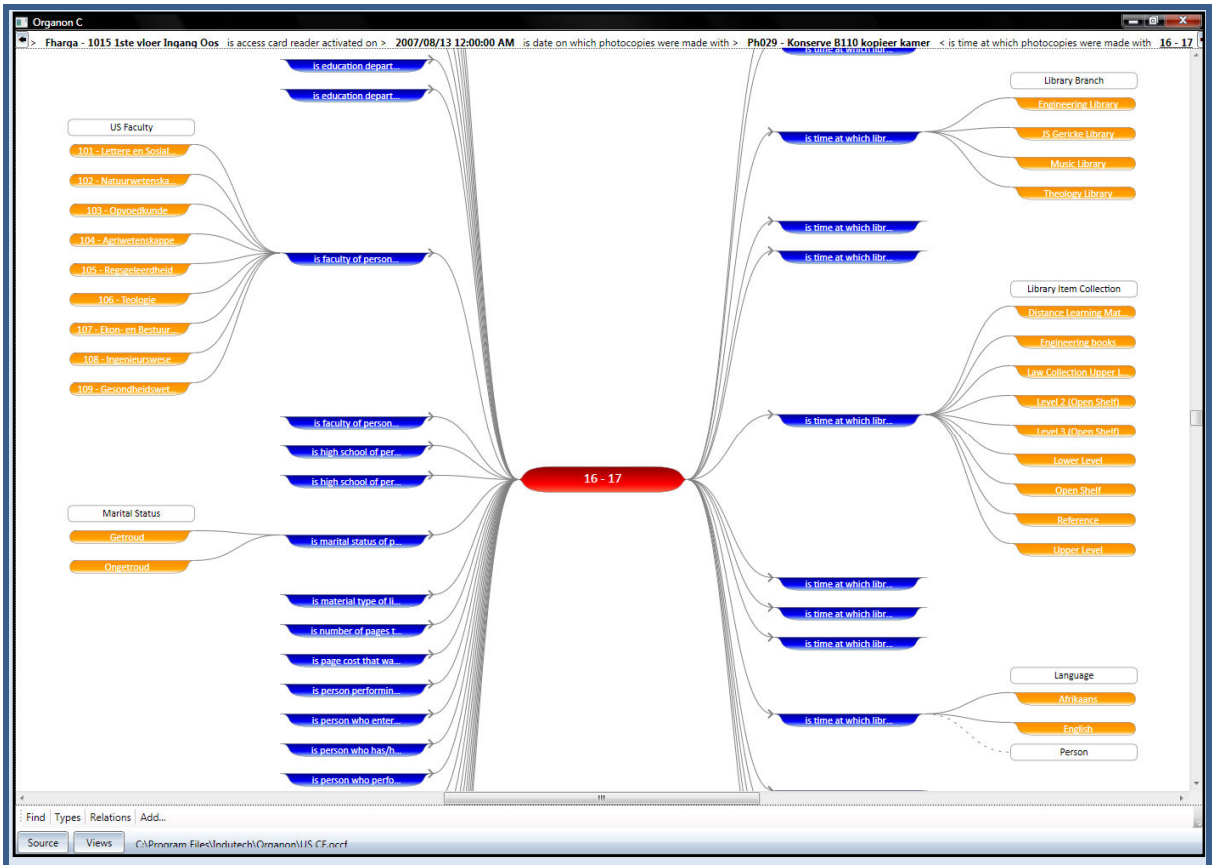


Figure 7.22: Screenshot 6 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.22 shows a **"Time band"** entity "16 - 17" (which is one of the time bands in which photocopier "Ph029 Konserv B110 kopieer kamer" is used, discussed in the previous screenshot), together with some of the relations pertaining to this entity. The following information relating to this particular hour in a day is also provided:

- The faculties with whom persons are enrolled who activated access card readers in this particular time band
- The marital statuses of persons who performed library transactions in this particular time band
- The library branches at which library transactions were performed in this particular time band
- The collections within the university library system to which the library items belong on which library transactions were performed in this particular time band
- The languages of the library items on which library transactions were performed in this particular time band

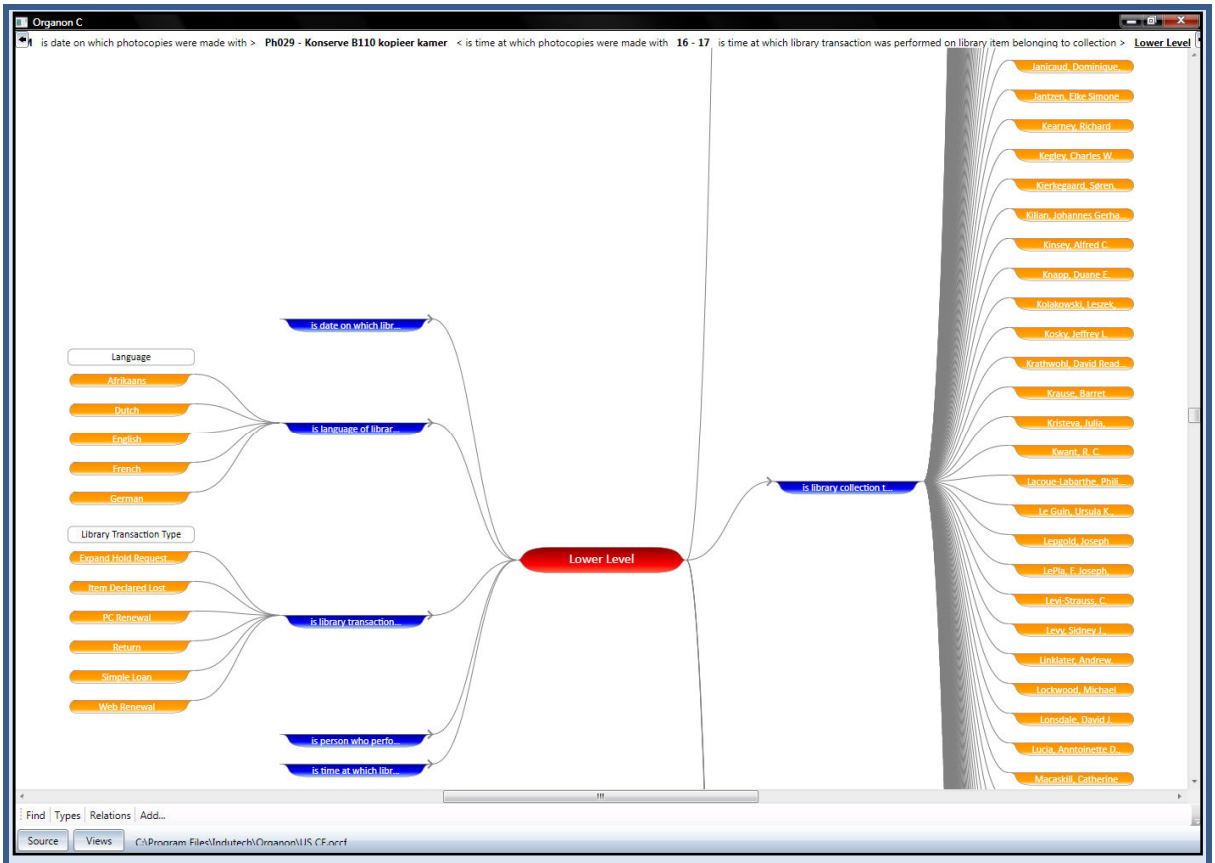


Figure 7.23: Screenshot 7 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.23 shows a **"Library Item Collection"** entity "Lower Level" (which is one of the library collections of library items that have had library transactions performed on them in the 16 – 17 time band, discussed in the previous screenshot), as well as some of the relations pertaining to this entity. The following information relating to this library item collection is also provided:

- The languages of the library items belonging to this library item collection
- The types of library transaction performed on items from this library item collection
- The authors of the library items that belong to this library item collection

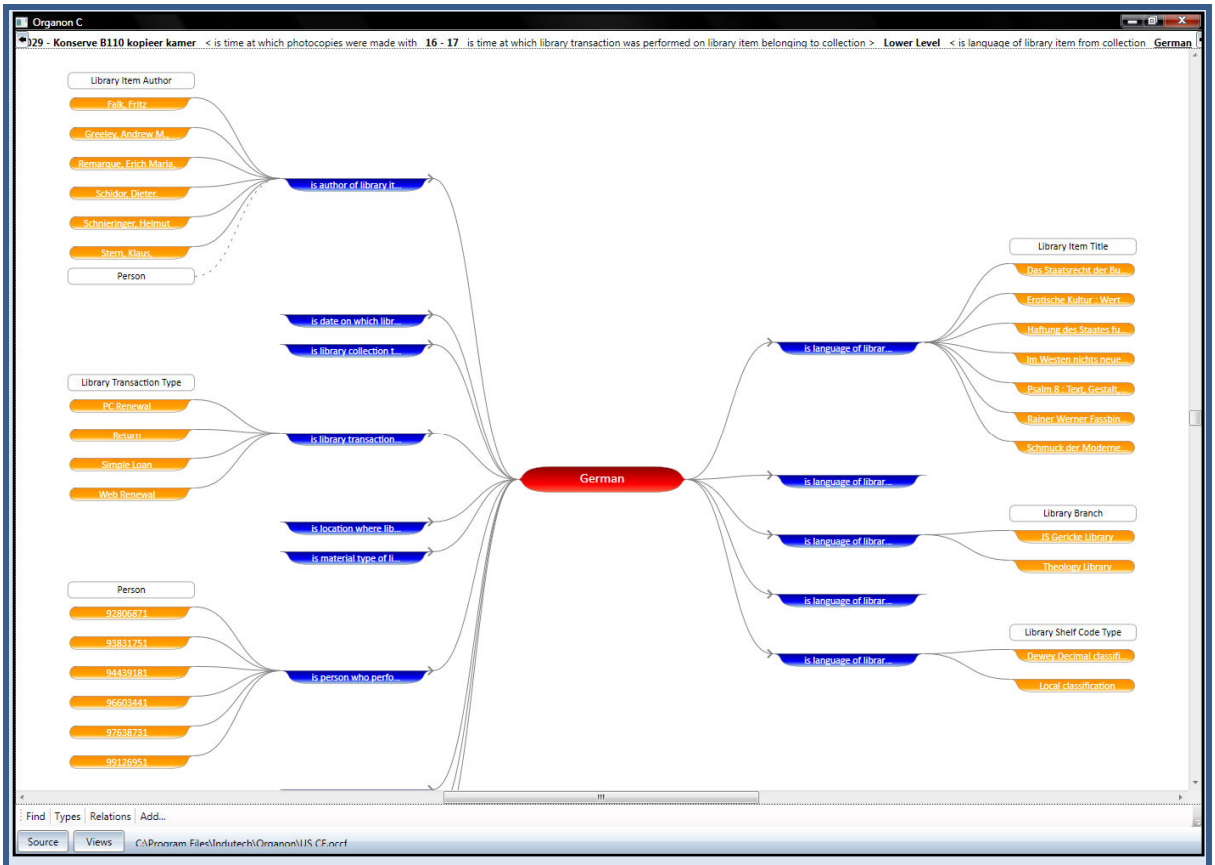


Figure 7.24: Screenshot 8 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.24 shows a **"Language"** entity "German" (which can denote either the home language of a person, or the language a library item was written in, and is one of the languages of the library items in the "Lower Level" library collection, discussed in the previous screenshot), as well as some of the relations pertaining to this entity. The following information relating to this language is also provided:

- The names of authors of library items who was written in this specific language
- The types of transactions performed on library items written in this specific language
- The persons who have performed library transactions on library items written in this specific language
- The titles of library items written in this specific language
- The library branches where library transactions have been performed on library items written in this language
- The type of shelf codes used in the shelving of library items written in this language

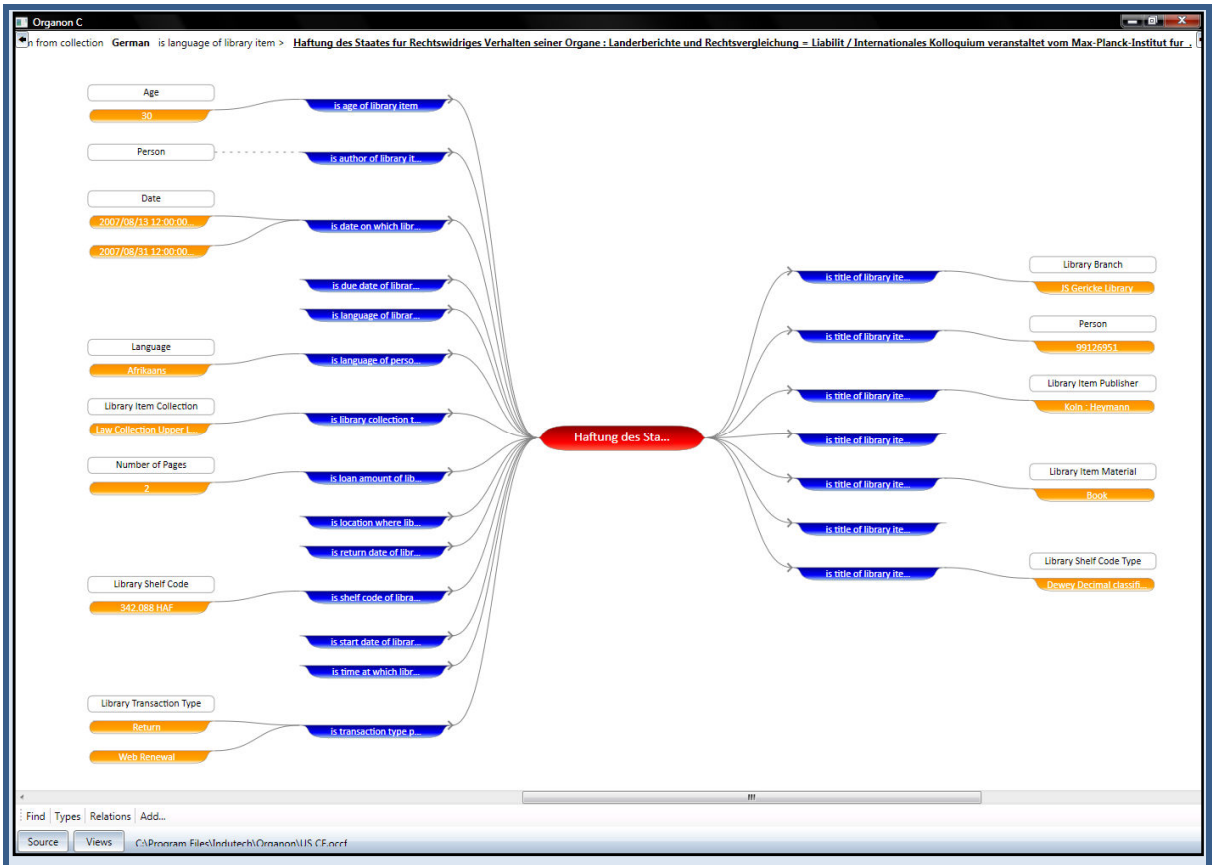


Figure 7.25: Screenshot 9 of Organon showing information of the SU IT conceptual framework implementation case study

Figure 7.25 shows a **"Library Item Title"** entity "Haftung des Staates fur Rechtswidriges Verhalter seiner Organe: Landerberichte und Reghtsvergleichung = Liabilit" (which represents the item itself, and is one of the library items written in the "German" language, discussed in the previous screenshot), as well as some of the relations pertaining to this entity. The following information relating to this library item is also provided:

- The age of this particular library item
- The dates on which library transactions were performed on this particular library item
- The languages of the persons who have performed library transactions on this particular library item
- The library collection to which this particular library item belongs
- The loan amount of this particular library item
- The library shelf code of this particular library item
- The types of library transactions that were performed on this particular library item
- The library branch where this particular library item can be found
- The persons who have performed library transactions on this particular library item
- The publisher of this particular library item
- The material type of this particular library item
- The shelf code type used in the shelving of this particular library item

- ***Observations***

This section will discuss the impressions and observations of the stakeholders and other participants that originated during the verification of the SU IT conceptual framework. Similar to the first case study, the discussion is structured around the main focus areas of the questionnaire that the participants had completed.

Mandatory questions

These questions (that formed part of the questionnaire that the participants in the conceptual framework verification had to complete) were structured in such a manner as to obtain very definite answers from the participants concerning key issues of the SU IT conceptual framework implementation. A participant could answer 'Yes', 'No' or 'Other' to each of these questions, and was provided with the option to comment on his/her answer, should it prove pertinent. The questions and their responses are discussed below:

- *Question 1: Does the conceptual framework that was presented provide an accurate view on the information contained in the implementation domain?*
  Response: 75% of participants[35] answered "Yes" to this question. The remaining 25% percent answered "Other", stating that the conceptual framework seemed accurate, but that it was difficult to assess the accuracy based on the representation of the information by Organon alone. The general impression was however that the conceptual framework created through application of the conceptual framework implementation methodology discussed in this thesis represented the organisation's information in an accurate and truthful manner.

- *Question 2: Do the types and relations that were defined provide an intuitive view on the information contained in the implementation domain?*
  Response: Again 75% of participants answered "Yes" to this question, although it was felt that the intuitiveness of the information could be improved by providing an indication of the number of entities that are associated with each relation pertaining to a specific entity. Other than that, 12.5% of participants answered "Other", citing uncertainty, and 12.5% answered "No", indicating that the types and relations are not in and of themselves intuitive.

[35] *From a total of eight, being made up of the principal stake holder, one primary stakeholder, four secondary stakeholders and two external participants (from the GCC) to the conceptual framework verification activities.*

- *Question 3: Is there benefit in viewing the organisation's information as a navigable conceptual framework?*

  Response: In this case only 62.5% of participants answered "Yes" (with some seeing the potential of conceptual frameworks for performing marketing research in their environments), and 37.5% of participants answered "Other". It was generally felt by the participants who answered "Other" that potential exists, but that benefit could be improved through the adding of functionalities to Organon like user-driven filtering, aggregation functionalities, pattern recognition and the ability to pose composite questions (and therefore view complex facts).

- *Question 4: Could the utilisation of a conceptual framework improve the efficiency of decision-making in the organisation?*

  Response: 50% of participants answered "Yes" to this question, and 50% answered "Other". Again, most of the comments focused on the need for additional functionality to the software, aimed at presenting the information in a more palatable manner (e.g. highlighting relevant information, whilst hiding irrelevant information). One comment noted the uniqueness of the approach in presenting "unknown" information to a specific user, providing him/her with better insight into the context surrounding a problem, and which may ultimately lead to improved decision making.

- *Question 5: Did the conceptual framework methodology address all the major issues required for a successful conceptual framework implementation?*

  Response: For this question 37.5% of participants answered "Yes", 25% of participants answered "No" and 37.5% of participants answered "Other". The reason for the large percentage of participants not answering "Yes" (62.5%) is because they were largely uninvolved with the actual execution of the implementation, and could not fairly comment on the issues encountered during implementation. It was furthermore realised that the fact that no clear requirements were defined at the start of the implementation resulted in a conceptual framework being created which contained a massive amount of information without any focus, which reduced the usefulness of the approach.

Optional Questions

Further optional questions were included in the questionnaire that were provided to participants of the conceptual framework verification activities, and were aimed at capturing more comments and insights surrounding the implementation and usefulness of the SU IT conceptual framework. These comments will forthwith be discussed:

Conceptual framework implementation methodology

Again it was highlighted that the decision to omit the initial information needs identification activity during the Planning phase of the conceptual framework methodology resulted in the approach lacking some focus, and therefore lacking usefulness for its users. It therefore gave a definite indication of the importance of this activity, as well as that the focus should also be defined with the intended users in mind.

Benefits of the SU IT conceptual framework

It was felt that the graphical representation of data can greatly improve the analytical capabilities of a user, especially if more business intelligence functionality can be present in the conceptual framework software tool used to visualise the conceptual framework. Furthermore, it has been noted that the conceptual framework can become very useful in specifically the university's library environment for tracking the behaviour of library patrons (e.g. their movements, active hours during the day, types of library material used by students from different subject disciplines, etc.) in order to plan the layout of the library to best service these patrons.

Shortcomings of the SU IT conceptual framework

No real shortcomings of the SU IT conceptual framework were identified, except for the fact that the absence of filtering and aggregation mechanisms makes it difficult for a user to interact with the masses information contained by the conceptual framework. This is however a shortcoming of the conceptual framework software tool, and not of the conceptual framework itself. The lack of focus in the SU IT conceptual framework that were experienced by the conceptual framework verification participants can obviously also be seen as a shortcoming.

Growing the SU IT conceptual framework

It was mentioned that adding administrative data of the university to the data currently represented by the SU IT conceptual framework, would provide a link between two traditional disparate information domains within the university (that is, between administrative and academic data), and would provide decision makers with a much more complete view to the university IT environment, and the manner in which persons interact with it. It was furthermore also noted that by providing additional information on the entities and relations defined in the original three information domains (like frequencies, totals, percentages, the significance of relations, etc.) an increase would be experienced in the value of the information represented by the conceptual framework.

Other observations

In general the conceptual framework verification participants (and especially the implementation stakeholders) found the conceptual framework approach to view organisational information quite useful, and its potential of connecting real-world transaction data (as a proxy for behaviour) to other structured data (such as surveys, results, and demographics) in the field of student tracking rather interesting. The overall feeling was that, with added reporting and business intelligence functionality, the conceptual framework approach could prove to be very useful in decision making within the organisation, as it succeeds in bringing (often) hidden information to light, thereby providing a much clearer view on the context of a certain matter, which should increase the overall quality of decision-making. Except for the fact that the final developed conceptual framework lacked focus because of the initial needs identification activity not having been performed, the Stellenbosch University: Information Technology conceptual framework implementation was deemed a success.

• *Conclusion – Stellenbosch University: Information Technology case study*

The SU IT conceptual framework implementation was seen as a success, and through the case study a number of conclusions could be made:

- Similar to the Indutech conceptual framework implementation case study, it was seen that by employing the conceptual framework implementation methodology discussed in this thesis, a conceptual framework of organisational information can successfully be created and implemented in the organisation.
- Organisational participants in the implementation could attest to the usefulness of the conceptual framework approach in representing organisational information, especially with regards to locating specific information and improving the quality of organisational decision making.
- It is imperative that the initial information needs identification activity is performed at the start of a conceptual framework implementation (during the Planning phase). If this activity is omitted, the risk exists that a conceptual framework is created that does not fully address the contexts of its intended target users, which will have a negative impact on its usefulness to them.
- Although a conceptual framework in itself can provide useful information to users within the organisation, its benefit will become much more relevant to decision makers should the

conceptual framework software tool they use possess reporting, filtering and aggregation functionality to assist them in the interrogation of the conceptual framework's information.

- The identification of conceptual framework relations in organisational information sources can become very laborious and time consuming if the number of identified conceptual framework entity types becomes very large. A possible solution to this problem may be to automate this (currently) manual activity that forms part of the Analysis phase of the conceptual framework implementation methodology.

Following this case study some changes were made to the conceptual framework implementation methodology and proposed for the Organon conceptual framework tool. The conceptual framework implementation methodology discussed in this thesis has hereby been validated from a practical viewpoint, and should be able to successfully implement a conceptual framework in any organisation. Suggested changes will be made to Organon as Indutech sees fit.

7.4 Conclusion – Methodology Validation

This chapter has conveyed information on the ways in which the conceptual framework implementation methodology introduced through this thesis has been validated. The methodology first underwent a theoretical validation, through which it was compared with similar methodologies aimed at the implementation of other information modelling approaches (typically ontologies) in organisations. It was found that the conceptual framework implementation methodology focused on the same issues as these methodologies, indicating that the reasoning behind this methodology is conceptually sound.

After this, the methodology underwent a practical validation in the form of two implementation case studies, the first at an organisation called Indutech, and the second with the Information Technology division of Stellenbosch University. In both of these cases the Implementation and Maintenance phases of the conceptual framework implementation methodology was not executed, but this did not have a definite influence on the outcome of the case studies, as the Implementation phase focused more on configuring the support functions of a conceptual framework implementation, and the Maintenance phase focused more on ensuring the sustained improvement and expansion of an organisational conceptual framework. The phases of the conceptual framework implementation methodology that were more significant in terms of the scope of this thesis and its hypothesis were performed successfully, thereby successfully validating the conceptual framework implementation methodology and providing the input required to successfully evaluate the hypothesis stated in section 5.3. Chapter 8 will conclude this thesis, and specifically discuss the hypothesis evaluation and its outcome.

Chapter 8 – Conclusion

8.1 Introduction

As indicated by the illustration below (Figure 8.1), this chapter will conclude this document by giving a summary of the thesis content, and ending with a discussion of the hypothesis test results, and some finishing remarks.

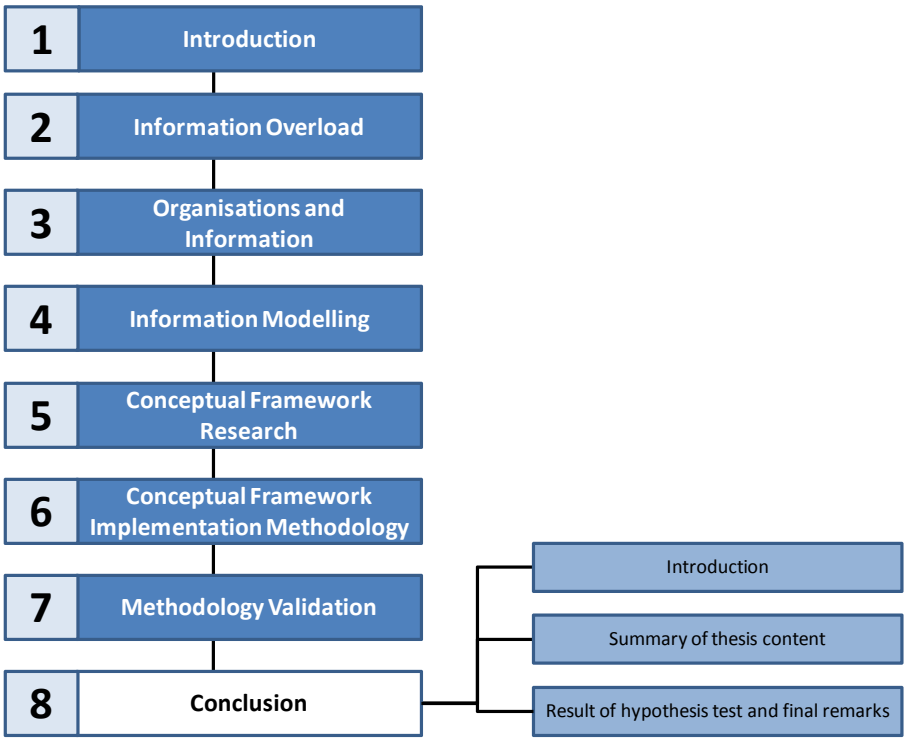
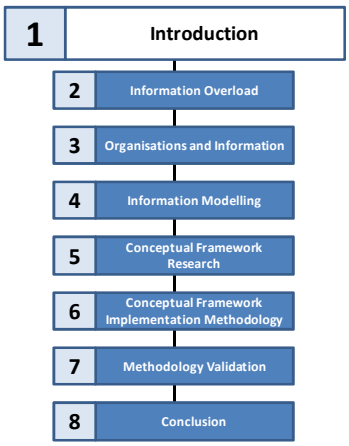


Figure 8.1: Image for navigating thesis and chapter 8

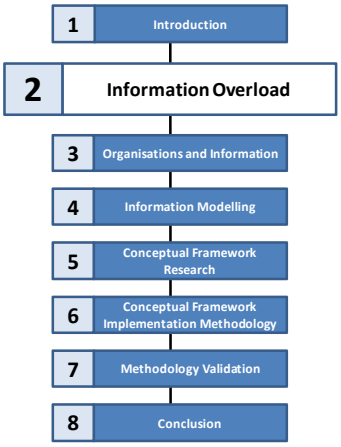
8.2 Summary of thesis content



Chapter 1 introduced this thesis by explaining how humanity's activities have shifted over the centuries from being survival-focused to being knowledge-focused, and that the masses of information that have been generated (and are still being generated) have become too much for people to process on their own.

It continued by noting that various tools and approaches have been developed to assist persons in the processing and application of information, but that these tools and approaches often fail to take into consideration the context in which a person might find him/herself, which then may result in the person's information need not being sufficiently addressed.

The chapter then introduced the rationale for the research project on which this thesis reports, which is the development of a methodology to facilitate the implementation of a context-sensitive information management approach in organisations. In order, however, to gain a better understanding of the environment and context for which this information management approach and implementation methodology was developed, a literature study was performed focusing on the concept of information overload, the manner in which organisations typically utilise and interact with information, and existing information modelling approaches, aimed at improving human-information interaction (Chapters 2 - 4).

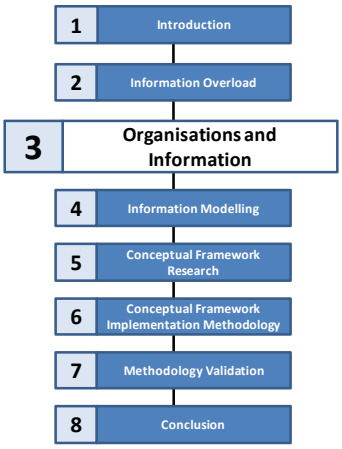


Chapter 2 focused on the topic of information overload, and started with considering the amounts of information currently available in the world, together with the rate that new information is being generated. It was indicated that these amounts of information, together with the fact that large variations are apparent in the quality of information being generated, make it difficult for a person to locate specific information in a short time, which may lead to the "condition" called information overload.

Information overload was then defined as being the provision of information in excess of the cognitive and emotional ability of an individual to process that information, and it was explained how the performance of a person will improve with an increase of information up to a certain point, from where it will again start to decrease as the person fails to assimilate amounts of information in excess of his/her capabilities. The manner in which this effect is experienced in an even more pronounced fashion in organisations, was also discussed.

Finally, the different causes and symptoms of information overload were discussed, together with different solutions proposed by experts in the information overload research field. Special mention was made of the following possible solutions to information overload, as they were very relevant to the theme of this thesis:

- Delivering the information in the most convenient way and format;
- Visualisation, compression and aggregation of the information;
- Using methods of intelligent information management for easier information prioritisation; and
- Using quality filters.

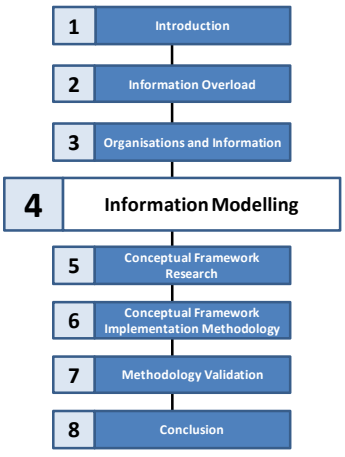


Chapter 3 discussed the manner in which organisations make use of information, and started by explaining the relevance of information to organisations. It also explained that, besides enabling an organisation to orientate itself within its environment, it enables the organisation to continuously increase its knowledge about itself. Different aspects of the manner in which organisational information utilisation takes place were discussed, particularly information as documents and transactions, information silos and integrated information, and information management tools and approaches typically employed by organisations.

The chapter furthermore proceeded to discuss in general the shortcomings of the manner in which organisations typically utilise information, but also mentioned various solutions that have been developed and implemented (with various levels of success) to address these shortcomings.

The chapter concludes by suggesting that the large variation in the tools and approaches aimed at providing effective information management capabilities to organisations may be an indication that the nature of information is still not sufficiently understood to effectively address this problem. A suggestion is made that the notion of considering information as a collection of atomic points of data which are related to each other in some manner might provide a way for persons in an organisation to interact much easier with information. This “modelling of information” was then discussed in the next chapter.

Chapter 4 focused on the concept of information modelling, and discussed the concepts behind a number of information modelling approaches currently used throughout the world, aimed at assisting persons in their interaction with information. The contexts and manners in which these approaches are typically applied were also discussed.

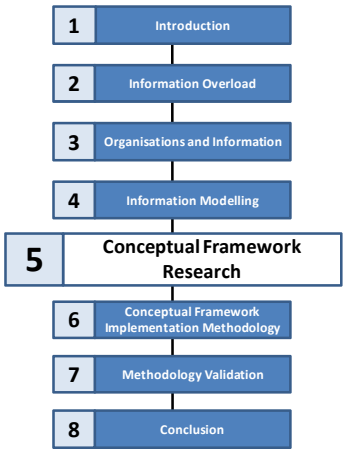


The discussion focused on the following approaches:

- Semantic networks;

- Thesauri;
- Taxonomies;
- Ontologies;
- Folksonomies;
- Concept maps; and
- Topic maps;

At the end of this chapter the approaches discussed were compared, based on the generality, flexibility and simplicity of each approach (as these features were deemed critical for an approach that provides an improved general human interaction with information). It was found that, from the approaches discussed, none could address all three features satisfactorily. The statement was made that a need still exists for an information representation technique which can be easily understood by most information users, whilst still providing enough functionality to address information modelling requirements over a broad spectrum of application. This led to the introduction of the conceptual framework approach discussed in Chapter 5.



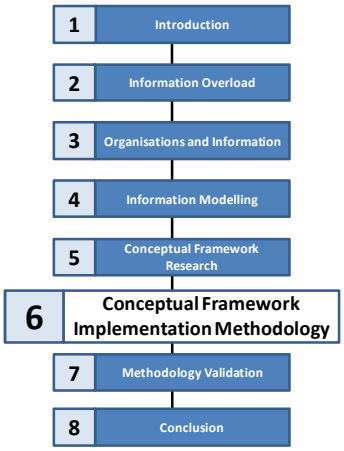
Chapter 5 introduced the concept of a conceptual framework, and explained its definition, together with the fact that the construction of a conceptual framework is based on only four concepts, namely that of entities, relations, facts and types (each of which were also explained).

The conceptual framework approach was compared with the information approaches discussed in the previous chapter, explaining the manners in which these approaches differ from the conceptual framework approach. It was further indicated that conceptual frameworks can be applied to both structured and unstructured information, and that it lends itself very well to the graphical representation of information (which is normally useful from a users' perspective).

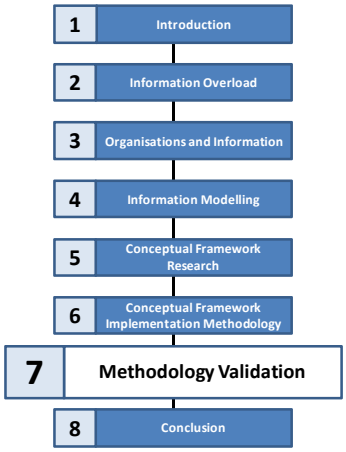
The chapter proceeded to provide more information on research into conceptual frameworks, most notably the two streams of research focusing on employing the conceptual approach in the representation of structured information and unstructured information. It further indicated that the research described in this thesis makes up part of the research stream focusing on structured information.

The chapter concludes by providing the hypothesis of this thesis, and the research method that was followed.

Chapter 6 provided the conceptual framework implementation methodology that was developed through the research on which this thesis reports. The methodology is composed of five phases, each with a number of activities that should be executed to successfully implement a conceptual framework in an organisation. The structure of the methodology is as follows:



- **A – Planning phase**
  - *Activity A1 – Identify stakeholders*
  - *Activity A2 – Initial information need identification*
- **B – Analysis phase**
  - *Activity B1 – Examination of information sources pertaining to stakeholder-specified information requirements*
  - *Activity B2 – Development of conceptual framework specification*
- **C – Development phase**
  - *Activity C1 – Fulfil conceptual framework hardware requirements*
  - *Activity C2 – Create and populate conceptual framework*
  - *Activity C3 – Verify information represented by conceptual framework*
- **D – Implementation phase**
  - *Activity D1 – Information infrastructure setup*
  - *Activity D2 – User setup*
- **E – Maintenance phase**
  - *Activity E1 – Grow conceptual framework*



Chapter 7 discussed the validation of the conceptual framework implementation methodology, discussed in the previous chapter.

Validation of the methodology was done in two ways. First, the methodology was validated from a theoretical point of view through comparison with existing implementation methodologies which facilitate the implementation of other types of information modelling approaches (like ontologies) in organisations. The aspects addressed by the conceptual framework implementation methodology was found to correspond very well with the aspects focused upon by the other approaches, and thus the conceptual framework implementation methodology was validated (from a theoretical point of view).

The second manner in which the conceptual framework was validated, was through the execution of two conceptual framework implementation case studies, through which conceptual frameworks were implemented in two actual organisations. Feedback from the stakeholders and participants of each of the

implementations affirmed that the conceptual frameworks were successfully implemented, and that the information represented by each conceptual framework gave an accurate view on the organisational information involved. Through the case studies the conceptual framework implementation methodology was therefore validated from a practical point of view, which released the methodology for wide-spread application.

Chapter 8 has provided a summary of the contents of this thesis, and will now continue to discuss the result of the hypothesis test, following with some final remarks.

8.3 Result of hypothesis test and final remarks

8.3.1 Hypothesis Test

In section 5.3 the following hypothesis (which this thesis aimed to prove or disprove) was stated:

Hypothesis: **Information represented by a conceptual framework implemented through the use of the developed conceptual framework implementation methodology, offer a truthful view on information found in an organisation.**

The only mechanism through which this hypothesis could be tested was through the conducting of actual conceptual framework implementations and assessing whether the information represented by each resultant conceptual framework mirrors the organisational information on which the implementation focused. As already discussed, such implementations were performed at a small industrial engineering organisation called Indutech (Pty) Ltd., and at the Information Technology division of Stellenbosch University (see sections 7.3.2 and 7.3.3). As part of the conceptual framework implementation methodology followed in both cases, stakeholders of each implementation were

required to verify the accuracy of their specific conceptual framework as thoroughly as possible. Because these verifications addressed exactly the same issue that the hypothesis mentioned above focuses on, the conclusions from these verifications could be directly applied in the hypothesis test.

Both sets of stake holders agreed that, based on their exposure to the conceptual framework that was developed through the implementation case study in which they took part, a conceptual framework developed by means of the conceptual framework implementation methodology discussed in this thesis offer a truthful view on organisational information. It can therefore be said that the hypothesis could not be disproven, and can therefore be assumed to be true. There are however two aspects of the implementations that should be clarified in order for this conclusion to have credibility:

- Stakeholders were not able to thoroughly analyse the information constituting each of the conceptual frameworks. This should however have no real effect on the outcome of the hypothesis test as the conceptual frameworks were constructed in a consistent manner, and the conceptual framework information that were not analysed by the stakeholders should therefore (from a conceptual point of view) be no different from the conceptual framework information that were.
- In neither of the two implementations were all the phases of the conceptual framework implementation methodology that were discussed in this thesis executed. As was mentioned in section 7.4, the Implementation and Maintenance phases of the conceptual framework implementation methodology were not executed in either of the two implementation case studies. This did however not have any impact on the outcome of the hypothesis test as these phases, though very relevant and important in the process of implementing conceptual frameworks, focus on activities that should take place *after* the organisational conceptual framework have been created and verified.

The hypothesis, which forms the main focus of this thesis, could therefore not be disproven, and can consequently be assumed to be true. It can thus be confidently stated that:

Information represented by a conceptual framework implemented through the use of the developed conceptual framework implementation methodology, offer a truthful view on information found in an organisation.

8.3.2 Final Remarks

This document has discussed a number of issues around the utilisation of information by individual persons and organisations. The search for an effective solution to efficiently interact with (especially large amounts of) information does not seem to be reaching its end, whilst the difficulties associated with information overload seem to be on the increase. It is the opinion of the author that this problem can only be sufficiently addressed by employing strategies focused around information structuring/modelling, together with the effective use of business and technical metadata.

This thesis was aimed at adding to the existing knowledge on information structuring, and on providing a methodology that can assist in the furthering of knowledge on conceptual frameworks and in improving the efficiency of organisational information utilisation. The development of the conceptual framework implementation methodology has now made it possible for researchers to create environments in which conceptual framework research can more easily be conducted, and for organisations to implement conceptual framework based tools in order to interact more efficiently with their own information. The increased interaction with conceptual frameworks will increase humanity's understanding of the value of conceptual frameworks, and could ultimately bring it closer to a solution for the global information overload problem.

Chapter 9 – References

1. *About Stellenbosch University* 2007. Retrieved May 19, 2008 from http://www.sun.ac.za/university/index.html

2. Appfluent Technology 2002, *A Study on Reporting and Business Intelligence Application Usage*. Retrieved August 25, 2006 from http://whitepapers.zdnet.co.uk/0,1000000651,260095138p,00.htm

3. ARMS 2006, *Glossary of Recordkeeping Terms*. Retrieved November 2, 2007 from http://secint33.un.org/unarms/en/unrecordsmgmt/unrecordsresources/glossaryofrecordkp.html

4. Auray, N 2007, 'Folksonomy: the New Way to Serendipity', *COMMUNICATIONS & STRATEGIES*, no. 65, pp. 67-89

5. Baupin, N & Zreik K 2000, Remote decision support system: a distributed information management system', *Knowledge-Based Systems*, vol. 13, pp. 37-46

6. Bechhofer, S, Broekstra, J, Decker, S, Erdmann, M, Fensel, D, Goble, C, Van Harmelen, F, Horrocks, I, Klein, M, McGuinness, D, Motta, E, Patel-Schneider, P, Staab, S & Studer, R 2000, An informal description of Standard OIL and Instance OIL. Retrieved February 18, 2006, from www.ontoknowledge.org/oil/downl/oil-whitepaper.pdf

7. Bentley, TJ 1998, *Managing Information – Avoiding Overload*, The Chartered Institute of Management Accountants, London

8. Berghel, H 1997, 'Cyberspace 2000: Dealing with information overload', Association for Computing Machinery, *Communications of the ACM*, February 1997, vol. 40, no. 2; *ProQuest Computing,* pg. 19

9. Bernard A & Tichkiewitch S (eds.) 2008, 'An Overview on Knowledge Management', *Methods and Tools for Effective Knowledge Life-Cycle Management*, Springer-Verlag, Berlin, Heidelberg, pp. 3-21

10. Biezunski, M & Newcomb, SR 2001, 'XML Topic Maps: Finding Aids for the Web', *IEEE MultiMedia*, vol. 8, no. 2, pp. 104-108

11. Blair, DC 2002, 'The challenge of document retrieval, Part I: Major issues, and a framework based on search exhaustivity, determinacy of representation and document collection size', *Information Processing and Management*, vol. 38, pp. 273-291

12. Böhm, K, Engelbach, W, Härtwig, J, Wilcken, M & Delp, M 2005, 'Modelling and Implementing Pre-built Information Spaces. Architecture and Methods for Process Oriented Knowledge Management', Journal of Universal Computer Science, vol. 11, no. 4, pp. 605-633

13. Borchers, A, Herlocker, J, Konstan J & Riedl, J 1998, 'Ganging up on Information Overload', IEEE Computer Society Press, *Computer*, vol. 31, no. 4 , pp. 106-108

14. Breu, M & Ding, Y 2004, *Modelling the World: Databases and Ontologies.* Retrieved February 12, 2006, from http://www.cogproject.org

15. Canadian Food Inspection Agency n.d.*, Appendix A: Glossary of Terms*. Retrieved September 6, 2006 from http://www.inspection.gc.ca/english/corpaffr/publications/riscomm/riscomm_appe.shtml

16. Cann, J 1997, Principles of Classification: Suggestions for a procedure to be used by ICIS in developing international classification tables for the construction industry. Retrieved November 8, 2007, from http://icis.org/siteadmin/rtdocs/images/5.pdf

17. Carlson, CN 2003, 'Information overload, retrieval strategies and Internet user empowerment', *Proceedings of The Good, the Bad and the Irrelevant (COST 269)*, University of Art and Design, Helsinki, Finland, pp. 169-173. Retrieved March 16, 2007, from E-prints in Library and Information Science online database

18. Carr, L, Bechhofer, S, Goble, C & Hall, W 2001, 'Conceptual Linking: Ontology-based Open Hypermedia', *Proceedings of the 10th international conference on World Wide Web*, ACM, New York, NY, pp. 334-342. Retrieved February 1, 2006, from http://www10.org/cdrom/papers/246/index.html

19. Carvalho, MR, Hewett, R & Cañas, AJ 2001. 'Enhancing web searches from concept map-based knowledge models', *Proceedings of SCI 2001: Fifth Multiconference on Systems, Cybernetics and Informatics*, International Institute of Informatics and Systemics, Orlando, FL, pp. 69–73

20. Celko, J 2006, *Discovering Relationships in Context*. Retrieved February 4, 2008 from http://www.ddj.com/database/189401684

21. Chan, CW 2004, 'From Knowledge Modeling to Ontology Construction', International Journal of Software Engineering and Knowledge Engineering, vol. 14, no. 6, pp. 603-624

22. Chan SY 2001, 'The use of graphs as decision aids in relation to information overload and managerial decision quality', *Journal of Information Science*, vol. 27, no. 6, pp. 417–425

23. Chen, C-T & Tai, W-S 2003, 'An information push-delivery system design for personal information service on the Internet', *Information Processing and Management*, vol. 39, pp. 873–888

24. Chen, PP-S 1976, 'The Entity-Relationship Model – Toward a Unified View of Data', *ACM Transactions on Database Systems*, vol. 1, no. 1, pp. 9-36

25. Chenoweth, T, Corral K & Demirkan H 2006,' Seven Key Interventions for Data Warehouse Success', *Communications of the ACM*, vol. 49, no. 1, pp. 115-119

26. Cheung, CF, Lee, WB & Wang Y 2005, 'A multi-facet taxonomy with applications in unstructured knowledge management', *Journal of Knowledge Management*, vol. 9, no. 6, pp. 76-91

27. Choo, CW 2000, Information management for the intelligent organisation: The art of scanning the environment, 2nd edn, Information Today, Inc., Medford

28. Cogito White Paper n.d., Graph-based relationship analytics: Eliminating data analysis barriers. Retrieved May 25, 2006 from www.cogitoinc.com.

29. *Combating Global Warming Mind Map* [image] 2007. Retrieved January 6, 2008, from http://learningfundamentals.com.au/resources/combating-global-warming-mind-map/

30. Conway, S & Sligar, C 2002, *Unlocking Knowledge Assets*, Microsoft Press, Washington

31. Dictionary.com 2008a, *Antonymy*. Retrieved February 24, 2008, from http://dictionary.reference.com/browse/antonymy

32. Dictionary.com 2008b, *Holonymy*. Retrieved February 24, 2008, from http://dictionary.reference.com/browse/holonymy

33. Dictionary.com 2008c, *Hypernymy*. Retrieved February 24, 2008, from http://dictionary.reference.com/browse/hypernymy

34. Dictionary.com 2008d, *Hyponymy*. Retrieved February 24, 2008, from http://dictionary.reference.com/browse/hyponymy

35. Dictionary.com 2008e, *Meronymy*. Retrieved February 24, 2008, from http://dictionary.reference.com/browse/meronymy

36. Dictionary.com 2008f, *Synonymy*. Retrieved February 24, 2008, from http://dictionary.reference.com/browse/synonymy

37. Dictionary.com 2008g, *Validation*. Retrieved October 30, 2008, from http://dictionary.reference.com/browse/validation

38. Edgington, T, Choi, B, Henson, K, Raghu, TS & Vinze, A 2004, 'Adopting Ontology To Facilitate Knowledge Sharing', *Communications of the ACM*, vol. 47, no. 11, pp. 85-90

39. Edmunds, A & Morris, A 2000, 'The problem of information overload in business organisations: a review of the literature', *International Journal of Information Management*, vol. 20, pp. 17-28

40. Elwood, A 2005, 'Setting the Records straight', *Knowledge Ireland*, April 2005, pp. 28-31

41. Eppler , MJ & Mengis, J 2004, 'The Concept of Information Overload: A Review of Literature from Organisation Science, Accounting, Marketing, MIS and Related Disciplines', *The Information Society*, vol. 20, pp. 325-344

42. Everyone's Tags [Image]2007. Retrieved December 28, 2007 from http://myweb.yahoo.com/myweb?ei=UTF-8&dg=6&dmode=vtags&sortby=top

43. Free Dictionary, The 2007, *intelligent*. Retrieved November 7, 2007, from http://www.thefreedictionary.com/intelligent

44. FREMA Ontology [Image] 2006. Retrieved November 12, 2007 from http://frema.ecs.soton.ac.uk/wiki/index.php?title=Image:Latest_Ontology.jpg

45. Gaines, BR & Shaw, MLG 1995, 'Concept maps as hypermedia components', *International Journal of Human-Computer Studies*, vol. 43, no. 3, pp. 323-361

46. Garshol, LM 2004, *Metadata? Thesauri? Taxonomies? Topic Maps!* Retrieved January 7, 2008 from http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html

47. Gilchrist, A 2003, 'Thesauri, taxonomies and ontologies – an etymological note', *Journal of Documentation*, vol. 59, no. 1, pp. 7-18

48. Golder, SA & Huberman, BA 2006, 'Usage patterns of collaborative tagging systems', *Journal of Information Science*, vol. 32, no. 2, pp. 198–208

49. Goodhue, DL, Wybo MD & Kirsch LJ 1992,' The Impact of Data Integration on the Costs and Benefits of Information Systems', *MIS Quarterly*, September 1992, pp. 293-311

50. Greenfield L 2006, *A Definition of Data Warehousing*. Retrieved August 25, 2006 from http://www.dwinfocenter.org/defined.html

51. Gruber, TR n.d., *What is an Ontology?* Retrieved February 1, 2006 from http://www.ksl.stanford.edu/kst/what-is-an-ontology.html

52. Gruber, TR 1993, 'A Translation Approach to Portable Ontology Specification', *Knowledge Acquisition*, vol. 5, pp. 199-220

53. Gruber, TR 2005, *Ontology of Folksonomy: A Mash-up of Apples and Oranges*. Retrieved December 19, 2007 from http://tomgruber.org/writing/ontology-of-folksonomy.htm

54. Hiltz, SR & Turoff, M 1985, 'Structuring computer-mediated communication systems to avoid information overload', Association for Computing Machinery, Communications of the ACM July 1985, vol. 28, no. 7, pp. pp. 680-689

55. *Hippopotamus Taxonomy* [image] n.d. Retrieved November 8, 2007, from http://www.nnf.org.na/RARESPECIES/InfoSys/IMAGES/Hippo/Fig2%20Taxonomy.gif

56. Ho, J & Tang, R 2001, 'Towards an Optimal Resolution to Information Overload: And Infomediary Approach', *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work.* Boulder, Colorado, USA, pp. 91-96

57. *How much information? 2003*. Retrieved March 8, 2007, from http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/

58. Hwang, MI, Lin, JW 1999. 'Information dimension, information overload and decision quality', *Journal of Information Science*, vol. 25, no. 3, pp. 213-218

59. HyBrow n.d., *Conceptual Framework Overview*. Retrieved February 7, 2007 from http://wartik19.biotec.psu.edu/Docs/Conceptual.html

60. *Information Management Independent Research Results 2006.* Information Management Solutions 2006, Retrieved March 13, 2007, from http://www.ims-show.co.uk/ol06/ims/files/research.brochure.1061001120112.pdf

61. *Information Overload Statistics* 2003, Humboldt State University Library. Retrieved February 13, 2007, from http://library.humboldt.edu/~ccm/fingertips/ioverloadstats.html

62. Inmon, W, O'Neill, B & Fryman, L 2008, *Business Metadata – Capturing Enterprise Knowledge*, Elsevier Inc., Morgan Kaufman Publishers, Burlington

63. ISO/IEC 13250 Topic Maps 2002. Retrieved January 13, 2008 from www.y12.doe.gov/sgml/sc34/document/0322_files/iso13250-2nd-ed-v2.pdf

64. Kendall, KE & Kendall JE 2002, *System Analysis and Design*, 5th edn, Pearson Education Inc., New Jersey

65. Kim, HM 2000, 'Enabling integrated decision making for electronic commerce by modelling an enterprise's sharable knowledge', *Internet Research: Electronic Networking Applications and Policy*, vol. 10, no. 5, pp. 418-423

66. Kirk, EE 1996, *Evaluating Information Found on the Internet.* Retrieved January 19, 2007, from http://www.library.jhu.edu/researchhelp/general/evaluating/

67. Kishore, R, Sharman, R & Ramesh, R 2004a, 'Computational Ontologies and Information Systems: I. Foundations', *Communications of the Association for Information Systems*, vol. 14, pp. 158-183

68. Kishore, R, Sharman, R & Ramesh, R 2004b, 'Computational Ontologies and Information Systems: II. Formal Specification', *Communications of the Association for Information Systems*, vol. 14, pp. 184-205

69. Kotze, DJ, Uys, JW, Schutte, C & Du Preez, ND 2007, 'The Implementation of a Conceptual Framework Based Approach for the Improved Viewing and Utilisation of Organisational Information'*, Proceedings of the International Conference on Competitive Manufacturing 2007 (COMA '07)*, Global Competitiveness Centre in Engineering, Stellenbosch, pp. 413-418

70. Lacher, MS & Decker, S 2001, 'RDF, Topic Maps, and the Semantic Web', *Markup Languages: Theory & Practice,* vol. 3, no. 3, pp. 313-331

71. Lau, T & Sure, Y 2002, 'Introducing Ontology-based Skills Management at a Large Insurance Company', Modellierung 2002, Modellierung in der Praxis - Modellierung für die Praxis, Tutzing, Deutschland, pp. 123-134

72. Leake, DB, Maguitman, A & Cañas, A 2002, 'Assessing Conceptual Similarity to Support Concept Mapping', *Proceedings of FLAIRS-2002*, AAAI Press, pp. 168-172. Retrieved November 1, 2007 from www.cs.indiana.edu/~leake/papers/p-02-02.pdf

73. Lee, S, Lee, T, Kim, D & Chun, J 2007, 'A Bottom-Up Strategy for Enterprise Ontology Implementation', Proceedings of the eighteenth IEICE Data Engineering Workshop and the fifth annual conference of the Database Society of Japan, The Database Society of Japan and the IEICE Data Engineering Technical Group. Retrieved February 19, 2008 from www.ieice.org/~de/DEWS/DEWS2007/pdf/m6-4.pdf

74. Mathes, A 2004, *Folksonomies - cooperative classification and communication through shared metadata.* Retrieved December 19, 2007, from http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html

75. McCarthy, WE 1982 'The REA Accounting Model: A Generalized Framework for Accounting Systems in a Shared Data Environment', *The Accounting Review*, vol. LVII, no. 3, pp. 554-578

76. Mertins, K, Heisig, P & Vorbeck, J (ed.) 2005, *Knowledge Management Concepts and Best Practices,* Springer-Verlag, Berlin

77. Meyer, J-A 1998 'Information overload in marketing management', *Marketing Intelligence & Planning*, vol. 16, no. 3, pp. 200-209

78. Newcomb, SR & Biezunski, M 2001, 'A Processing Model for Topic Maps', *Presentation at the first Knowledge Technologies conference*, Austin, TX. Retrieved September 24, 2007 from http://www.infoloom.com/tmartic.htm

79. Noruzi, A 2006, 'Folksonomies: (Un)Controlled Vocabulary?', Knowledge Organisation, vol. 33, no. 4, pp. 199-203

80. Novak, JD & Cañas, AJ 2006, *The Theory Underlying Concept Maps and How to Construct Them*, Technical Report IHMC CmapTools 2006-01, Florida Institute for Human and Machine Cognition. Retrieved September 24, 2007 from http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf

81. Noy, NF & McGuinness, DL 2002, *Ontology Development 101: A Guide to Creating Your First Ontology.* Stanford University, Stanford, CA, Stanford Medical Informatics Report SMI-2001-0880.

82. Oakes, M n.d., *Thesauri: Ontologies for Information Retrieval*. Retrieved November 8, 2007 from http://osiris.sunderland.ac.uk/~cs0moa/iistut2.doc

83. Paes, P 1994, 'Agents that reduce work and information overload', Association for Computing Machinery, *Communications of the ACM*, July 1994, vol. 37, no. 7, *ProQuest Computing,* p. 30

84. Pepper, S 2002, *The TAO of Topic Maps*. Retrieved February 12, 2007, from http://www.ontopia.net/topicmaps/materials/tao.html

85. Pipe, P 1997,'The Data Mart: A New Approach to Data Warehousing', *International Review of Law Computers & Technology*, vol. 11, no. 2, pp. 251-261

86. Rath, HH 2003, *The Topic Maps Handbook*, Empolis GmbH, Gutersloh, Germany 2003. Retrieved January 7, 2008 from http://www.empolis.com/downloads/empolis_TopicMaps_Whitepaper20030206.pdf

87. *Recordkeeping Manual - Glossary of Recordkeeping Terms* 2007. Retrieved October 4, 2007, from http://www.records.curtin.edu.au/recordkeeping_manual/glossary.html

88. Rob, P & Coronel, C 2000, *Database Systems: Design, Implementation and Management*, 4th edn, Course Technology, Cambridge

89. Rowley, J & Farrow, J 2000, *Organising Knowledge: An Introduction to Managing Access to Information*, 3rd edn, Gower Publishing Limited, Hampshire

90. Rubin, DL, Hewett, M, Oliver, DE, Klein, TE & Altmann, RB 2002, 'Automating Data Acquisition into Ontologies from Pharmacogenetics Relational Data Sources using Declarative Object Definitions and XML', *Proceedings of the seventh Pacific Symposium on Biocomputing*, pp.88-99. Retrieved October 23, 2007, from http://psb.stanford.edu/psb-online/proceedings/psb02/rubin.pdf

91. Schreiber, Z 2003, Semantic Information Management (SIM): Solving the Enterprise Data Problem by Managing Data Based on its Business Meaning. Retrieved February 12, 2006, from http://www.cogproject.org

92. SearchOracle.com 2005, *Definition – data mart*. Retrieved August 25, 2006, from http://searchoracle.techtarget.com/gDefinition/0,,sid41_gci211900,00.html

93. Sherman, R 2005a, 'The Enterprise Data Warehouse Strikes Again, Part 1', *DM* Review, November 2005, pp. 12, 44

94. Sherman, R 2005b, 'The Enterprise Data Warehouse Strikes Again, Part 2, *DM* Review, December 2005, pp. 88, 90

95. Shirky, C 2005, *Ontology is overrated: Categories, Links and Tags*. Retrieved August 28, 2006 from http://www.shirky.com/writings/ontology_overrated.html

96. Sinclair, J & Cardew-Hall, M 2007, `The folksonomy tag cloud: When is it useful?', Journal of Information Science, pp. 0165551506078083+

97. Skoutas, D & Simitsis, A 2006, 'Designing ETL processes using semantic web technologies', *Proceedings of the 9th ACM international workshop on Data warehousing and OLAP*, Arlington, Virginia, USA, pp. 67-74. Retrieved April 9, 2008 from http://portal.acm.org/citation.cfm?id=1183512.1183526

98. Sowa, JF 2006, *Semantic Networks*. Retrieved November 7, 2007, from http://www.jfsowa.com/pubs/semnet.htm

99. Speller, E 2007, *Collaborative tagging, folksonomies, distributed classification or ethnoclassification: a literature review*. Retrieved December 19, 2007 from http://informatics.buffalo.edu/org/lsj/articles/speller_2007_2_collaborative.php

100. Spiteri, LF 2007, 'Structure and form of folksonomy tags: The road to the public library catalogue', *Webology*, vol. 4, no. 2. Retrieved December 19, 2007 from http://www.webology.ir/2007/v4n2/a41.html

101. STARlab – Research n.d., Research interests: The Semantic Web, Databases and Ontologies. Retrieved January 20, 2006 from http://www.starlab.vub.ac.be/research/index.htm

102. Stock, WG 2007, 'Folksonomies and science communication. A mash-up of professional science databases and Web 2.0 services', *Information Services & Use*, vol. 27, no. 3, pp. 97-103

103. Thinkmap Visual Thesaurus [Image] 2007, *intelligent*. Retrieved November 7, 2007, from http://www.visualthesaurus.com/

104. *Topic Map Example* [Image] 2007. Retrieved January 15, 2008 from http://www.wandora.org/wandora/wiki/index.php?title=Image:Topicmap_example01.gif

105. Uys, W 2005, 'A Conceptual Framework based Approach for the Improved Viewing and Utilisation of Organisational Databases', *Presentation given at the 19th SAIIE and 35th ORSSA Annual Conference,* Vanderbijlpark, South Africa

106. Vatant, B 2004, *Ontology-driven topic maps.* Retrieved February 1, 2006 from www.idealliance.org/europe/04/call/xmlpapers/03-03-03.91/.03-03-03.html

107. Vrandečić, D, Pinto, S, Tempich, C & Sure, Y 2005, 'The DILIGENT knowledge processes', Journal of Knowledge Management, vol. 9, no. 5, pp. 85-96

108. *What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model?* 2003. Retrieved August 30, 2006 from http://www.metamodel.com/article.php?story=20030115211223271

109. *What We Offer* 2008. Retrieved April 29, 2008 from http://www.indutech.co.za/whatweoffer.htm

110. Wikipedia, the free encyclopedia 2007a, *Concept Map*. Retrieved February 12, 2007 from http://en.wikipedia.org/wiki/Concept_map

111. Wikipedia, the free encyclopedia 2007b, *Database.* Retrieved October 8, 2007, from http://en.wikipedia.org/wiki/Database

112. Wikipedia, the free encyclopedia 2007c, *Information Management*. Retrieved October 4, 2007, from http://en.wikipedia.org/wiki/Information_management

113. Wikipedia, the free encyclopedia 2007d, *Knowledge Management*. Retrieved October 10, 2007, from http://en.wikipedia.org/wiki/Knowledge_management

114. Wikipedia, the free encyclopedia 2007e, *Semantic Network*. Retrieved February 12, 2007, from http://en.wikipedia.org/wiki/Semantic_network

115. Wikipedia, the free encyclopedia 2007f, *Topic Map*. Retrieved February 12, 2007 from http://en.wikipedia.org/wiki/Topic_map

116. Withrow, S 2004, *How project champions can rally support for your team*. Retrieved March 16, 2008 from http://www.builderau.com.au/manage/soa/How_project_champions_can_rally_support_for_your_team/0,39024662,39130323,00.htm

117. WordNet Search 3.0 n.d., *Word to search for: database.* Retrieved October 8, 2007, from http://wordnet.princeton.edu/perl/webwn?s=database

118. Zantout, H & Marir, F 1999, 'Document management systems from current capabilities towards intelligent information retrieval: an overview', *International Journal of Information Management*, vol 19, pp. 471-484

119. Zhong, N 2002, 'Representation and Construction of Ontologies for Web Intelligence', International Journal of Foundations of Computer Science, vol. 13, no. 4, pp. 555-570

Appendices

Appendix A

Organon

INDUTECH
Your Key to Manage Innovation

ORGANON

CONCEPTUAL FRAMEWORK EDITOR

Tel +27 (0)21 887 1180 **Fax** +27 (0)21 887 1185
Pinotage House, Brandwacht Office Park, Trumali Road
Stellenbosch, 7600 **www.indutech.co.za** **info@indutech.co.za**

Mind maps and concept maps have gained popularity during the past few years as handy tools for structuring ideas before focusing on detail. These maps consist of a root topic and several lower-order topics - in essence representing a tree of topics or concepts. Usually, each topic can be explained with an associated note, while relations between related topics can also be drawn; it is even possible to attach binary files to a topic using some editors.

Although mind maps are excellent tools for structuring ideas around a certain topic, they are somewhat limiting when it comes to representing more complex environments consisting of several topics with explicit and implicit relations among them. In order to adequately represent such an environment, a network structure - consisting of entities and relations - is required.

Organon was developed as a tool to help its users to structure unstructured, qualitative problems to gain an understanding of the problem without losing the context of the detailed elements of the entities in the makeup of the problem. The network structures that can be created, explored and edited using Organon are called conceptual frameworks. Using Organon, various users can explore the conceptual framework (CF) of a certain environment, starting from an entity known to them, and reaching related, unknown entities by clicking on self-explanatory relations gaining an understanding of the environment as they go along. Users can further expand the CF by adding new entities or by associating existing entities - using new or existing relations - adding their own understanding to the CF while doing so.

A simple example illustrating the value of a CF is presented overleaf.

Organon has the following features:

- Explore the CF - by clicking on entities and associated relations - to gain an understanding of the environment represented.
- Expand the CF by relating existing entities to new entities using existing or new relations.
- Search for all entities matching a given string and start exploring the CF from any entity in the result set.
- Find how two sets of entities are connected in the CF.
- Build a tree view starting at a given entity and expanding on specified relations.
- Create templates for frequently used entities to speed up the process of expanding the CF.
- Get more information about a given entity in one click by using the Lookup on Google or Lookup on Wikipedia commands.

It will shortly be possible to link CF entities to actual text in electronic documents to provide an additional dimension to the context of an entity.

SHOULD YOU REQUIRE MORE INFORMATION ABOUT ORGANON, PLEASE CONTACT US AT INFO@INDUTECH.CO.ZA.

© Indutech (Pty) Ltd

ORGANON

CONCEPTUAL FRAMEWORK EXAMPLE

Suppose one wants to create a conceptual model of a manufacturing company to serve as a mechanism to capture the various complexities involved in the relevant environment with the aim to gain a better understanding of the environment among the employees to generate ideas for possible simplifications.

One may then identify certain generic entity types that would form part of the model (e.g. departments, employees, products, raw materials, manufacturing equipment, suppliers, etc.) as well as initial relations applying among the entity types identified.

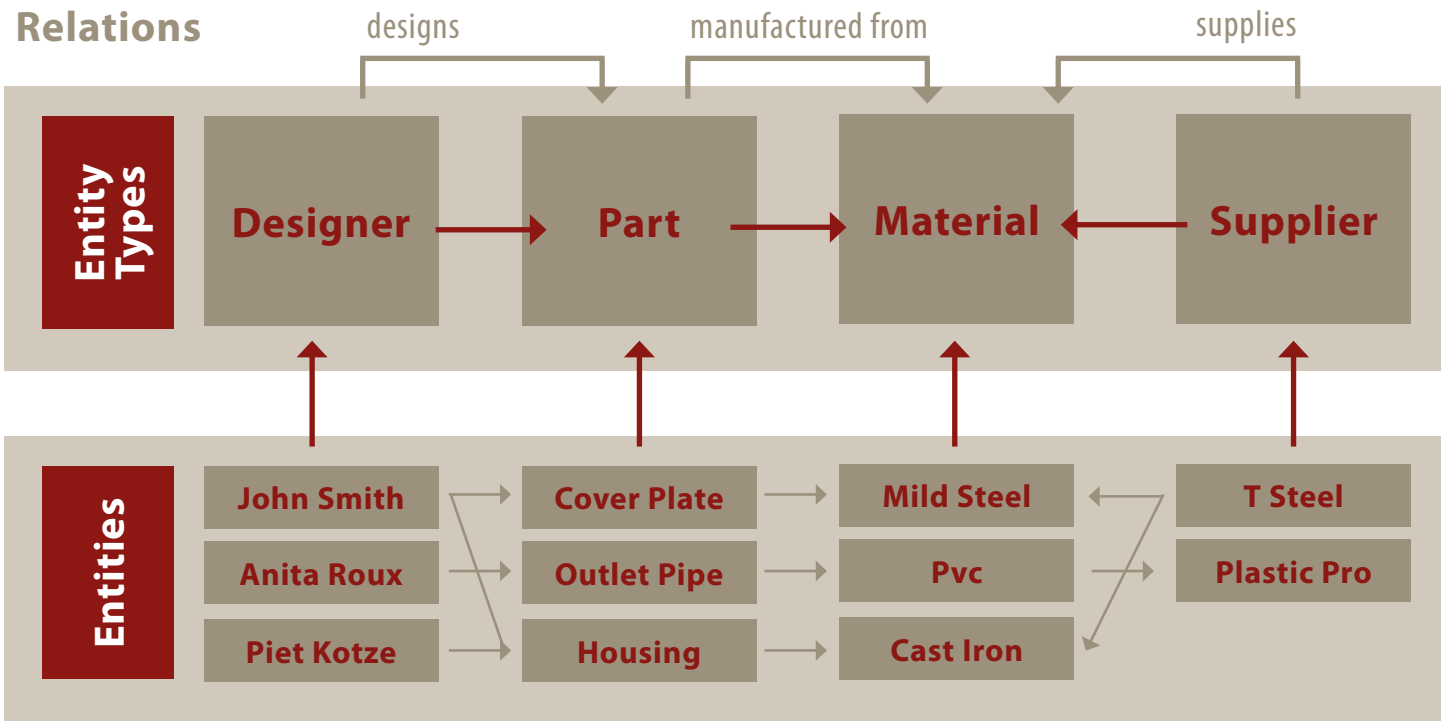
The network resulting from the combination of entity types and accompanying relations is called a Conceptual Framework (CF). Any entity [type] - relation - entity [type] pair is called a fact (e.g. 'Part manufactured from Material' or 'John Smit is a Designer' where 'Part' and 'Material' are entity types, 'manufactured from' and 'is a' are relations and 'John Smit' is a entity).

Typically, several departments will be involved in the construction of the CF, e.g. the HR department will add all employee information, the Procurement department will specify all suppliers and materials, whereas the Production department will add all products, machines, etc. Each party involved expands the CF with his/her view providing more information, context and dependencies.

For example, the Production department may add the entity type 'Manufacturing Processes' to capture which processes are used to produce which products; the same department may also associate certain materials with certain products and manufacturing processes - building on the information (i.e. the various materials) added by the Procurement department.

Using the CF, the Procurement department can now learn what the various materials they order everyday are used for in terms of products and processes, just as the Production department can learn from whom they receive the various materials used in the production of the various products.

The CF can therefore be used to understand the manufacturing environment as a whole, as well as the interactions between the various entities that make up the environment.



© Indutech (Pty) Ltd

Appendix B

Implementation I List of facts

Entity	Relation	Entity
Department	do work for	Client
Department	department uses	Keyword
Department	is made up of	Person
Department	is responsible for	Project
Department	takes part in project	Project
Document	was booked in on	Date
Document	was booked out on	Date
Document	was changed on	Date
Document	was created on	Date
Document	was deleted on	Date
Document	was read on	Date
Document	is found in EDEN	EDEN
Document	has rights assigned	EDEN Rights
Document	is worked on by role	EDEN Role
Document	is described by keyword	Keyword
Document	is worked on by	Person
Document	is created for	Project
EDEN	has role	EDEN Role
EDEN	is distinguished by	Keyword
EDEN	is used by	Person
EDEN	is used in	Project
EDEN	use required by	Project Role
EDEN Rights	is allocated to	EDEN Role
EDEN Rights	is ascribed to	Person
EDEN Role	consists of	Person
EDEN Role	rights of EDEN role is allocated to	Project Role
E-mail	e-mail belongs to	Client
Keyword	was added to document on	Date
Person	was at work on	Date
Person	has e-mail	E-mail
Person	is characterised by	Keyword
Person	is involved in	Project
Person	in project acts as	Project Role
Person	entered building at	Time
Project	is carried out for	Client
Project	is worked on date	Date
Project	was ended on	Date
Project	was started on	Date
Project	has project roles	Project Role
Project	has individual time spent to date	Work Duration
Project	has total time spent up to date	Work Duration
Work Duration	is individual project duration on date	Date
Work Duration	is total project duration on date	Date
Work Duration	spent up to date by	Person

Appendix C

Implementation I Questionnaire

Questionnaire of a Conceptual Framework Implementation at Indutech

This questionnaire is aimed at determining the success of the Indutech Conceptual Framework Implementation. Please feel free to elaborate on any point – all feedback is important!!

Impressions

- Benefits: *Comment on the usefulness of the CF (not Organon!). Referring to specific CF facts that proved/might prove to be useful should be stated.*

- Shortcomings: *Shortcomings of the content or structure should be stated in as much detail as possible.*

Growing the CF

What information can be added to the CF to make it even more useful/insightful?

Organon Bugs/Suggestions

Please list all Organon bugs uncovered and possible improvement suggestions based on the use of Organon to view the CF.

Methodology comments

Please list any suggestions with regards to the methodology that was followed to develop the CF.

Other comments

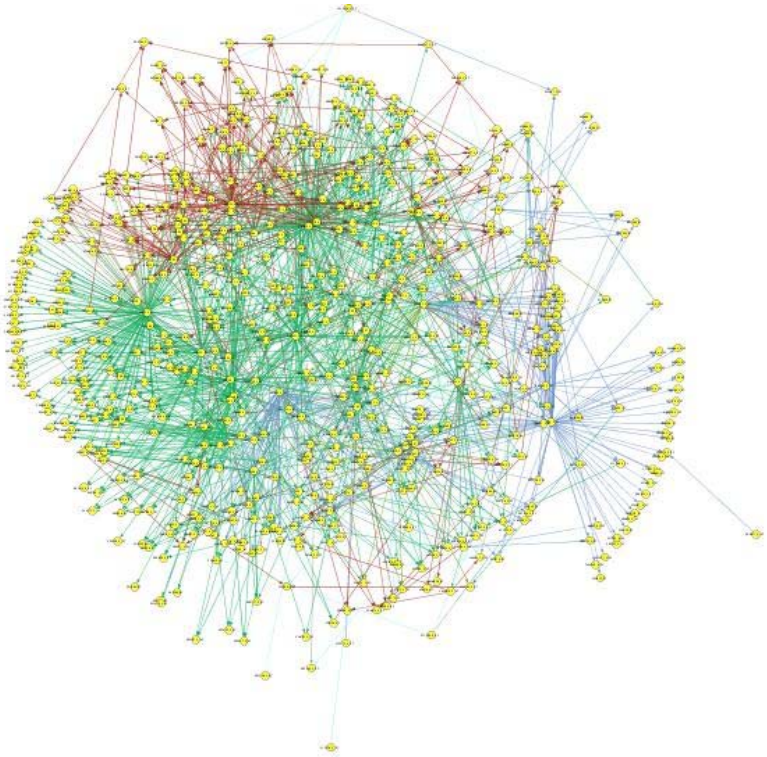
Please list any other comments.

Appendix D

Implementation 2 Design document

Conceptual Framework Implementation

Documentation



Planning and Analysis

Document

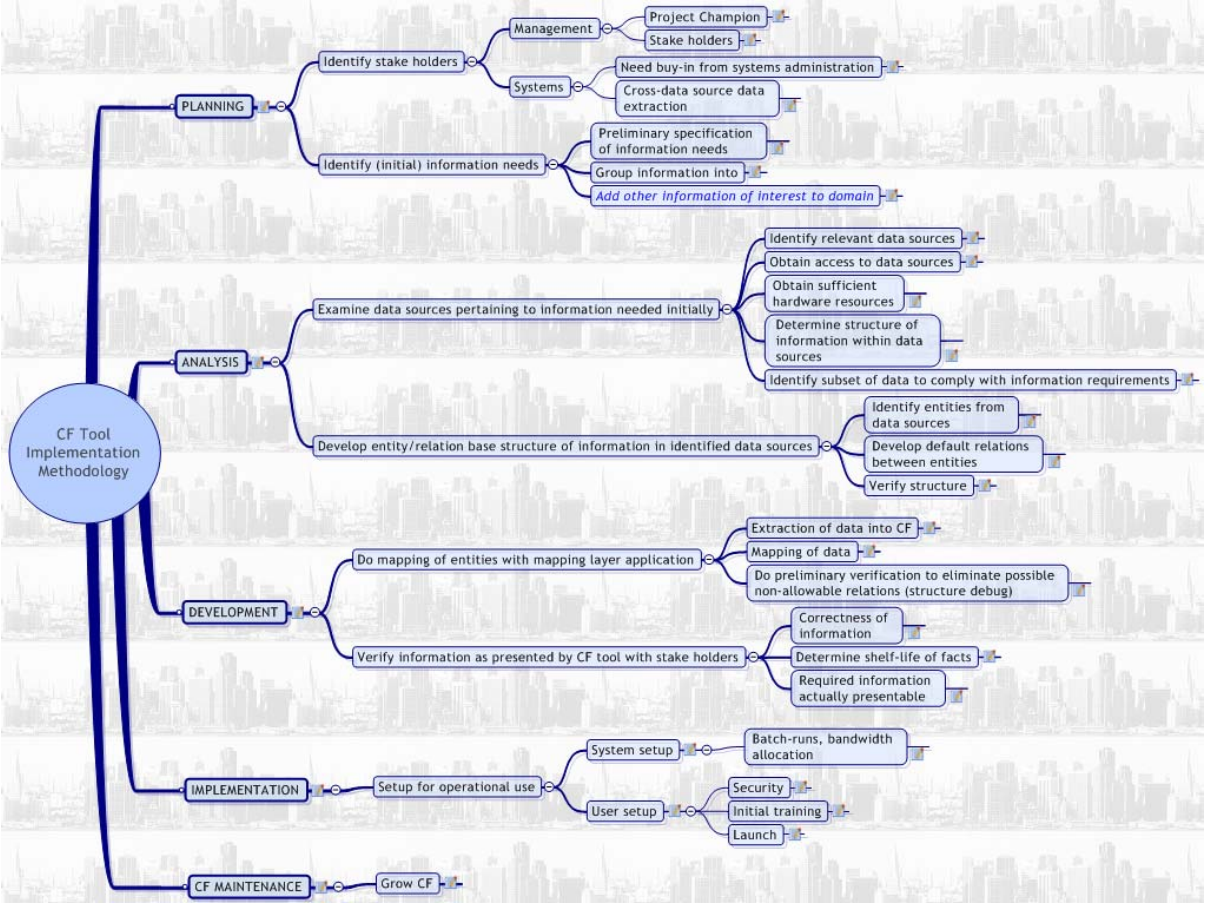
Implementation: **Case Study - University Of Stellenbosch**

Date: **9 September 2007**

Author: **Dirk Kotze**

CONTENTS

Methodology



This document contains outputs from the PLANNING step of the Conceptual Framework Tool Implementation Methodology.

Reason for the project

The project is conducted primarily as a research case study by the Industrial Engineering Department of the University of Stellenbosch and Indutech (Pty) Ltd. for testing an implementation methodology developed for establishing a Conceptual Framework approach for the utilising of institutionalised information in an organisation. A secondary reason for the case study is to get a feeling for whether the particular organisation finds the information portrayed by the Conceptual Framework, as well is the manner in which it is portrayed, useful for decision making and whether the ease of locating specific information has been improved. The Information Technology department of the University of Stellenbosch has tentatively agreed to the case study in order to have a fresh look on the

information in its possession, as well as assisting in furthering the scientific prowess of the University of Stellenbosch.

Stakeholders

This section identifies the stakeholders of the Conceptual Framework implementation process.

Principal Stakeholder:

Helmi Dreijer – Senior Director: Information Technology

Tel: *+27 21 808 4246*

E-mail: *mwd@sun.ac.za*

Primary Stakeholders:

Johann Kistner – Director: Administrative Information

Tel: *+27 21 808 4597*

Email: *jck@sun.ac.za*

Joe Smit – Director: User Services

Tel: *+27 21 808 4209*

Email: *jsmit@sun.ac.za*

Ralph Pina – Director: e-Business and Card Facility Manager

Tel: *+27 21 808 4271*

Email: *ralph@sun.ac.za*

Secondary Stakeholders:

Anne-Marie du Toit – Manager: Student Information Systems

Tel: *+27 21 808 4256*

Email: *ams@sun.ac.za*

Danie Malan – Head Engineer: Systems Integration

Tel: *+27 21 808 4367*

Email: *dgm@sun.ac.za*

Gerhard van Wageningen – (WebCT usage information)

Tel: *+27 21 808 4554*

Email: *gerhardv@sun.ac.za*

Hendrik le Roux – Manager: Access Control

Tel: *+27 21 808 3795*

Email: *hleroux@sun.ac.za*

Le Roux Franken – (Printing transaction information)

Tel: *+27 21 808 4048*

Email: *lrf@sun.ac.za*

Wouter Klapwijk – Senior Software Specialist: Library Services

Tel: *+27 21 808 4378*

Email: *wklap@sun.ac.za*

Domain of interest

The domain to be mapped is aimed at studying the impact that users at the university (both students and personnel) have on the university's information infrastructure. Sub-domains to be included in the project are:

- Card reader transaction information
- Demographic and study programme information
- Internet usage cost information
- Library transaction and item information
- Printing transaction information
- WebCT usage information

The US Conceptual Framework - Entities

Through a study made of the domain to be mapped (including the sub-domains listed in the previous section), the following entities have been identified:

Entity Name	Description
Age	Age of a US[1] student or a library item.
Date	Any references made to dates in the US IT[2] environment being mapped. This include the following: • Dates of card reader transactions; • Dates of internet transactions; • Dates of library transactions; • Dates of printing transactions; and • Dates of WebCT actions.
Document Type	Types of documents, e.g. MS Word documents, MS Excel documents, PDF documents, etc. (mainly related to printing operations).
Education department	Denoting a South African education department, from which a scholar can receive a senior certificate used in applying for study at the US.
File Size	Byte size of a computer file (mainly related to printing operations).
Gender	Male/Female.
Handicap	Manner of handicap experienced by a US student (this includes wheelchair handicap).
High school	High school attended by a US student.
Language	Language spoken by a US student.
Library Branch	Branch of the US library services where items can be loaned to students.
Library Database Name	Name of an on-line information database managed by the US library, and which can be accessed by students through a web-portal.
Library Dewey Classification	Number of a library item corresponding to a numbering system utilised by the US library to categorise its items.
Library Item Author	Author of a library item.
Library Item Collection	Indication of the type of US library material (e.g. reference material).

[1] University of Stellenbosch
[2] Information Technology

Entity Name	Description
Library Item Loan Amount	Amount of times library items have been loaned out.
Library Item Material	Media format of a US library item.
Library Item Publisher	Publisher of a US library item.
Library Item Status	Status of a US library item.
Library Item Subject Topic	Broad topic that includes a specific US library item.
Library Item Title	Title of a US library item.
Library Item Vendor	Vendor of a US library item.
Library Shelf Code	Shelf code where a US library item is stored.
Library Shelf Code Type	Distinguishes between different shelf code systems used by the US library.
Library Transaction Location	Indicates whether a US library transaction was conducted at the counter in the library, or through a web portal.
Library Transaction Patron Type	Type of person performing a certain transaction at the US library.
Library Transaction Type	Type of transaction performed at the US library.
Marital status	Marital status of a US student.
Nationality	Nationality of a US student.
Network traffic	Network usage experienced by US IT systems (mainly indicates internet usage and usage of US library on-line information sources in bytes per hour).
Number of pages	Number of pages used through printing and photo-copying activities.
Page Cost	Page cost of printing and photocopying activities.
Person	User of the US IT system, accessing the system through UT number or computer username. This includes the usage of card readers, internet access, printing and library services and WebCT. It further provides access to academic and demographic information stored by US on its students.
Population group	Population group of a US student.
Postal code	Postal code of US student.

Entity Name	Description
Time band	Any references to time made in the US IT system, grouped in time bands. This include the following: • Times of card reader transactions; • Times of internet transactions; • Times of library transactions; • Times of printing transactions; and • Times of WebCT actions.
Town	Home town of US students as well as cities where certain publishers of US library items are situated.
US Campus	Indicates one of the three main US campuses: Stellenbosch, US Business School and Saldanha Military Academy.
US Card Reader	Name of a US card reader.
US Card Reader Type	Indicating type of US card reader (e.g. door access, photo copier, washing equipment, etc.)
US Department	Department of the US.
US Faculty	Faculty of the US.
US Module	Module (or subject) taken by US student and presented by a US department.
US Photocopier	Name of a US photocopier.
US Printer	Name of a US printer.
US Programme	Name of a programme presented by the US.
US Programme Continuity	Indicates whether a programme is studied full-time or part-time.
US Qualification	Type of US qualification that can be attained by a student, e.g. degree, diploma, certificate, etc.
US Residence	Name of a US residence. Only applicable to the usage of card readers with regards to washing equipment.
US Residence Type	Indicates a type of US residence, e.g. hostel, student house, etc.
US Vehicle license plate	License plate number of US vehicles in the US motor pool.
US Washing Equipment	Type of US washing equipment that can be used by US students, e.g. washing machine, tumble drier, etc.

Entity Name	Description
US Workstation/PC	Name of a US workstation or PC used for printing of documents and internet access.
WebCT Action	Action performed by a user in the US WebCT environment.
WebCT Dwell time	Time spent by a user in the US WebCT environment between actions.
WebCT Event Role	Role of a user when performing an action in the WebCT environment.
WebCT Learning Context	Learning context in the WebCT environment.
WebCT Learning Context Administrative Period	Period for which a specific WebCT learning context is accessible.
WebCT Learning Context Role	The role of a user in a specific learning context.
WebCT Learning Context Type	Indicates the level where a learning context can be found in the WebCT environment.
WebCT Number	Any references to numbers in the WebCT environment, usually used in reporting the contents of WebCT templates and WebCT learning contexts.
WebCT Organization	Name of organisation using specific WebCT learning contexts.
WebCT Organization Type	Value distinguishing general categories of the organisation.
WebCT Organization Unit	Name of sponsoring/administering unit within the organisation.
WebCT Size	Any references to sizes in the WebCT environment, usually used in reporting the file sizes.
WebCT Template	Component of WebCT learning context that can be utilised by users.

The US Conceptual Framework – Relations

Different types of relations can be defined between the various entities listed in the previous section. The relation types which will be used in this mapping are **basic**, **inferred** and **aggregation** relations.

Basic relations are the relations that are inherent to the various sub-domains being mapped and which already existed before the mapping was done. **Inferred relations** are relations identified through the over-lapping of domains. Because an immense number of inferred relations can be developed from overlapping domains, only the relations that make most sense at the time of mapping will be included, but others can be added later on if so requested by the stakeholders who will ultimately review the final Conceptual Framework. **Aggregate relations** are relations providing links to aggregate values that can be included in the Conceptual Framework in order to provide more useful information to decision makers who use the Conceptual Framework. Aggregate values such as averages, totals, minimums, and maximums are just some of the values that can be provided and obviously focus more on entities of a numerical nature.

Due to the fact that the relations are quite numerous in nature, the relations (grouped per sub-domain) are listed in appendices to this document as follows:

Appendix A – Basic Relations

Relations inherent to the following sub-domains (where relevant, aggregate relations are included):

1. Card reader transaction information
2. Demographic and study programme information
3. Internet usage cost information
4. Library transaction and book information
5. Printing transaction information
6. WebCT usage information

Appendix B – Inferred Relations

Relations inherent to the following overlapping of sub-domains (where relevant, aggregate relations are included):

- Card reader transaction information
    1. ←→ Internet usage cost information
    2. ←→ Library transaction and book information
    3. ←→ Printing transaction information
    4. ←→ WebCT usage information
- Demographic and study programme information
    5. ←→ Card reader transaction information
    6. ←→ Internet usage cost information
    7. ←→ Library transaction and item information
    8. ←→ Printing transaction information

9. ←→ WebCT usage information

- Internet usage cost information
  10. ← → Library transaction and book information
  11. ← → WebCT usage information
- Library transaction and book information
  12. ← → Printing transaction information
  13. ← → WebCT usage information
- Printing transaction information
  14. ← → WebCT usage information

Decisions

The following are decisions that must be made before the next phase of the Conceptual Framework implementation can be undertaken:

Student privacy concerns

A possibility exists that the usage and presentation of student numbers by the Conceptual Framework might intrude on the privacy of students. The stakeholders should decide whether this poses a risk to university-student relations, and how this risk may be mitigated.

Location of the Conceptual Framework

A network location should be identified where the Conceptual Framework would reside once constructed. Considerations for this decision includes network security, data transfer speed, and ease of access for both users and developers of the Conceptual Framework.

Delivery date of Conceptual Framework

A decision is needed on the delivery date of the Conceptual Framework, especially to allow for enough time that stakeholders can sufficiently evaluate the information presented by the Conceptual Framework, and the usefulness of the Conceptual Framework concept.

Next Steps

Once decisions regarding the aspects listed in the previous section have been made, information must be acquired from the six sub-domains. With this information, a Conceptual Framework is to be constructed, and presented to the stakeholders on the delivery date.

Conclusion

This document provided some initial information on the implementation of a Conceptual Framework at the University of Stellenbosch. Information on the stakeholders of the implementation was provided, as well as the sub-domains within the university environment to be mapped, the entities that exist within these sub-domains and the relations that exist between the entities. The next phase of the implementation entails the mapping of university data onto the defined entity-relation structure as defined in the Conceptual Framework.

Appendix A – Basic Relations

A1 - Card reader transaction information

Entity	Relation	Entity	Relation Type
Date	is date on which door was entered by	Person	Basic
Date	is date on which photocopies were made by	Person	Basic
Date	is date on which vehicle was taken from motor pool by	Person	Basic
Date	is date on which washing equipment was used by	Person	Basic
Date	is date on which door was entered at	Time Band	Basic
Date	is date on which meal was taken at	Time Band	Basic
Date	is date on which photocopies were made at	Time Band	Basic
Date	is date on which vehicle was taken from motor pool at	Time Band	Basic
Date	is date on which washing equipment was used at	Time Band	Basic
Date	is date on which photocopies were made with	US Photocopier	Basic
Date	is date on which washing equipment was used at residence	US Residence	Basic
Date	is date on which vehicle was taken from motor pool with licence plate	US Vehicle License Plate	Basic
Number of Pages	is number of pages that was photocopied on	Date	Basic
Number of Pages	is number of pages photocopied at page cost	Page Cost	Basic
Number of Pages	is number of pages that was photocopied by	Person	Basic
Number of Pages	is number of pages that was photocopied at	Time Band	Basic
Number of Pages	is number of pages printed through activation of card reader	US Card Reader	Basic
Number of Pages	is number of pages photocopied with	US Photocopier	Basic
Page Cost	is page cost that was incurred through photocopying on	Date	Basic
Page Cost	is page cost that was incurred through photocopying at	Time Band	Basic
Page Cost	is page cost that was incurred through photocopying with	US Photocopier	Basic
Person	is person who did not take meal on	Date	Basic
Person	is person who took meal on	Date	Basic
Person	is person who makes the most photocopies with per person average daily number of pages photocopied	Number of Pages	Aggregation
Person	is most expensive photocopier of documents with per person average daily page cost per document	Page Cost	Aggregation
Person	is person who made photocopies with page cost	Page Cost	Basic
Person	is person on daily average access card reader-controlled doors the most at	Time Band	Aggregation
Person	is person on daily average making the most photocopies at	Time Band	Aggregation
Person	is person on daily average takes meals the most at	Time Band	Aggregation
Person	is person on daily average taking vehicles from the vehicle pool the most at	Time Band	Aggregation
Person	is person on daily average using washing equipment the most at	Time Band	Aggregation
Person	is person who entered door at	Time Band	Basic
Person	is person who took meal at	Time Band	Basic
Person	is person who makes most use of card reader	US Card Reader	Aggregation
Person	is person who activated card reader type	US Card Reader Type	Basic
Person	is person who makes most use of card reader type	US Card Reader Type	Aggregation
Person	is person who is most frequent user of photocopier	US Photocopier	Aggregation
Person	is person who made photocopies with	US Photocopier	Basic
Person	is person who mostly uses photocopier	US Photocopier	Aggregation
Person	is person who took vehicle from motor pool with licence plate	US Vehicle License Plate	Basic
Time Band	is time at which photocopies were made by	Person	Basic
Time Band	is time at which vehicle was taken from vehicle pool by	Person	Basic
Time Band	is time at which washing equipment was used by	Person	Basic
Time Band	is time at which photocopies were made with	US Photocopier	Basic
Time Band	is time at which vehicle was taken from motor pool with licence plate	US Vehicle License Plate	Basic
US Card Reader	is card reader activated on	Date	Basic
US Card Reader	is card reader responsible for most photocopies with average daily number of pages photocopied	Number of Pages	Aggregation
US Card Reader	is card reader activated by	Person	Basic
US Card Reader	is card reader activated at	Time Band	Basic
US Card Reader	is card reader most frequently activated at	Time Band	Aggregation
US Card Reader	is card reader that activates photocopier	US Photocopier	Basic
US Card Reader	is card reader that activates washing equipment	US Washing Equipment	Basic
US Card Reader Type	is card reader type activated on	Date	Basic
US Card Reader Type	is card reader type activated by person at	Time Band	Basic
US Card Reader Type	is card reader type most frequently activated at	Time Band	Aggregation
US Card Reader Type	is card reader type of card reader	US Card Reader	Basic
US Photocopier	is most active photocopier with average daily number of pages photocopied	Number of Pages	Aggregation
US Photocopier	is most expensive photocopier with average daily page cost per document	Page Cost	Aggregation
US Residence	is residence where washing equipment was used by	Person	Basic
US Residence	is residence where washing equipment was used at	Time Band	Basic
US Vehicle License Plate	is license plate of vehicle taken from motor pool through activation of card reader	US Card Reader	Basic
US Washing Equipment	is washing equipment used on	Date	Basic
US Washing Equipment	is washing equipment used by	Person	Basic
US Washing Equipment	is washing equipment used at	Time Band	Basic
US Washing Equipment	is washing equipment activated by card reader	US Card Reader	Basic
US Washing Equipment	is washing equipment used at residence	US Residence	Basic

A2 - Demographic and study programme information

Entity	Relation	Entity	Relation Type
Age	is age of oldest person with gender	Gender	Aggregation
Age	is age of youngest person with gender	Gender	Aggregation
Age	is average age of persons with gender	Gender	Aggregation
Age	is age of oldest person with handicap	Handicap	Aggregation
Age	is age of person with handicap	Handicap	Basic
Age	is age of youngest person with handicap	Handicap	Aggregation
Age	is average age of persons with handicap	Handicap	Aggregation
Age	is age of person who attended high school	High School	Basic
Age	is age of oldest person with marital status	Marital Status	Aggregation
Age	is age of youngest person with marital status	Marital Status	Aggregation
Age	is average age of persons with marital status	Marital Status	Aggregation
Age	is age of oldest person with nationality	Nationality	Aggregation
Age	is age of youngest person with nationality	Nationality	Aggregation
Age	is average age of persons with nationality	Nationality	Aggregation
Age	is age of	Person	Basic
Age	is age of oldest person	Person	Aggregation
Age	is age of youngest person	Person	Aggregation
Age	is average age of persons and specifically	Person	Aggregation
Age	is age of oldest person within population group	Population Group	Aggregation
Age	is age of person that belongs to population group	Population Group	Basic
Age	is age of youngest person within population group	Population Group	Aggregation
Age	is average age of persons within population group	Population Group	Aggregation
Age	is age of person with home town	Town	Basic
Age	is age of person attending class at	US Campus	Basic
Age	is age of oldest person taking module at department	US Department	Aggregation
Age	is age of youngest person taking module at department	US Department	Aggregation
Age	is average age of persons taking modules at department	US Department	Aggregation
Age	is age of oldest person enrolled with faculty	US Faculty	Aggregation
Age	is age of person enrolled with faculty	US Faculty	Basic
Age	is age of youngest person enrolled with faculty	US Faculty	Aggregation
Age	is average age of persons enrolled with faculty	US Faculty	Aggregation
Age	is age of oldest person taking module	US Module	Aggregation
Age	is age of youngest person taking module	US Module	Aggregation
Age	is average age of persons taking module	US Module	Aggregation
Age	is age of oldest person studying programme	US Programme	Aggregation
Age	is age of youngest person studying programme	US Programme	Aggregation
Age	is average age of persons studying programme	US Programme	Aggregation
Age	is age of person studying programme	US Programme Continuity	Basic
Age	is age of person studying towards qualification	US Qualification	Basic
Age	is age of oldest person staying in residence type	US Residence Type	Aggregation
Age	is age of youngest person staying in residence type	US Residence Type	Aggregation
Age	is average age of persons staying in residence type	US Residence Type	Aggregation
Education Department	is education department of person with age	Age	Basic
Education Department	is education department of person with gender	Gender	Basic
Education Department	is education department of person with handicap	Handicap	Basic
Education Department	is education department with highest occurrence of handicap	Handicap	Aggregation
Education Department	is education department of person who attended high school	High School	Basic
Education Department	is education department of person speaking	Language	Basic
Education Department	is education department of person with marital status	Marital Status	Basic
Education Department	is education department of person with nationality	Nationality	Basic
Education Department	is education department of person that belongs to population group	Population Group	Basic
Education Department	is education department of person with postal code	Postal Code	Basic

A2 - Demographic and study programme information (cont.)

Entity	Relation	Entity	Relation Type
Education Department	is education department of person studying	US Programme	Basic
Education Department	is education department of person studying programme	US Programme Continuity	Basic
Education Department	is education department of person studying towards qualification	US Qualification	Basic
Education Department	is education department of person staying in	US Residence Type	Basic
Gender	is gender of person with age	Age	Basic
Gender	is gender of person with handicap	Handicap	Basic
Gender	is gender of person speaking	Language	Basic
Gender	is gender of person with marital status	Marital Status	Basic
Gender	is gender of person with nationality	Nationality	Basic
Gender	is gender of person that belongs to population group	Population Group	Basic
Gender	is gender of person with home town	Town	Basic
Gender	is gender of person attending class at	US Campus	Basic
Gender	is gender of person enrolled with faculty	US Faculty	Basic
Gender	is gender of person taking module	US Module	Basic
Gender	is gender of person studying	US Programme	Basic
Gender	is gender of person studying programme	US Programme Continuity	Basic
Gender	is gender of person studying towards qualification	US Qualification	Basic
Handicap	is handicap most common in education department	Education Department	Aggregation
Handicap	is handicap of person speaking	Language	Basic
Handicap	is handicap of person with	Marital Status	Basic
Handicap	is handicap of	Person	Basic
Handicap	is handicap of person with home town	Town	Basic
Handicap	is handicap of person attending class at	US Campus	Basic
Handicap	is handicap of person enrolled with faculty	US Faculty	Basic
Handicap	is handicap of person taking	US Module	Basic
Handicap	is handicap of person studying	US Programme	Basic
Handicap	is handicap of person studying programme	US Programme Continuity	Basic
Handicap	is handicap of person studying towards qualification	US Qualification	Basic
Handicap	is handicap of person staying in residence type	US Residence Type	Basic
High School	is high school attended by person with gender	Gender	Basic
High School	is high school attended by person with handicap	Handicap	Basic
High School	is high school attended by person with nationality	Nationality	Basic
High School	is high school attended by	Person	Basic
High School	is high school attended by person with postal code	Postal Code	Basic
High School	is high school attended by person with home town	Town	Basic
High School	is high school attended by person attending class at	US Campus	Basic
High School	is high school attended by person enrolled with faculty	US Faculty	Basic
High School	is high school attended by person studying programme	US Programme Continuity	Basic
High School	is high school attended by person studying towards qualification	US Qualification	Basic
Language	is language of person with age	Age	Basic
Language	is language spoken by most persons with handicap	Handicap	Aggregation
Language	is language of person who attended high school	High School	Basic
Language	is language of person with marital status	Marital Status	Basic
Language	is language of person with nationality	Nationality	Basic
Language	is language of	Person	Basic
Language	is language of person that belongs to population group	Population Group	Basic
Language	is language of person with postal code	Postal Code	Basic
Language	is language of person with home town	Town	Basic
Language	is language of person attending class at	US Campus	Basic
Language	is language of person enrolled with faculty	US Faculty	Basic
Language	is language of person studying programme	US Programme Continuity	Basic
Language	is language of person studying towards qualification	US Qualification	Basic
Language	is language most common in residence type	US Residence Type	Aggregation
Marital Status	is marital status of person with age	Age	Basic
Marital Status	is marital status of person who attended high school	High School	Basic
Marital Status	is marital status most common in people speaking	Language	Aggregation
Marital Status	is marital status of person with nationality	Nationality	Basic
Marital Status	is marital status of most people in population group	Population Group	Aggregation
Marital Status	is marital status of person with postal code	Postal Code	Basic
Marital Status	is marital status of person attending class at	US Campus	Basic
Marital Status	is marital status of person studying	US Programme	Basic
Marital Status	is marital status of person studying programme	US Programme Continuity	Basic
Marital Status	is marital status of person studying towards qualification	US Qualification	Basic
Marital Status	is marital status of person staying in residence type	US Residence Type	Basic
Nationality	is nationality of person with age	Age	Basic
Nationality	is nationality of most persons with handicap	Handicap	Aggregation
Nationality	is nationality of person with handicap	Handicap	Basic
Nationality	is nationality of person with home town	Town	Basic
Nationality	is nationality of person attending class at	US Campus	Basic
Nationality	is nationality of person studying programme	US Programme Continuity	Basic
Nationality	is nationality of person studying towards qualification	US Qualification	Basic
Person	Is person that was born on	Date	Basic
Person	is person that received senior certificate from education department	Education Department	Basic
Person	is person with gender	Gender	Basic
Person	is person with marital status	Marital Status	Basic
Person	is person with nationality	Nationality	Basic
Person	is person belonging to population group	Population Group	Basic
Person	is person taking module at department	US Department	Basic
Person	is person enrolled with faculty	US Faculty	Basic
Person	is person taking module	US Module	Basic
Person	is person studying programme	US Programme Continuity	Basic
Person	is person studying towards qualification	US Qualification	Basic
Person	is person staying in residence type	US Residence Type	Basic
Population Group	is population group of most persons with handicap	Handicap	Aggregation
Population Group	is population group of person with handicap	Handicap	Basic
Population Group	is population group of person who attended high school	High School	Basic
Population Group	is largest population group in group of persons with nationality	Nationality	Aggregation
Population Group	is population group of person with nationality	Nationality	Basic
Population Group	is population group of person with home town	Town	Basic
Population Group	is population group of person taking module	US Module	Basic
Population Group	is population group of person studying towards qualification	US Qualification	Basic
Postal Code	is postal code of person with age	Age	Basic
Postal Code	is postal code of person with gender	Gender	Basic
Postal Code	is postal code of person with handicap	Handicap	Basic
Postal Code	is postal code of person with nationality	Nationality	Basic
Postal Code	is postal code of	Person	Basic
Postal Code	is postal code of person belonging to population group	Population Group	Basic
Postal Code	is postal code associated with town	Town	Basic
Postal Code	is postal code of person enrolled with faculty	US Faculty	Basic
Postal Code	is postal code of person taking module	US Module	Basic
Postal Code	is postal code of person studying	US Programme	Basic
Postal Code	is postal code of person staying in residence type	US Residence Type	Basic
Town	is home town of person with marital status	Marital Status	Basic
Town	is home town of	Person	Basic
Town	is home town of person attending class at	US Campus	Basic
Town	is town of person taking module at department	US Department	Basic
Town	is home town of person studying programme	US Programme Continuity	Basic
Town	is home town of person studying towards qualification	US Qualification	Basic
Town	is home town of person staying in residence type	US Residence Type	Basic
US Campus	is campus where class is attended by	Person	Basic
US Campus	is campus where class is attended by person belonging to population group	Population Group	Basic
US Campus	is campus where class is attended by person with postal code	Postal Code	Basic
US Campus	is campus where class is attended by person taking module	US Module	Basic
US Campus	is campus where class is attended by person studying	US Programme	Basic
US Campus	is campus where class is attended by person studying towards qualification	US Qualification	Basic
US Department	is department of person with age	Age	Basic
US Department	is department most favoured by education department	Education Department	Aggregation
US Department	is department of person with education department	Education Department	Basic
US Department	is department of person with nationality	Nationality	Basic
US Department	is most favoured department of persons with nationality	Nationality	Aggregation
US Department	is department favoured by persons from population group	Population Group	Aggregation
US Department	is department of person belonging to population group	Population Group	Basic

A2 - Demographic and study programme information (cont.)

Entity	Relation	Entity	Relation Type
US Faculty	is faculty most favoured by education department	Education Department	Aggregation
US Faculty	is faculty most favoured by person with handicap	Handicap	Aggregation
US Faculty	is faculty of person with nationality	Nationality	Basic
US Faculty	is most favoured faculty of persons with nationality	Nationality	Aggregation
US Faculty	is faculty favoured by persons from population group	Population Group	Aggregation
US Faculty	is faculty of person belonging to population group	Population Group	Basic
US Faculty	is faculty favoured by persons with home town	Town	Aggregation
US Faculty	is faculty of person with home town	Town	Basic
US Faculty	is faculty of person attending class at	US Campus	Basic
US Faculty	is faculty of person taking module	US Module	Basic
US Faculty	is faculty with favoured module	US Module	Aggregation
US Faculty	is faculty with favoured programme	US Programme	Aggregation
US Faculty	is faculty of person studying towards qualification	US Qualification	Basic
US Faculty	is faculty of person staying in residence type	US Residence Type	Basic
US Faculty	is faculty of persons that favour residence type	US Residence Type	Aggregation
US Module	is module taken by person with age	Age	Basic
US Module	is module most favoured by education department	Education Department	Aggregation
US Module	is module most favoured by gender	Gender	Aggregation
US Module	is module taken by person who attended high school	High School	Basic
US Module	is module taken by person speaking	Language	Basic
US Module	is module taken by person with marital status	Marital Status	Basic
US Module	is module taken by person with nationality	Nationality	Basic
US Module	is most favoured module taken by persons with nationality	Nationality	Aggregation
US Module	is module favoured by persons from population group	Population Group	Aggregation
US Module	is module favoured by persons with home town	Town	Aggregation
US Module	is module taken by person with home town	Town	Basic
US Module	is module person takes at department	US Department	Basic
US Module	is module favoured by persons studying programme	US Programme	Aggregation
US Module	is module taken by person studying	US Programme	Basic
US Module	is module taken by person studying programme	US Programme Continuity	Basic
US Module	is module taken by person staying in residence type	US Residence Type	Basic
US Programme	is programme being studied by person with age	Age	Basic
US Programme	is programme most favoured by education department	Education Department	Aggregation
US Programme	is programme most favoured by person with handicap	Handicap	Aggregation
US Programme	is programme being studied by person who attended high school	High School	Basic
US Programme	is programme most favoured by person who attended high school	High School	Aggregation
US Programme	is programme being studied by person speaking	Language	Basic
US Programme	is programme being studied by person with marital status	Marital Status	Basic
US Programme	is most favoured programme studied by persons with nationality	Nationality	Aggregation
US Programme	is programme being studied by person with nationality	Nationality	Basic
US Programme	is programme being studied by	Person	Basic
US Programme	is programme favoured by persons from population group	Population Group	Aggregation
US Programme	is programme being studied by person with home town	Town	Basic
US Programme	is programme favoured by persons with home town	Town	Aggregation
US Programme	is programme being studied by person enrolled with faculty	US Faculty	Basic
US Programme	is programme being studied by person	US Programme Continuity	Basic
US Programme	is programme being studied by person for qualification	US Qualification	Basic
US Programme	is programme being studied by person staying in residence type	US Residence Type	Basic
US Programme Continuity	is manner in which qualification is acquired by person belonging to population group	Population Group	Basic
US Programme Continuity	is manner in which qualification is acquired by person with postal code	Postal Code	Basic
US Programme Continuity	is manner in which qualification is acquired by person enrolled with faculty	US Faculty	Basic
US Programme Continuity	is manner in which qualification is acquired by person studying programme	US Programme Continuity	Basic
US Qualification	is qualification being studied towards by person belonging to population group	Population Group	Basic
US Qualification	is qualification favoured by persons from population group	Population Group	Aggregation
US Qualification	is qualification being studied towards by person with postal code	Postal Code	Basic
US Qualification	is qualification being studied towards by person taking module at department	US Department	Basic
US Qualification	is qualification being studied towards by person taking module	US Module	Basic
US Qualification	is qualification being studied towards by person studying programme	US Programme Continuity	Basic
US Qualification	is qualification mostly studied with programme continuity	US Programme Continuity	Aggregation
US Residence Type	is residence type inhabited by person with age	Age	Basic
US Residence Type	is residence type inhabited by person with gender	Gender	Basic
US Residence Type	is residence type most favoured by person with handicap	Handicap	Aggregation
US Residence Type	is residence type inhabited by person who attended	High School	Basic
US Residence Type	is residence type inhabited by person speaking	Language	Basic
US Residence Type	is most favoured residence type of persons with nationality	Nationality	Aggregation
US Residence Type	is residence type inhabited by person with nationality	Nationality	Basic
US Residence Type	is residence type favoured by persons from population group	Population Group	Aggregation
US Residence Type	is residence type favoured by persons with home town	Town	Aggregation
US Residence Type	is residence type inhabited by person attending class at	US Campus	Basic
US Residence Type	is residence type favoured by persons studying programme	US Programme	Aggregation
US Residence Type	is residence type inhabited by person studying programme	US Programme Continuity	Basic
US Residence Type	Is residence type favoured by persons pursuing qualification	US Qualification	Aggregation
US Residence Type	is residence type inhabited by person studying towards qualification	US Qualification	Basic

A3 - Internet usage cost information

Entity	Relation	Entity	Relation Type
Network Traffic	is internet usage on	Date	Basic
Network Traffic	is internet usage of	Person	Basic
Person	is person who accessed internet on	Date	Basic
Person	is heaviest internet user with average daily internet usage of	Network Traffic	Aggregation
Person	is person performing on daily average the most internet transactions at	Time Band	Aggregation
Person	is person who accessed internet from workstation/PC	US Workstation/PC	Basic
Person	is person who performs most internet transactions at workstation	US Workstation/PC	Aggregation
Time Band	is time at which internet was accessed on	Date	Basic
Time Band	is time band in which internet usage was	Network Traffic	Basic
Time Band	is time at which internet was accessed by	Person	Basic
Time Band	is time at which internet was accessed from workstation/PC	US Workstation/PC	Basic
US Workstation/PC	is workstation/PC used to access the internet on	Date	Basic
US Workstation/PC	is workstation/PC used for internet usage of	Network Traffic	Basic

A4 - Library transaction and book information

Entity	Relation	Entity	Relation Type
Date	is return date of library item which is after due date on	Date	Basic
Date	is date on which library transaction was performed on library item written in	Language	Basic
Date	is date on which library transaction was performed on library item at	Library Branch	Basic
Date	is date on which library transaction was performed on library item authored by	Library Item Author	Basic
Date	is date on which library transaction was performed on library item belonging to collection	Library Item Collection	Basic
Date	is date on which library transaction was performed on library item with loan amount	Library Item Loan Amount	Basic
Date	is date on which library transaction was performed on library item published by	Library Item Publisher	Basic
Date	is date on which library transaction was performed on library item with subject topic	Library Item Subject Topic	Basic
Date	is date on which library transaction was performed on library item	Library Item Title	Basic
Date	is due date of library item	Library Item Title	Basic
Date	is last returned date of library item	Library Item Title	Basic
Date	is return date of library item	Library Item Title	Basic
Date	is start date of library item	Library Item Title	Basic
Date	is update date of library item	Library Item Title	Basic
Date	is date on which library transaction was performed on library item provided by	Library Item Vendor	Basic
Date	is date on which library transaction was performed on library item	Library Shelf Code	Basic
Date	is due date of library item	Library Shelf Code	Basic
Date	is last returned date of library item	Library Shelf Code	Basic
Date	is return date of library item	Library Shelf Code	Basic
Date	is start date of library item	Library Shelf Code	Basic
Date	is update date of library item	Library Shelf Code	Basic
Date	is date on which library transaction was performed on library item with shelf code type	Library Shelf Code Type	Basic
Date	is date on which library transaction was performed at location	Library Transaction Location	Basic
Date	is date on which library transaction was performed by patron type	Library Transaction Patron Type	Basic
Date	is date of highest database usage corresponding to network traffic	Network Traffic	Aggregation
Date	is date of lowest database usage corresponding to network traffic	Network Traffic	Aggregation
Date	is date on which library item on loan is/was due by	Person	Basic

A4 - Library transaction and book information (cont.)

Entity	Relation	Entity	Relation Type
Language	is language of library item with Dewey classification	Library Dewey Classification	Basic
Language	is language of library item from collection	Library Item Collection	Basic
Language	is language of library item with loan amount	Library Item Loan Amount	Basic
Language	is language of library item with transaction status	Library Item Status	Basic
Language	is language of library item with subject topic	Library Item Subject Topic	Basic
Language	is language of library item	Library Item Title	Basic
Language	is language of library item with shelf code type	Library Shelf Code Type	Basic
Language	is language of library item published in	Town	Basic
Library Branch	is library branch where library items are kept with shelf code type	Library Shelf Code Type	Basic
Library Branch	is library branch situated on campus	US Campus	Basic
Library Database Name	is library database that had network activity on	Date	Basic
Library Dewey Classification	is library Dewey classification of library item published by	Library Item Publisher	Basic
Library Dewey Classification	is library Dewey classification of library item	Library Item Title	Basic
Library Dewey Classification	is library Dewey classification of library item provided by	Library Item Vendor	Basic
Library Dewey Classification	is library Dewey classification of library item	Library Shelf Code	Basic
Library Dewey Classification	is library Dewey classification of library item with shelf code type	Library Shelf Code Type	Basic
Library Dewey Classification	is library Dewey classification of library item on which library transaction was performed at location	Library Transaction Location	Basic
Library Item Author	is author of library item written in	Language	Basic
Library Item Author	is author of library item kept at	Library Branch	Basic
Library Item Author	is author with most library items in library branch	Library Branch	Aggregation
Library Item Author	is author of library item with Dewey classification	Library Dewey Classification	Basic
Library Item Author	is author with a highest average library item loan amount	Library Item Loan Amount	Aggregation
Library Item Author	is author of library item published by	Library Item Publisher	Basic
Library Item Author	is author of library item with transaction status	Library Item Status	Basic
Library Item Author	is author of library item	Library Item Title	Basic
Library Item Author	is author of library item	Library Shelf Code	Basic
Library Item Author	is author of library item on which library transaction was performed at	Time Band	Basic
Library Item Author	is author of library item published in	Town	Basic
Library Item Collection	is library collection that contains item written in	Language	Basic
Library Item Collection	is library collection that contains item with Dewey classification	Library Dewey Classification	Basic
Library Item Collection	is library collection that contains item authored by	Library Item Author	Basic
Library Item Collection	is library collection that contains item published by	Library Item Publisher	Basic
Library Item Collection	is library collection that contains item	Library Item Shelf Code	Basic
Library Item Collection	is library collection that contains item with shelf code type	Library Item Shelf Code Type	Basic
Library Item Collection	is library collection that contains item with current status	Library Item Status	Basic
Library Item Collection	is library collection that contains item with subject topic	Library Item Subject topic	Basic
Library Item Collection	is library collection that contains item	Library Item Title	Basic
Library Item Collection	is library collection that contains item provided by	Library Item Vendor	Basic
Library Item Loan Amount	is loan amount of library item authored by	Library Item Author	Basic
Library Item Loan Amount	is loan amount of library item published by	Library Item Publisher	Basic
Library Item Loan Amount	is loan amount of library item	Library Item Shelf Code	Basic
Library Item Loan Amount	is loan amount of library item with shelf code type	Library Item Shelf Code Type	Basic
Library Item Loan Amount	is loan amount of library item with current status	Library Item Status	Basic
Library Item Loan Amount	is loan amount of library item	Library Item Title	Basic
Library Item Material	is material type of library item on which library transaction was performed on	Date	Basic
Library Item Material	is material type of library item written in	Language	Basic
Library Item Material	is material type of library item kept at	Library Branch	Basic
Library Item Material	is material type of library item with Dewey classification	Library Dewey Classification	Basic
Library Item Material	is material of library item published by	Library Item Publisher	Basic
Library Item Material	is material type of library item with current status	Library Item Status	Basic
Library Item Material	is material type of library item on which library transaction was performed at	Library Transaction Location	Basic
Library Item Material	is material type of library item on which library transaction was performed by patron type	Library Transaction Patron Type	Basic
Library Item Material	is material type of library item on which library transaction was performed of type	Library Transaction Type	Basic
Library Item Material	is material type of library item on which library transaction was performed by	Person	Basic
Library Item Material	is material type of library item on which library transaction was performed at	Time Band	Basic
Library Item Material	is material type of library item published in	Town	Basic
Library Item Publisher	is publisher of library item published on	Date	Basic
Library Item Publisher	is publisher of library item written in	Language	Basic
Library Item Publisher	is publisher with a highest average library item loan amount	Library Item Loan Amount	Aggregation
Library Item Publisher	is publisher of library item with transaction status	Library Item Status	Basic
Library Item Publisher	is favoured publisher for subject topic	Library Item Subject Topic	Aggregation
Library Item Publisher	is publisher of library item on which library transaction was performed by	Person	Basic
Library Item Status	is item status of library item on which library transaction was performed on	Date	Basic
Library Item Status	is transaction status of library item on which library transaction was performed on	Date	Basic
Library Item Status	is current status of library item written in	Language	Basic
Library Item Status	is current status of library item authored by	Library Item Author	Basic
Library Item Status	is transaction status of library item that belongs to collection	Library Item Collection	Basic
Library Item Status	is transaction status of library item with loan amount	Library Item Loan Amount	Basic
Library Item Status	is transaction status of library item with material type	Library Item Material	Basic
Library Item Status	is current status of library item published by	Library Item Publisher	Basic
Library Item Status	is transaction status of library item with subject topic	Library Item Subject Topic	Basic
Library Item Status	Is current status of library item	Library Item Title	Basic
Library Item Status	is transaction status of library item provided by	Library Item Vendor	Basic
Library Item Status	is current status of library item	Library Shelf Code	Basic
Library Item Status	is current status of library item with shelf code type	Library Shelf Code Type	Basic
Library Item Status	is status of library item because of transaction performed at location	Library Transaction Location	Basic
Library Item Status	is transaction status of library transaction performed by patron type	Library Transaction Patron Type	Basic
Library Item Status	is current status of library item on which transaction was performed of type	Library Transaction Type	Basic
Library Item Status	is item status of library item on which library transaction was performed by	Person	Basic
Library Item Status	is transaction status of library item on which library transaction was performed by	Person	Basic
Library Item Status	is current status of library item on which library transaction was performed at	Time Band	Basic
Library Item Status	is transaction status of library item on which library transaction was performed at	Time Band	Basic
Library Item Status	is current status of library item published in	Town	Basic
Library Item Subject Topic	is subject topic of library item with Dewey classification	Library Dewey Classification	Basic
Library Item Subject Topic	is subject topic of library item authored by	Library Item Author	Basic
Library Item Subject Topic	is subject topic of library item with loan amount	Library Item Loan Amount	Basic
Library Item Subject Topic	is subject topic with a highest average library item loan amount	Library Item Loan Amount	Aggregation
Library Item Subject Topic	is subject of library item published by	Library Item Publisher	Basic
Library Item Subject Topic	is subject topic of library item with current status	Library Item Status	Basic
Library Item Subject Topic	is subject topic of library item provided by	Library Item Vendor	Basic
Library Item Subject Topic	is subject topic with library items mostly provided by vendor	Library Item Vendor	Aggregation
Library Item Subject Topic	is subject topic of library item on which library transaction was performed by	Person	Basic
Library Item Subject Topic	is subject topic of library item on which library transaction was performed at	Time Band	Basic
Library Item Title	is title of library item published on	Date	Basic
Library Item Title	is title of library item kept at	Library Branch	Basic
Library Item Title	is title of library associated with database	Library Database Name	Basic
Library Item Title	is title of library item with material type	Library Item Material	Basic
Library Item Title	is title of library item published by	Library Item Publisher	Basic
Library Item Title	is title of library item with shelf code type	Library Item Shelf Code Type	Basic
Library Item Title	is title of library item with transaction status	Library Item Status	Basic
Library Item Title	is title of library item with subject topic	Library Item Subject topic	Basic
Library Item Title	is title of library item on which library transaction was performed by	Person	Basic
Library Item Title	is title of library item with due time	Time band	Basic
Library Item Title	is title of library item with return time	Time band	Basic
Library Item Title	is title of library item published in	Town	Basic
Library Item Vendor	is provider of library item written in	Language	Basic
Library Item Vendor	is provider of library item authored by	Library Item Author	Basic
Library Item Vendor	is provider of library item published by	Library Item Publisher	Basic
Library Item Vendor	is provider of library item with current status	Library Item Status	Basic
Library Item Vendor	is provider of library item	Library Item Title	Basic
Library Item Vendor	is provider of library item	Library Shelf Code	Basic
Library Item Vendor	is provider of library item on which library transaction was performed at	Library Transaction Location	Basic
Library Item Vendor	is provider of library item on which library transaction was performed by patron type	Library Transaction Patron Type	Basic
Library Item Vendor	is provider of library item on which library transaction was performed of type	Library Transaction Type	Basic
Library Item Vendor	is provider of library item on which library transaction was performed by	Person	Basic
Library Item Vendor	is provider of library item published in	Town	Basic
Library Shelf Code	is shelf code of library item published on	Date	Basic
Library Shelf Code	is shelf code of library item written in	Language	Basic
Library Shelf Code	is shelf code of library item kept at	Library Branch	Basic
Library Shelf Code	is shelf code of library item with material type	Library Item Material	Basic
Library Shelf Code	is shelf code of library item published by	Library Item Publisher	Basic
Library Shelf Code	is shelf code of library item with transaction status	Library Item Status	Basic

A4 - Library transaction and book information (cont.)

Entity	Relation	Entity	Relation Type
Library Shelf Code	is shelf code of library item on which library transaction was performed by patron type	Library Transaction Patron Type	Basic
Library Shelf Code	is shelf code of library item on which library transaction was performed of type	Library Transaction Type	Basic
Library Shelf Code	is shelf code of library item with due time	Time band	Basic
Library Shelf Code	is shelf code of library item with return time	Time band	Basic
Library Shelf Code	is shelf code of library item published in	Town	Basic
Library Shelf Code Type	is shelf code type of library item with material type	Library Item Material	Basic
Library Shelf Code Type	is shelf code type of library item with transaction status	Library Item Status	Basic
Library Shelf Code Type	is shelf code type of library item with subject topic	Library Item Subject Topic	Basic
Library Shelf Code Type	is shelf code type of library item	Library Shelf Code	Basic
Library Transaction Location	is location where library transaction was performed on library item written in	Language	Basic
Library Transaction Location	is location where library transaction was performed in	Library Branch	Basic
Library Transaction Location	is location where library transaction was performed on library item published by	Library Item Publisher	Basic
Library Transaction Location	is location where library transaction was performed on library item	Library Item Title	Basic
Library Transaction Patron Type	is patron type that performed library transaction on library item written in	Language	Basic
Library Transaction Patron Type	is patron type who performed library transaction at	Library Branch	Basic
Library Transaction Patron Type	is patron type who performed library transaction on library item with Dewey Classification	Library Dewey Classification	Basic
Library Transaction Patron Type	is patron type that performed library transaction on library item authored by	Library Item Author	Basic
Library Transaction Patron Type	is patron type who performed library transaction on library item belonging to collection	Library Item Collection	Basic
Library Transaction Patron Type	is patron type who performed library transaction on library item published by	Library Item Publisher	Basic
Library Transaction Patron Type	is patron type who performed library transaction on library item with current status	Library Item Status	Basic
Library Transaction Patron Type	is patron type who performed library transaction on library item with subject topic	Library Item Subject Topic	Basic
Library Transaction Patron Type	is patron type who performed library transaction on library item	Library Item Title	Basic
Library Transaction Patron Type	is patron type that performed library transaction on library item at location	Library Transaction Location	Basic
Library Transaction Patron Type	is patron type who performed library transaction on library item at	Time Band	Basic
Library Transaction Type	is library transaction type performed by patron type	Library Transaction Patron Type	Basic
Library Transaction Type	is library transaction type performed at	Library Transaction Type	Basic
Library Transaction Type	is transaction type performed on library item by	Person	Basic
Library Transaction Type	is transaction type performed on	Date	Basic
Library Transaction Type	is library transaction type performed on library item written in	Language	Basic
Library Transaction Type	is transaction type performed at	Library Branch	Basic
Library Transaction Type	is transaction type performed on library item with Dewey Classification	Library Dewey Classification	Basic
Library Transaction Type	is library transaction type performed on library item authored by	Library Item Author	Basic
Library Transaction Type	is library transaction type performed on library item belonging to collection	Library Item Collection	Basic
Library Transaction Type	is library transaction type performed on library item published by	Library Item Publisher	Basic
Library Transaction Type	is library transaction type performed on library item with transaction status	Library Item Status	Basic
Library Transaction Type	is library transaction type performed on library item with subject topic	Library Item Subject Topic	Basic
Library Transaction Type	is transaction type performed on library item	Library Item Title	Basic
Library Transaction Type	is transaction type performed at	Time Band	Basic
Library Transaction Type	is transaction type performed at location	Time Band	Basic
Library Transaction Type	is library transaction type performed on library item published in	Town	Basic
Network Traffic	is database traffic on	Date	Basic
Network Traffic	is on-line magazine traffic on	Date	Basic
Network Traffic	is network traffic associated with library item written in	Language	Basic
Network Traffic	is daily network traffic of	Library Database Name	Basic
Network Traffic	is daily network traffic associated with electronic magazine	Library Item Title	Basic
Person	is person who performed library transaction on library item written in language	Language	Basic
Person	is person who performed library transaction at	Library Branch	Basic
Person	is person who performed library transaction on library item with Dewey classification	Library Dewey Classification	Basic
Person	is person who performed library transaction on library item authored by	Library Item Author	Basic
Person	is person who performed library transaction on library item belonging to collection	Library Item Collection	Basic
Person	is person who performed library transaction on library item with loan amount	Library Item Loan Amount	Basic
Person	is person who performed library transaction on library item	Library Item Shelf Code	Basic
Person	is person who performed library transaction on library item with shelf code type	Library Item Shelf Code Type	Basic
Person	is person who performed library transaction at location	Library Transaction Location	Basic
Person	is person who performed library transaction on library item as patron type	Library Transaction Patron Type	Basic
Person	is person performing on daily average the most library transactions at	Time Band	Aggregation
Person	is person who has/had library items due at	Time Band	Basic
Person	is person who performed library transaction at	Time Band	Basic
Person	is person who returned library item at	Time Band	Basic
Time Band	is time at which library item were due on	Date	Basic
Time Band	is time at which library item were returned on	Date	Basic
Time Band	is time at which library transaction was performed on	Date	Basic
Time Band	is time at which library transaction was performed on library item written in	Language	Basic
Time Band	is time at which library transaction was performed at	Library Branch	Basic
Time Band	is time at which library transaction was performed on library item belonging to collection	Library Item Collection	Basic
Time Band	is time at which library transaction was performed on library item with loan amount	Library Item Loan Amount	Basic
Time Band	is time at which library transaction was performed on library item published by	Library Item Publisher	Basic
Time Band	is last returned time of library item	Library Item Shelf Code	Basic
Time Band	is time at which library transaction was performed on library item	Library Item Shelf Code	Basic
Time Band	is time at which library transaction was performed on library item with shelf code type	Library Item Shelf Code Type	Basic
Time Band	is last returned time of library item	Library Item Title	Basic
Time Band	is time at which library transaction was performed on library item	Library Item Title	Basic
Time Band	is time at which library transaction was performed on library item provided by	Library Item Vendor	Basic
Time Band	is time at which library transaction was performed on library item published in	Town	Basic
Town	is city in which library item was published on	Date	Basic
Town	is city in which library item was published which is now kept at	Library Branch	Basic
Town	is city in which library item was published belonging to collection	Library Item Collection	Basic
Town	is city in which library item was published with loan amount	Library Item Loan Amount	Basic
Town	is city in which library item was published on which library transaction was performed by patron type	Library Item Patron Type	Basic
Town	is city in which library item was published by	Library Item Publisher	Basic
Town	is city in which library item was published on with shelf code type	Library Item Shelf Code Type	Basic
Town	is city in which library item was published with transaction status	Library Item Status	Basic
Town	is city in which library item was published with subject topic	Library Item Subject Topic	Basic
US Department	is department subscribed to on-line magazine	Library Item Title	Basic

A5 - Printing transaction information

Entity	Relation	Entity	Relation Type
Date	is date on which file was printed with number of pages	Number of Pages	Basic
Date	is date on which file was printed with page cost	Page Cost	Basic
Date	is date on which file was printed by	Person	Basic
Date	is date on which file was printed at	Time Band	Basic
Document Type	is document type printed on	Date	Basic
Document Type	is document type printed with file size	File Size	Basic
Document Type	is document type printed with number of pages	Number of Pages	Basic
Document Type	is document type printed with page cost	Page Cost	Basic
Document Type	is document type printed by	Person	Basic
Document Type	is document type printed at	Time band	Basic
Document Type	is document type printed with printer	US Printer	Basic
Document Type	is document type printed from workstation/PC	US Workstation/PC	Basic
File Size	is size of file printed on	Date	Basic
File Size	is file size of largest document printed with document type	Document Type	Aggregation
File Size	is file size of smallest document printed with document type	Document Type	Aggregation
File Size	is file size of largest document printed by printer	US Printer	Aggregation
File Size	is size of file printed with	US Printer	Basic
File Size	is size of file printed from workstation/PC	US Workstation/PC	Basic
Number of Pages	is average number of pages printed of document type	Document Type	Aggregation
Number of Pages	is largest number of pages printed of document type	Document Type	Aggregation
Number of Pages	is least number of pages printed of document type	Document Type	Aggregation
Number of Pages	is number of pages that was printed with file size	File Size	Basic
Number of Pages	is number of pages printed at page cost	Page Cost	Basic
Number of Pages	is number of pages printed from workstation/PC	US Workstation/PC	Basic
Page Cost	is average page cost of document printed with document type	Document Type	Aggregation
Page Cost	is highest page cost of document printed with document type	Document Type	Aggregation
Page Cost	is lowest page cost of document printed with document type	Document Type	Aggregation
Page Cost	is page cost of file printed with file size	File Size	Basic
Page Cost	is page cost of print job incurred from workstation/PC	US Workstation/PC	Basic
Person	is person who prints least documents of document type	Document Type	Aggregation
Person	is person who prints mostly documents of document type	Document Type	Aggregation
Person	is person that printed file with file size	File Size	Basic
Person	is person that printed file with number of pages	Number of Pages	Basic
Person	is person who prints the most with average daily number of pages printed	Number of Pages	Aggregation

A5 - Printing transaction information (cont.)

Entity	Relation	Entity	Relation Type
Time Band	is time at which file was printed with file size	File Size	Basic
Time Band	is time at which file was printed with number of pages	Number of Pages	Basic
Time Band	is time at which file was printed with page cost	Page Cost	Basic
Time Band	is time at which file was printed by	Person	Basic
Time Band	is time at which print job was processed by printer	US Printer	Basic
US Printer	is most active printer with per person average daily number of pages printed	Number of Pages	Aggregation
US Printer	is printer that processed print job on	Date	Basic
US Printer	is printer used to print mostly	Document Type	Aggregation
US Printer	is printer that printed largest document with file size	File Size	Aggregation
US Printer	is printer that processed print job with number of pages	Number of Pages	Basic
US Printer	is most expensive printer with average daily page cost per document	Page Cost	Aggregation
US Printer	is printer that processed print job with page cost	Page Cost	Basic
US Printer	is printer used by	Person	Basic
US Printer	is printer that was sent a print job from workstation/PC	US Workstation/PC	Basic
US Workstation/PC	is workstation/PC used to print document on	Date	Basic
US Workstation/PC	is workstation/PC that prints most of document type	Document Type	Aggregation
US Workstation/PC	is workstation/PC that prints mostly documents of document type	Document Type	Aggregation
US Workstation/PC	is workstation/PC used for printing by	Person	Basic
US Workstation/PC	is workstation/PC used to print document at	Time Band	Basic
US Workstation/PC	is workstation/PC that sends most print jobs to printer	US Printer	Aggregation

A6 - WebCT usage information

Entity	Relation	Entity	Relation Type
Date	is end date of WebCT learning context that started on	Date	Basic
Date	is start date of WebCT learning context ending	Date	Basic
Date	is start date of WebCT learning context ending on	Date	Basic
Date	is date on which WebCT action was performed by person with role	WebCT Event Role	Basic
Date	is start date of WebCT learning context with administrative period	WebCT Learning Context Administrative Period	Basic
Date	is date on which WebCT learning context was accessed by user with context role	WebCT Learning Context Role	Basic
Person	is person that performed action in WebCT on	Date	Basic
Person	is person performing on daily average the most transactions in WebCT at	Time Band	Aggregation
Person	is person with the highest average WebCT dwell time	WebCT Dwell Time	Aggregation
Person	is person with the lowest average WebCT dwell time	WebCT Dwell Time	Aggregation
Person	is person who performed action in WebCT with event role	WebCT Event Role	Basic
Person	is person with access to WebCT learning context	WebCT Learning Context	Basic
Person	is person with WebCT learning context role	WebCT Learning Context Role	Basic
Person	is person with access to WebCT template	WebCT Template	Basic
Time band	is time at which WebCT action was performed on	Date	Basic
Time Band	is time at which action in WebCT was performed by	Person	Basic
Time Band	is time favoured by users to perform WebCT Action	WebCT Action	Aggregation
Time Band	is time at which WebCT user has been dwelling for	WebCT Dwell Time	Basic
Time Band	is time favoured by users to access WebCT learning context	WebCT Learning Context	Aggregation
WebCT Action	is action performed in WebCT on	Date	Basic
WebCT Action	is action performed in WebCT by	Person	Basic
WebCT Action	is action performed in WebCT at	Time Band	Basic
WebCT Action	is action performed in WebCT after dwell time	WebCT Dwell Time	Basic
WebCT Action	is action performed in WebCT y user with event role	WebCT Event Role	Basic
WebCT Action	is action performed in WebCT learning context	WebCT Learning Context	Basic
WebCT Action	is action performed most in WebCT learning context	WebCT Learning Context	Aggregation
WebCT Action	is action performed most by persons WebCT learning context role	WebCT Learning Context Role	Aggregation
WebCT Action	is action performed in WebCT Template	WebCT Template	Basic
WebCT Dwell Time	is time dwelled by user in WebCT on	Date	Basic
WebCT Dwell Time	is time dwelled in WebCT by	Person	Basic
WebCT Dwell Time	is the average dwell time in WebCT learning context	WebCT Learning Context	Aggregation
WebCT Dwell Time	is the average dwell time for WebCT learning context role	WebCT Learning Context Role	Aggregation
WebCT Dwell Time	is time dwelled by user in WebCT learning context type	WebCT Learning Context Type	Basic
WebCT Event Role	is WebCT event role that performed action at	Time Band	Basic
WebCT Event Role	is WebCT event role that performed action in learning context	WebCT Learning Context	Basic
WebCT Learning Context	is WebCT learning context accessed on	Date	Basic
WebCT Learning Context	is WebCT learning context with end date	Date	Basic
WebCT Learning Context	is WebCT learning context with start date	Date	Basic
WebCT Learning Context	is WebCT learning context accessed at	Time Band	Basic
WebCT Learning Context	is WebCT learning context dwelled in for	WebCT Dwell Time	Basic
WebCT Learning Context	is WebCT learning context which is child of	WebCT Learning Context	Basic
WebCT Learning Context	is WebCT learning context which is parent of	WebCT Learning Context	Basic
WebCT Learning Context	is WebCT learning context with context type	WebCT Learning Context Type	Basic
WebCT Learning Context	is WebCT learning context owned by organisation	WebCT Organisation	Basic
WebCT Learning Context	is WebCT learning context owned by organisation type	WebCT Organisation Type	Basic
WebCT Learning Context	is WebCT learning context owned by organisation unit	WebCT Organisation Unit	Basic
WebCT Learning Context Administrative Period	is administrative period of WebCT learning context ending on	Date	Basic
WebCT Learning Context Administrative Period	is administrative period of WebCT learning context	WebCT Learning Context	Basic
WebCT Learning Context Role	is context role of user that accessed WebCT learning context at	Time Band	Basic
WebCT Learning Context Role	is context role that preformed WebCT action	WebCT Action	Basic
WebCT Learning Context Role	is context role that dwelled in WebCT for	WebCT Dwell Time	Basic
WebCT Learning Context Role	is context role associated with WebCT learning context	WebCT Learning Context	Basic
WebCT Learning Context Role	is context role associated with WebCT learning context type	WebCT Learning Context type	Basic
WebCT Learning Context Type	is WebCT learning context type accessed on	Date	Basic
WebCT Learning Context Type	is WebCT learning context type with end date	Date	Basic
WebCT Learning Context Type	is WebCT learning context type with start date	Date	Basic
WebCT Learning Context Type	is WebCT learning context type which can be accessed by	Person	Basic
WebCT Learning Context Type	is WebCT learning context type accessed at	Time Band	Basic
WebCT Learning Context Type	is WebCT learning context type on which was performed action	WebCT Action	Basic
WebCT Learning Context Type	is WebCT learning context type on which action was performed by event role	WebCT Event Role	Basic
WebCT Learning Context Type	is WebCT learning context type with administrative period	WebCT Learning Context Administrative Period	Basic
WebCT Number	is average daily number of occurrences of WebCT action	WebCT Action	Aggregation
WebCT Number	is number of children of WebCT learning context	WebCT Learning Context	Basic
WebCT Number	is number of members of WebCT learning context	WebCT Learning Context	Basic
WebCT Number	is number of templates in WebCT learning context	WebCT Learning Context	Basic
WebCT Number	is number of content pages found in WebCT template	WebCT Template	Basic
WebCT Number	is number of media libraries found in WebCT template	WebCT Template	Basic
WebCT Number	is number of organisers found in WebCT template	WebCT Template	Basic
WebCT Number	is number of questions found in WebCT template	WebCT Template	Basic
WebCT Number	is number of SCORM packages found in WebCT template	WebCT Template	Basic
WebCT Organisation	is organisation that owns WebCT learning context type	WebCT Learning Context Type	Basic
WebCT Organisation	is organisation that has an organisation type	WebCT Organisation Type	Basic
WebCT Organisation Type	is organisation type that owns WebCT learning context type	WebCT Learning Context Type	Basic
WebCT Organisation Unit	is organisation unit that owns WebCT learning context type	WebCT Learning Context Type	Basic
WebCT Organisation Unit	is organisation unit that forms part of organisation	WebCT Organisation	Basic
WebCT Organisation Unit	is organisation unit that forms part of organisation with type	WebCT Organisation Type	Basic
WebCT Size	is files size of WebCT learning context	WebCT Learning Context	Basic
WebCT Size	is templates size of WebCT learning context	WebCT Learning Context	Basic
WebCT Size	is total size of WebCT learning context	WebCT Learning Context	Basic
WebCT Size	is file size of WebCT template	WebCT Template	Basic
WebCT Template	**is WebCT template that was accessed on**	**Date**	**Basic**
WebCT Template	is WebCT template with dwell time	WebCT Dwell Time	Basic
WebCT Template	is WebCT template that is part of learning context	WebCT Learning Context	Basic
WebCT Template	is WebCT template with number of assessments	WebCT Number	Basic
WebCT Template	is WebCT template with number of chat rooms	WebCT Number	Basic
WebCT Template	is WebCT template with number of discussion topics	WebCT Number	Basic
WebCT Template	is WebCT template with number of learning modules	WebCT Number	Basic
WebCT Template	is WebCT template with number of proxy tools	WebCT Number	Basic

Appendix B – Inferred Relations

B1 - Card reader transaction information ←→ Internet usage cost information

Entity	Relation	Entity	Relation Type
Date	is date internet was accessed by person who also activated card reader	US Card Reader	Inferred
Date	is date internet was accessed by person who also activated card reader type	US Card Reader Type	Inferred
Date	is date on which the internet was accessed by person who also used printer	US Printer	Inferred
US Card Reader	is card reader activated by person who had internet usage of	Network Traffic	Inferred
US Photocopier	is photocopier that was used by a person who also accessed the internet on	Date	Inferred
US Photocopier	is photocopier that was used by a person who also performed library transaction on	Date	Inferred

B2 - Card reader transaction information ←→ Library transaction and book information

Entity	Relation	Entity	Relation Type
Date	is date library transaction was performed by person who also activated card reader	US Card Reader	Inferred
Date	is date library transaction was performed by person who also activated card reader type	US Card Reader Type	Inferred
Person	is person who performed library transaction and on the same date activated card reader	US Card Reader	Inferred

B3 - Card reader transaction information ←→ Printing transaction information

Entity	Relation	Entity	Relation Type
Date	is date file was printed by person who also made photocopies with	US Photocopier	Inferred
US Card Reader	is card reader activated by person who printed file on	Date	Inferred
US Card Reader Type	is card reader type activated by person who printed file on	Date	Inferred

B4 - Card reader transaction information ←→ WebCT usage information

Entity	Relation	Entity	Relation Type
Date	is date action was performed in WebCT by person who also activated card reader	US Card Reader	Inferred
Date	is date action was performed in WebCT by person who also activated card reader type	US Card Reader Type	Inferred
US Photocopier	is photocopier that was used by a person who also performed action in WebCT on	Date	Inferred

B5 - Demographic and study programme information ←→ Card reader transaction information

Entity	Relation	Entity	Relation Type
Education Department	is education department of person who activated card reader at	Time Band	Inferred
Education Department	is education department of persons responsible for the most card reader activity at	Time Band	Aggregation
Gender	is gender of most persons that took meal on	Date	Aggregation
Handicap	is handicap of person who activated card reader on	Date	Inferred
Handicap	is handicap of person who activated card reader	US Card Reader	Inferred
Handicap	is handicap of person who activated card reader type	US Card Reader Type	Inferred
High School	is high school attended by person who activated card reader on	Date	Inferred
High School	is high school of person who photocopied number of pages	Number of Pages	Inferred
High School	is high school of person who activated card reader at	Time Band	Inferred
Language	is language spoken by person who activated card reader on	Date	Inferred
Nationality	is nationality of person who photocopied number of pages	Number of Pages	Inferred
Nationality	is nationality of persons with highest average number of pages photocopied	Number of Pages	Aggregation
Nationality	is nationality of persons responsible for the most card reader activity at	Time Band	Aggregation
Nationality	is nationality of person who activated card reader	US Card Reader	Inferred
Nationality	is nationality of person who activated card reader type	US Card Reader Type	Inferred
Population Group	is population group who makes the most photocopies with per person average daily number of pages photocopied	Number of Pages	Aggregation
Population Group	is population group making most use of card reader	US Card Reader	Aggregation
Population Group	is population group of person who activated card reader type	US Card Reader Type	Inferred
Town	is town of person who activated card reader	US Card Reader	Inferred
Town	is town of person who activated card reader type	US Card Reader Type	Inferred
Town	is home town of person who used washing equipment	US Washing Equipment	Inferred
US Campus	is campus with card reader	US Card Reader	Inferred
US Campus	is campus with photocopier	US Photocopier	Inferred
US Campus	is campus with washing equipment	US Washing Equipment	Inferred
US Faculty	is faculty of person who activated card reader on	Date	Inferred
US Faculty	is faculty of person who photocopied number of pages	Number of Pages	Inferred
US Faculty	is faculty responsible for most photocopies with per person average daily number of pages photocopied	Number of Pages	Aggregation
US Faculty	is faculty which is the most expensive photocopier of documents with average page cost per document	Page Cost	Aggregation
US Faculty	is faculty of person who activated card reader at	Time Band	Inferred
US Faculty	is faculty of person who made photocopies with	US Photocopier	Inferred
US Module	is module taken by person who activated card reader on	Date	Inferred
US Programme	is programme studied by person who activated card reader on	Date	Inferred
US Programme	is programme with heaviest photocopying having per person average daily number of pages photocopied	Number of Pages	Aggregation
US Programme	is programme with heaviest photocopying having per person average daily page cost per document	Page Cost	Aggregation
US Programme	is programme studied by person who activated card reader	US Card Reader	Inferred
US Residence Type	is residence type inhabited by person who activated card reader at	Time Band	Inferred

B6 - Demographic and study programme information ←→ Internet usage cost information

Entity	Relation	Entity	Relation Type
Education Department	is education department of person who accessed the internet at	Time Band	Inferred
Education Department	is education department of persons responsible for the most internet activity at	Time Band	Aggregation
High School	is high school attended by person who had internet usage of	Network Traffic	Inferred
High School	is high school of person who accessed the internet at	Time Band	Inferred
Nationality	is nationality of person who accessed internet on	Date	Inferred
Nationality	is nationality of persons with highest average internet usage network traffic	Network Traffic	Aggregation
Nationality	is nationality of persons responsible for the most internet activity at	Time Band	Aggregation
Population group	is population group with heaviest internet usage having average daily internet usage of	Network Traffic	Aggregation
Population Group	is population group having the most internet usage at	Time Band	Aggregation
Population Group	is population group of person who accessed the internet from workstation/PC	US Workstation/PC	Inferred
US Faculty	is faculty of person who accessed the internet on	Date	Inferred
US Faculty	is faculty with heaviest internet usage having average daily internet usage of	Network Traffic	Aggregation
US Faculty	is faculty of person who accessed the internet at	Time Band	Inferred
US Faculty	is faculty of person who accessed the internet using workstation	US Workstation/PC	Inferred
US Module	is module taken by person who accessed the internet on	Date	Inferred
US Programme	is programme studied by person who also accessed the internet on	Date	Inferred
US Programme	is programme with heaviest internet users having average daily internet usage of	Network Traffic	Aggregation
US Residence Type	is residence type inhabited by person who accessed the internet on	Date	Inferred
US Residence Type	is residence type inhabited by person who accessed the internet at	Time Band	Inferred
US Residence Type	is residence type of person who accessed the internet with workstation/PC	US Workstation/PC	Inferred

B7 - Demographic and study programme information ←→ Library transaction and item information

Entity	Relation	Entity	Relation Type
Age	is age of oldest library item with subject topic	Library Item Subject Topic	Aggregation
Age	is age of youngest library item with subject topic	Library Item Subject Topic	Aggregation
Age	is average age of library items with subject topic	Library Item Subject Topic	Aggregation
Age	is age of library item	Library Item Title	Inferred
Age	is age of oldest library item	Library Item Title	Aggregation
Age	is age of youngest library item	Library Item Title	Aggregation
Age	is average age of all library items and specifically library item	Library Item Title	Aggregation
Age	is age of library item	Library Shelf Code	Inferred
Education Department	is education department of person who performed library transaction at	Time Band	Inferred
Education Department	is education department of persons responsible for the most library activity at	Time Band	Aggregation
High School	is high school attended by person who performed library transaction on library item with subject topic	Library Item Subject Topic	Inferred
High School	is high school of person who performed library transaction at	Time Band	Inferred
Language	is language of person who performed library transaction on library item published by	Library Item Publisher	Inferred
Language	is language of person who performed library transaction on library item	Library Item Title	Inferred
Marital Status	is marital status of person who performed action in WebCT at	Time Band	Inferred
Marital Status	is marital status of person who performed library transaction on library item at	Time Band	Inferred
Nationality	is nationality of person who performed library transaction on library item authored by	Library Item Author	Inferred
Nationality	is nationality of person who performed library transaction on library item yielding a status	Library Item Status	Inferred
Nationality	is nationality of persons responsible for the most library activity at	Time Band	Aggregation
Population Group	is population group of person who performed library transaction at	Time Band	Inferred
Population Group	is population group performing the most library transactions at	Time Band	Aggregation
US Faculty	is faculty of person who performed library transaction on	Date	Inferred
US Faculty	is faculty resulting in the most library transactions being conducted on library items with Dewey classification	Library Dewey Classification	Aggregation
US Faculty	is faculty of person who performed library transaction on library item published by	Library Item Publisher	Inferred
US Faculty	is faculty of person who performed library transaction on library item at	Time Band	Inferred
US Module	is module taken by person who performed library transaction on	Date	Inferred
US Programme	is programme studied by person who also performed action in WebCT on	Date	Inferred

B8 - Demographic and study programme information ←→ Printing transaction information

Entity	Relation	Entity	Relation Type
US Residence Type	is residence type inhabited by person who performed library transaction on library item at	Time Band	Inferred
Education Department	is education department of person who printed file at	Time Band	Inferred
Education Department	is education department of persons responsible for the most printing activity at	Time Band	Aggregation
High School	is high school of person who printed file with number of pages	Number of Pages	Inferred
High School	is high school of person who printed file at	Time Band	Inferred
Language	is language of person who printed document on	Date	Inferred
Nationality	is nationality of person who printed file with number of pages	Number of Pages	Inferred
Nationality	is nationality of persons with highest average number of pages printed	Number of Pages	Aggregation
Population Group	is population group who prints the most with per person average daily number of pages printed	Number of Pages	Aggregation
Postal Code	is postal code of person who printed file on	Date	Inferred
Town	is home town of person who printed file using workstation	US Workstation/PC	Inferred
US Department	is department of person who printed file on	Date	Inferred
US Faculty	is faculty of person who printed number of pages	Number of Pages	Inferred
US Faculty	is faculty who prints the most with per person average daily number of pages printed	Number of Pages	Aggregation
US Faculty	is faculty which is the most expensive printer of documents with average page cost per document	Page Cost	Aggregation
US Faculty	is faculty of person who printed file at	Time Band	Inferred
US Faculty	is faculty of person who printed file using workstation	US Workstation/PC	Inferred
US Module	is module taken by person who used printer	US Printer	Inferred
US Programme	is programme responsible for the printing of the most documents with document type	Document Type	Aggregation
US Programme	is programme studied by person who printed file with file size	File Size	Inferred
US Programme	is programme studied by person who photocopied number of pages	Number of Pages	Inferred
US Programme	is programme studied by person who printed number of pages	Number of Pages	Inferred
US Programme	is programme with heaviest printing having per person average daily number of pages printed	Number of Pages	Aggregation
US Programme	is programme with heaviest printing having per person average daily page cost per document	Page Cost	Aggregation
US Programme	is programme studied by person who used printer	US Printer	Inferred
US Residence Type	is residence type inhabited by person who printed file at	Time Band	Inferred
US Residence Type	is residence type of person who printed file with workstation/PC	US Workstation/PC	Inferred

B9 - Demographic and study programme information ←→ WebCT usage information

Entity	Relation	Entity	Relation Type
Education Department	is education department of person who performed action in WebCT at	Time Band	Inferred
Education Department	is education department of persons responsible for the most WebCT activity at	Time Band	Aggregation
High School	is high school of person who performed action in WebCT at	Time Band	Inferred
High School	is high school attended by persons who have the longest personal average WebCT dwell time	WebCT Dwell Time	Aggregation
Nationality	is nationality of person who performed WebCT action	WebCT Action	Inferred
Nationality	is nationality of person who with WebCT dwell time	WebCT Dwell Time	Inferred
Nationality	is nationality of persons who have the longest personal average WebCT dwell time	WebCT Dwell Time	Aggregation
Population Group	is population group performing the most actions in WebCT at	Time Band	Aggregation
Population Group	is population group having the greatest average per person WebCT dwell time	WebCT Dwell Time	Aggregation
Population Group	is population group of person with WebCT dwell time	WebCT Dwell Time	Inferred
Town	is home town of person who performed action in WebCT on	Date	Inferred
Town	is home town of person with WebCT dwell time	WebCT Dwell Time	Inferred
US Department	is department where module is taken by person with WebCT dwell time	WebCT Dwell Time	Inferred
US Faculty	is faculty of person who performed action in WebCT on	Date	Inferred
US Faculty	is faculty of person who accessed WebCT at	Time Band	Inferred
US Module	is module taken by person who performed action in WebCT on	Date	Inferred
US Module	is module taken by person who accessed WebCT learning context	WebCT Learning Context	Inferred
US Programme	is programme studied by person who performed WebCT action	WebCT Action	Inferred
US Programme	is programme studied by person with WebCT dwell time	WebCT Dwell Time	Inferred
US Residence Type	is residence type inhabited by person who accessed WebCT at	Time Band	Inferred
US Residence Type	is residence type of person who accessed WebCT learning context	WebCT Learning Context	Inferred

B10 - Internet usage cost information ←→ Library transaction and book information

Entity	Relation	Entity	Relation Type
Library Item Title	is title of library item on which library transaction was performed by person who also accessed the internet on	Date	Inferred

B11 - Internet usage cost information ←→ WebCT usage information

Entity	Relation	Entity	Relation Type
Date	is date internet was accessed by person who also performed WebCT action	WebCT Action	Inferred
US Workstation/PC	is workstation/PC used to access the internet by person who performed WebCT action	WebCT Action	Inferred
WebCT Dwell Time	is WebCT dwell time of person with internet usage of	Network Traffic	Inferred
WebCT Learning Context	is WebCT learning context that was accessed by person who also accessed internet on	Date	Inferred
WebCT Learning Context	is WebCT learning context accessed by person who accessed the internet with workstation/PC	US Workstation/PC	Inferred

B12 - Library transaction and book information ←→ Printing transaction information

Entity	Relation	Entity	Relation Type
Date	is date file was printed by person who performed library transaction at	Library Branch	Inferred
Date	is date file was printed by person who performed library transaction on item	Library Item Title	Inferred
Date	is date library transaction was performed by person who also used printer	US Printer	Inferred
Library Item Author	is author of library item on which library transaction was performed by person who also printed file on	Date	Inferred

B13 - Library transaction and book information ←→ WebCT usage information

Entity	Relation	Entity	Relation Type
Date	is date library transaction was performed by person who also performed WebCT action	WebCT Action	Inferred
Library Item Subject Topic	is subject topic of library item on which library transaction was performed by person who accessed WebCT learning context	WebCT Learning Context	Inferred
Library Item Title	is title of library item on which library transaction was performed by person who also performed action in WebCT on	Date	Inferred
Library Transaction Patron Type	is library patron type who performed action in WebCT with context role	WebCT Learning Context Role	Inferred
WebCT Action	is action performed in WebCT by person who performed library transaction on library item with subject topic	Library Item Subject Topic	Inferred
WebCT Action	is WebCT action performed by person who also performed library transaction type	Library Transaction Type	Inferred
WebCT Learning Context	is WebCT learning context that was accessed by person who also performed library transaction on	Date	Inferred

B14 - Printing transaction information ←→ WebCT usage information

Entity	Relation	Entity	Relation Type
Date	is date action was performed in WebCT by person who also used printer	US Printer	Inferred
Date	is date file was printed by person who also accessed WebCT Learning context	WebCT Learning Context	Inferred
US Workstation/PC	is workstation/PC used to print file by person who performed WebCT action	WebCT Action	Inferred
WebCT Learning Context	is WebCT learning context accessed by person who printed file with workstation/PC	US Workstation/PC	Inferred

Appendix E

Implementation 2 Questionnaire

Impressions of a Conceptual Framework Implementation at the University of Stellenbosch

This document is aimed at the capturing the impressions of stakeholders involved in a conceptual framework implementation at the University of Stellenbosch.

Name: _____

QUESTIONS (Mandatory):

This feedback captured by this section is important as it plays a major role in proving (or disproving) the hypothesis of the research project which this implementation forms part of. Please tick the appropriate boxes and provide additional information where relevant.

	Yes	No	Other
1. Does the conceptual framework that was presented provide an accurate view on the information contained in the implementation domain?	☐	☐	☐
Comment:			
2. Do the types and relations that were defined provide an intuitive view on the information contained in the implementation domain?	☐	☐	☐
Comment:			
3. Is there benefit in viewing the organisation's information as a navigable conceptual framework?	☐	☐	☐
Comment:			
4. Could the utilisation of a conceptual framework improve the efficiency of decision-making in the organisation?	☐	☐	☐
Comment:			
5. Did the conceptual framework methodology address all the major issues required for a successful conceptual framework implementation?	☐	☐	☐
Comment:			

Comments (Optional):

Although this section is optional, any contribution will be appreciated, as it may influence the conceptual framework implementation methodology, the conceptual framework concept itself, or the results of the research project this implementation forms part of.

Methodology comments

Please list any suggestions with regards to the methodology that was followed to develop the CF.

Impressions

- Benefits: *Comment on the usefulness of the CF (not Organon!). Referring to specific CF facts that proved/might prove to be useful should be stated.*

- Shortcomings: *Shortcomings of the content or structure should be stated in as much detail as possible.*

Growing the CF

What information can be added to the CF to make it even more useful/insightful?

Other comments:

Please list any other comments.