The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Epistemology as the Basis for a Corporate Memory Model

A dissertation submitted to the
Department of Computer Science,
Faculty of Science, University of Cape Town
in fulfilment of the requirements
for the degree of
Master of Science

By
Ilan Azbel
February 2001

Supervised by
Sonia Berman

University of Cape Town

University of Cape Town

© Copyright 2001

by

Ilan Azbel

Abstract

Epistemology is the philosophical study of knowledge; it attempts to understand which beliefs a person can claim to know. This thesis discusses how theories found in epistemology can be used to understand what knowledge an organisation possesses, and shows how epistemic theories can be used as a basis for a corporate memory model. Work in artificial intelligence knowledge representations and corporate memory knowledge representations are compared with ideas found in epistemology. A corporate memory model based on epistemology is proposed, and it is shown how this model can be used as a basis for development of corporate memory systems.

University of Cape Town

Dedications

This thesis is dedicated to the people that supported me through the most testing time of my life: my parents, Mark and Gena, my brother Arik and his wife Michelle, my grandmother Luba, and my dear friends Alon, Filip, Sam and Virginia.

University of Cape Town

Acknowledgements

Thank you to my friend John Ziniades who sparked my interest in knowledge management. Thank you to my supervisor, Sonia Berman, for her insight and invaluable contributions to this research.

University of Cape Town

Contents

University of Cape Town

University of Cape Town

University of Cape Town

University of Cape Town

List of Figures

University of Cape Town

1 Introduction

Development of complex systems involves many people, each with their own perspectives, skills, responsibilities, knowledge and expertise [EFKN94]. Many specialised fields of expertise suffer from a thin spread of domain knowledge, fluctuating and conflicting requirements and breakdowns in communication and coordination [CKI88]. Such reasons are responsible for creating a need for organizations to manage the intellectual capital of their employees, and thereby fueling the growth of the field of knowledge management.

There are numerous definitions of knowledge management that can be found in literature. Davenport and Prusak define *knowledge management* as "A fluid mix of framed experiences, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of the knowers. In organizations, it often becomes embedded not only in documents or repositories but also in organizational routines, processes, practices and norms" [DP98]. Van der Spek and Spijkervet define knowledge management as consisting of activities that "... aim at optimal use and development of knowledge, now and in the future. Knowledge management determines which knowledge, where, in which form and at which point of time, should be available within an organization, company or network of institutions. It employs a broad spectrum of techniques and instruments to improve the performance of knowledge operations and the learning capabilities of a system..." [vdSS97].

It is clear from the above definitions that the field of knowledge management concerns itself with the management processes involved in managing intellectual capital within an organization. This thesis is focused on investigating the field of *corporate memories*, which are the mechanisms used for storage of knowledge during knowledge management processes.

University of Cape Town

Many organizations are finding a need for corporate memory technology that can be used not only for storage and retrieval of shared information, but also for storage and retrieval of individual ideas and thoughts held by employees. Such systems have a strong need to understand and facilitate *human factors* of information re-use.

The field of epistemology has been around for thousands of years; many of the greatest philosophers have contemplated thoughts and ideas about the essence of human knowledge. A great majority of philosophers believe that knowledge is defined as *justified true belief*, that is, a person can only have knowledge of something if the belief in question is justified. Epistemology attempts to define human knowledge by trying to understand what a person is justified in believing.

Unlike research in artificial intelligence and corporate memory systems, the objective of epistemology is not to understand how humans represent knowledge, nor is it to understand how humans re-use knowledge. The objective of epistemology is to investigate the essence of what a person can claim to know.

University of Cape Town

1.1 Research Goals

There are two primary goals of this thesis:

- To compare the knowledge representations found in artificial intelligence systems and corporate memory models to theories of knowledge found in epistemology.
- To present a new corporate memory model that is based on epistemic theories of knowledge.

To attain these goals it is necessary to firstly outline the most important ideas found in three different fields of research: epistemology, knowledge representations in artificial intelligence, and corporate memory systems. Secondly, to compare work done in artificial intelligence and corporate memory models to theories found in epistemology. Epistemic theories of knowledge are then used to design a new corporate memory model and to compare this model to existing corporate memory models.

Through the description of an epistemically grounded corporate memory model, *EpiK*, it will be shown that such a model reflects many of the characteristics that have shown to be desirable in corporate memory systems. EpiK will make use of epistemic theories of Reliabilism, Consequentialism, Pragmatism and Ontology to represent corporate knowledge and to facilitate learning. It will be shown how such a model can also be used in understanding the social dynamics of an organization.

1.2 Thesis Outline

Chapter 2 (*Epistemology*) will give an introduction to epistemology by discussing the traditional account of knowledge as *justified true belief*. A number of theories will be presented about the ways in which knowledge can be justified, ranging from rationalist theories that require complete certainty of knowledge, through to empiricist theories that rely on less reliable sensory experiences as a source of knowledge and justification.

Chapter 3 (*Knowledge Representations in Artificial Intelligence*) presents summaries of the most popular knowledge representations used in artificial intelligence. Such representations include logic, semantic networks, rule-based systems, frames, and connectionist models. Each of these representations is described according to data models used to represent information rather than the reasoning processes that act on the models. The knowledge representations presented are then discussed from an epistemological perspective by drawing parallels to concepts found in Chapter 2 (*Epistemology*).

Chapter 4 (*Corporate Memory Models*) is focused around presenting pertinent issues in corporate memory technology such as degrees of knowledge formality, declarative versus procedural knowledge representation, workplace integration, and knowledge maintenance issues. Data models used in corporate memory technology are presented and specific corporate memory systems are discussed. The corporate memory models are then analysed from an epistemological perspective by comparing them to theories found in Chapter 2 (*Epistemology*).

University of Cape Town

Chapter 5 (*An Epistemically Grounded Corporate Memory Model*) presents a new corporate memory model, EpiK, based on epistemological concepts of reliabilism and consequentialism. This chapter outlines the general structure and components of the new model, how it enables users to learn, and how it can be used to evaluate people and knowledge in an organization. Parallels are drawn between the new model and epistemic theories found in Chapter 2 (*Epistemology*) as well as to other corporate memory models presented in Chapter 4 (*Corporate Memory Models*).

Chapter 6 (*From Model to Corporate Memory*) extends the overview of EpiK by presenting a more detailed explanation of each of the components through the use of an entity-relationship diagram. A concrete realisation is described in terms of procedures that a user would step through in using a corporate memory system based on this model. This chapter also shows that it is feasible to develop a working corporate memory system based on EpiK by discussing a number of prototype implementations.

Chapter 7 (*Evaluating and Extending EpiK*) describes two surveys, along with their results, that were conducted to investigate the possibility of integrating EpiK into existing email infrastructure of an organization.

Chapter 8 (*Conclusion*) concludes by summarizing each of the chapters of this thesis, presenting the benefits of an epistemically grounded corporate memory, as well as outlining possibilities for future work.

University of Cape Town

2 Epistemology

To act wisely in the world, it is necessary to know that world and understand it. Epistemology is the study of the nature and validity of knowledge and examines the degrees of certainty and the difference between knowing (with certainty) and believing (without certainty). Epistemology does not attempt to understand how humans represent knowledge, nor is it to understand how humans re-use knowledge. The objective of epistemology is to investigate what a person can claim to know.

People often assert confidently what they take to be the truth, only to find later that the assertions taken have been false – hence the need not only to know the world, but also to know oneself, and in particular, to understand the character and reliability of human cognitive powers. Epistemology seeks to provide this needed understanding about knowledge.

There are a number of questions that epistemology attempts to answer: What is knowledge? What do we know? What is it for a belief to be justified? Which of our beliefs are justified? To answer these questions one needs to define the concept of knowledge, determine the extent of the knowledge, analyse when beliefs are justified and investigate what one is justified in believing [STE96].

There are philosophers that are sceptical whether a person is able to have any knowledge whatsoever. Such sceptical theories range from doubting the existence of an external world, doubting knowledge about one's own mind, doubting knowledge of past and future experiences, and doubting knowledge of unperceived events [MOS96]. Although such sceptical arguments play a very important role in epistemology, this dissertation assumes that humans are capable of attaining knowledge about the world.

This chapter will provide an overview of the primary views and opinions held by philosophers regarding the nature of knowledge. This chapter is by no means a complete

University of Cape Town

history of epistemology, but focuses rather on outlining the most important concepts that have emerged over the years with a focus on contemporary epistemology. Section 2.1 outlines the concept of epistemic justification and discusses some mainstream theories about knowledge. Section 2.2 describes a number of non-mainstream theories about knowledge. Section 2.3 discusses typical knowledge sources and explains their role in a person's justification system. Section 2.4 describes the concept of social epistemology, and Section 2.5 briefly outlines the theory of Ontology creation.

Section 2.6 concludes this chapter by summarizing the important aspects of epistemology that will be re-used in later chapters for comparison with knowledge representation technologies, as well as for a basis for a new corporate memory model.

2.1 Justification and Knowledge as Justified True Belief

A definition of knowledge that is accepted by many contemporary philosophers is that of knowledge as *justified true belief*. The definition of knowledge as *justified true belief* is comprised of three conditions. The truth condition: what is known must be true, or in other words, it is impossible to know a proposition that is false. The belief condition: that one needs to believe (voluntarily or involuntarily) that the proposition is true. The justification condition: true beliefs amount to knowledge only if they are justified [STE96]. The justification condition excludes lucky guesses: if one believes a proposition, and the proposition is true but one is not justified in that belief, then the belief cannot be regarded as knowledge because the belief is a lucky guess [MOS96].

A distinction between belief and acceptance can be drawn. It is argued that acceptance is formed in the interest of obtaining truth and avoiding error, whereas holding a belief might be for other purposes such as usefulness, happiness, or fulfilment of a moral duty [LEH90]. This means that holding a belief does not mean that belief has been accepted.

It is clear to see that the argument about the nature of knowledge comes down to an analysis of what processes people take to be sufficiently reliable in producing information that corresponds to perceptions of the world around us.

University of Cape Town

The activity of justifying a belief is when one engages in a process that attempts to convince others that a belief is justified. This justification may, for example, be in the form of a mathematical proof, or in the form of sense experiences. One may even reference a "classical" piece of literature about a subject that has been written by a well-known author as justification for a belief.

Beliefs can belong to one of three major certainty levels: *necessarily true*, *necessarily false*, or *contingent*. Necessary truths are beliefs that are true without the need for justification; they are obviously true, for example "all bachelors are unmarried". Self-evident truths are *epistemically immediate* - one believes without having to go through any justification processes [AUD98]. Necessarily false beliefs are beliefs that are obviously false without the need for justification; for example: "I have a colourless red car". Contingent propositions are propositions that are neither necessarily false nor necessarily true. Different forms of justification produce different levels of contingency for beliefs. For example, if a person has rationalist tendencies then sense experiences would not be seen as a very reliable form of justification for a belief.

Philosopher Immanuel Kant introduced the terms *a priori* and *a posteriori*. Kant defines a-priori knowledge as knowledge that is prior to, or gained independently of experience, whereas a posteriori knowledge is knowledge gained through experience [STE96]. The idea of a priori knowledge is closely related to that of a priori justification because what makes a piece of knowledge a priori is the way in which is has been justified. Justification can take place through experience or independently of experience. Therefore, a priori knowledge can be seen as knowledge whose justification is a priori [STE96].

A priori justification cannot be based on sensory experiences such as visual, auditory, tactile, or gustatory, nor can it be based on memorial or introspective experiences. This leads one to the idea that all a priori justification is based on necessary truths. This means that simply thinking about the proposition is sufficient for recognising its truth. If one is

University of Cape Town

justified in believing a proposition on the basis of authority, then one's justification for believing that proposition is a posteriori [STE96].

The process of justification of a belief results in a truth level for a belief. This truth level is the primary point of contention among philosophers. Rationalists argue that only completely truthful beliefs that are beyond any doubt can be regarded as knowledge, while empiricists' viewpoint is that knowledge does not require complete certainty and in many cases they maintain that pragmatism plays an important role in one's judgment of whether a belief is truthful or not.

2.1.1 Foundationalism

Foundationalism is the theory about the structure of a belief system, it is not an analysis of epistemic justification. This theory assumes that there are basic beliefs that form the foundation for all other beliefs in a person's belief structure. Non-basic beliefs form the super-structure of the belief system. Every non-basic belief owes its justification to one or more beliefs below it in the belief structure (either basic or non-basic) [STE96] [MOS96]. This type of belief structure is clearly illustrated in Figure 1 below.

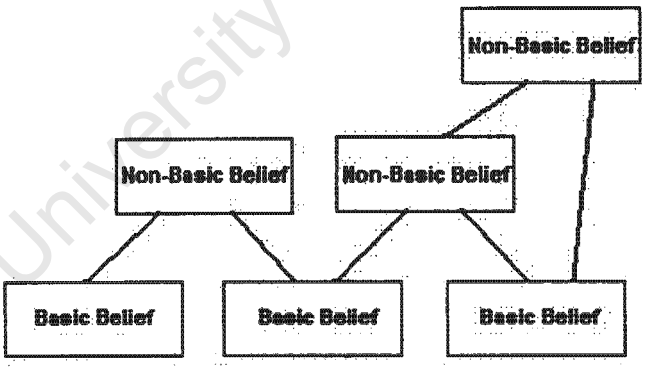


Figure 1: Foundationalism represented graphically

Basic beliefs do not need to be indubitable, infallible, nor certain. The reason for this is that basic beliefs can take their justification from introspection, perception and memory. All these factors add to the concept of reliability of a person to form a basic belief [AUD93a]. If a person is reliable in making certain observations, then it is highly

University of Cape Town

probable that basic beliefs of a certain nature made by this person are correct. Therefore, the justification of a basic belief depends highly on the reliability of the individual as well as the process.

The core of foundationalism is the existence of *basic* knowledge. The regress argument is a popular argument used to prove the existence of such knowledge. The argument is: Suppose that a belief B1 is justified. This means that B1 is either basic or non-basic, if it is non-basic, then there must exist a belief B2 from which B1 gets its justification. Now B2 itself can be basic or non-basic. If no basic beliefs exist, then this justification tree would continue at infinitum. But an infinite regress of justification does not justify anything [AUD98], [STE96]. Therefore, there must exist some basic beliefs from which all other beliefs in a belief system can derive their justification.

Foundationalists differ on a number of issues. These include the explanation of foundational knowledge, the explanation of how justification can be transmitted from foundational to non-foundational beliefs and the effects of changes to the belief structure [MOS96]. Some foundationalists regard only beliefs gained through immediate experience and logically intuitive beliefs as suitable for a foundation. In less restrictive foundationalist theories, one could allow beliefs about material objects and beliefs about one's past experiences [AUD93b]. There are several theories about the relationship between foundational beliefs and non-foundational beliefs. Some might stipulate that beliefs in the superstructure are only sustained by one foundational belief, while other theories might allow multiple foundational beliefs to support a belief in the superstructure [AUD93b].

Furthermore, there is also contention about how changes to the belief structure can affect beliefs in the rest of the structure. Some theories may stipulate that loss of a foundational belief always results in loss of at least one superstructure belief, other theories might allow superstructure beliefs to be "over-determined", that is, elimination of a foundational belief does not result in the loss of a superstructure belief. Yet other theories maintain that superstructure beliefs cannot displace foundational beliefs unless the

University of Cape Town

superstructure beliefs that they maintain become inconsistent with a certain foundational belief but consistent with other foundational beliefs [AUD93b].

According to *classical foundationalism*, a belief cannot be justified unless its justification can guarantee its truth. To guarantee a belief's truth, the belief needs to meet two conditions: basic beliefs need to be infallible, and non-basic beliefs would need to follow deductively from basic beliefs. Classical foundationalism leaves no room for truths that are based on observations; it insists that all beliefs are physically justified [STE96]. . *Modern foundationalism* differs from classical Foundationalism in the fact that it allows for non-deductive transmission of justification from basic to non-basic beliefs [STE96]. "... foundational beliefs need not possess or yield certainty, and need not deductively support justified non-foundational beliefs" [MOS96]. *Minimal foundationalism* throws out the ideas of classical and modern foundationalism by simplifying the difference between basic and non-basic beliefs. According to minimal foundationalism, the difference between basic and non-basic beliefs is solely a function of how the belief received its justification. If the belief received its justification from experience and not from other beliefs, it is basic, otherwise it is non-basic [STE96], thereby allowing non-basic beliefs to receive their justification using any method a person believes will maintain truth.

University of Cape Town

2.1.2 Coherentism

The theory of coherentism rejects the fact that knowledge and justification rest on a foundation of basic beliefs. Coherentists argue that what justifies a belief must always be one or more other beliefs, but in no way do they define beliefs as "basic". A graphical representation of this structure can be seen in Figure 2 below.

Because coherentists agree with foundationalists that an infinite regress of reasons cannot justify a belief, they try to make sense of circular justification; that is, a circular set of beliefs that justify each other without a 'foundation'. According to some coherentists, circular justification is acceptable if the circle is sufficiently large [PLA93]. It is however difficult to understand how the size of a justification circle has any effect on its capacity

to justify. The argument against circular justification rests on two ideas: Firstly, self-justification cannot occur, and secondly, that a relation "x justifies y" is transitive. For example "B1 -> B2 ->B3 ->B1" is false because this circle simplifies to "B1->B1" by the idea of transitivity and "B1->B1" cannot be true because self-justification is impossible [STE96], [AUD98].

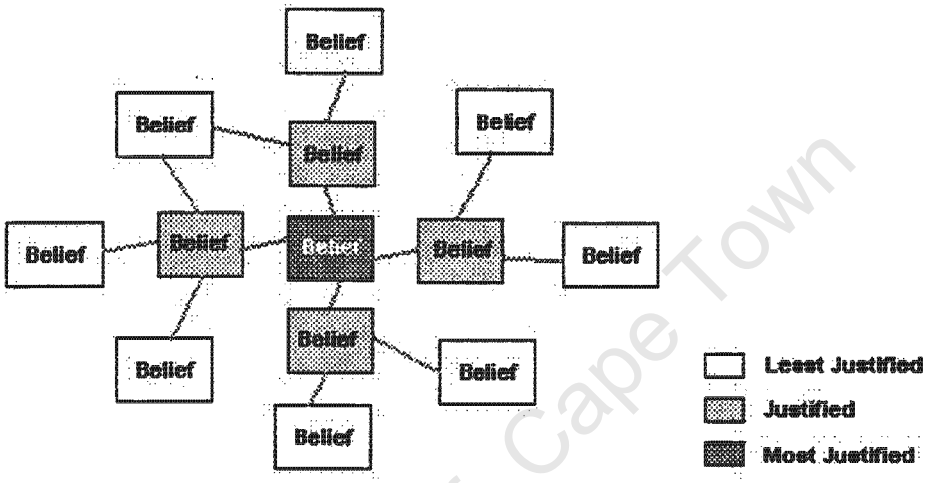


Figure 2: Coherentism represented graphically

Coherentists maintain that beliefs are better justified the more the system exhibits what is termed "explanatory coherence". A belief has explanatory coherence either if it explains other beliefs or it is explained by other beliefs. Some coherentists hold the view that explanatory coherence is not necessary for coherence, it does however add to a belief's justification. Robert Audi argues that incoherence is a defeater of justification, but this does not imply that coherence has a positive effect on justification [AUD93a].

Keith Lehrer proposes a theory of coherentism that depends greatly on conceptions of oneself as a knower. He proposes that one evaluates the trustworthiness of one's own belief producing faculties and thereby assess the probability of beliefs formed by these faculties [LEH86].

University of Cape Town

Coherentism should not be viewed in a linear form but rather in a holistic form, "Essentially systematic or holistic in character: beliefs are justified by being inferentially related to other beliefs in the overall context of a coherent system." [BON85].

2.1.3 Reliabilism

Reliabilism is a theory of justification rather than the structure of knowledge. It is therefore unlike the theories of foundationalism and coherentism, whose primary objective is to define the structure of knowledge.

Reliabilism measures the reliability of a belief according to the reliability of the justification methods that produced or substantiates that belief. A generic form of Reliabilism might be worded as follows:

> "S's belief that a proposition *p* at time *t* is justified if it is the outcome of a process of belief acquisition or retention which is reliable, or leads to a sufficiently high preponderance of true beliefs over false beliefs" [SOS91]

One of the problems with reliabilism is how to avoid using processes that have produced only one output ever, or an artificially selected process that yields only true beliefs [SOS91]. If such specific processes are regarded as reliable, then every belief could be seen as being produced by such a reliable process and all beliefs would be true. On the other hand, one must avoid making use of processes that are too general, such as perception, which can produce not only justified beliefs but also unjustified ones, even if in general, perception is a reliable form of knowledge acquisition for a human under normal circumstances [SOS91].

Alston claims that some practices, such as sense perception are superior to others, for example, crystal ball gazing. "Selection of doxastic[1] practices is not dictated by rules but must be based on sensitivity, experience, trained judgement and intuition" [ALS89]. With

[1] In Foundationalism, the term 'doxastic' refers to beliefs that have justification from other beliefs.

University of Cape Town

this sort of judgement, one accepts doxastic practices as rational and eliminates them as deficiencies are exposed, for example, conflict with more well established practices.

There are a number of more specific reliabilist theories such as *process reliabilism* and *virtue reliabilism*. Process reliabilism judges belief-producing processes that produce either reliable or unreliable beliefs. Examples of unreliable processes include confused reasoning, wishful thinking, reliance on emotional attachment, guesswork, and hasty generalisation. Some examples of reliable processes include perceptual processes, remembering, good reasoning and introspection [GOL79].

Virtue reliabilism explains belief justification in terms of virtues and vices. Intellectual virtues are cognitive processes that are seen as reliable, while intellectual vices are cognitive processes that are seen to be unreliable. The theory of virtue reliabilism proposes that a belief is justified if it results from an intellectual virtue, and an unjustified belief result from an intellectual vice [STE96]. When one evaluates beliefs, the psychological process with which the belief was generated needs to be evaluated against a list of virtues and vices. If the belief was produced using a 'virtue', then the belief is justified, and if the belief was produced by a vice, the belief is unjustified [GOL91].

2.1.4 Contextualism

"Man is a social animal, and yet when it comes to the justification of beliefs philosophers tend to ignore this fact. But this is one contextual parameter that no adequate theory of justification can overlook. .... Thus social information – the beliefs, information , and theories of others – plays an important part in justification, for it is part determines what objections will be raised, how a person will respond to them, and what responses the objectors will accept" [ANN78].

The theory of contextualism emphasises pragmatic and social factors in the acquisition of knowledge. In contextualism, people assume the existence of contextually basic propositions that are starting points for inquiry about a subject [MOS96]. Such

University of Cape Town

contextually basic propositions may lack evidential support, but can themselves justify other beliefs.

An example of a contextualist theory is that presented by David Annis. His theory is based on notions of beliefs and appropriate *objector groups* [ANN78]. The objector group is a group of individuals that are qualified to raise objections about a belief. In this theory, a belief is *contextually basic* if a person, at a specific point in time, relative to an objector group, does not require any reason for being in possession of a belief. A person may in some cases be justified in believing a proposition relative to one issue context but not with respect to another. If a member of the objector group raises an objection to the belief then one would need to justify that belief based on currently available evidence [ANN78].

2.1.5 Causal Theory of Knowledge

The causal theory of knowledge is seen as a very important theory in epistemology because it overcomes the Gettier problem. In 1963, Gettier argued that *true justified belief* was not a sufficient condition for knowledge. Gettier did this by giving an example of a situation when one is justified in believing something true, yet does not *know* it. A detailed discussion of the argument is beyond the scope of this discussion, but the reader is referred to Gettier's classical paper [GET63].

Recently, philosopher Alvin Goldman has proposed a causal theory of knowledge that claims to overcome the Gettier problem by proposing that one views justification of beliefs as causal conditions. The idea is that a person knows a proposition *p* if that person believes *p*, *p* is true, and *p* is causally produced and sustained by the fact that makes *p* true [MOS96]. For example, if one perceives one thing influencing another, and one also perceives the effect of this influence, then one is justified in having knowledge about the cause-effect relationship between these two things.

University of Cape Town

This definition of knowledge is based on knowledge as *justified true belief*, but grounds its method of justification on the existence of causal connections between actual occurrences and beliefs.

2.1.6 State of Affairs

Sayre believes that knowledge can be expressed in terms of *State of Affairs* (SOA) and *attitudes* [SAY97]. Sayre believes that human cognition is not completely based on the truth of propositions (as assumed by contemporary epistemology). He takes the stance that cognition should be described by attitudes such as believing fearing, wanting, forgetting, understanding, predicting, reporting, commanding, being aware, realising, thinking, judging, believing and knowing (to name just a few). In the example "I predict that the sky will turn cloudy" storms are eventualities and it does not make sense to speak of eventualities as propositions [SAY97]. That is, the speaker's utterance says nothing about the truth of the sky becoming cloudy. Another example given to demonstrate state of affairs is "Camille believes the sun is shining" [SAY97]. This is a proposition; if though, the word "believes" is replaced by the word "notices", then the statement becomes a SOA, not a proposition.

Propositional attitudes are distinguished via truth values, these include believe, assert, deny etc. "Attitudes of cognitive access" are things such as discovering, finding, noticing and do not admit to qualification in terms of truth [SAY97].

Sayre does make some very strong arguments for his view. For example, it is correct to speak of *true* and *false beliefs*, but it is incorrect to speak of *right* or *wrong* discoveries. Furthermore, one normally asks "Why?" and not "How?" in probing attitudes of propositional stance such as believing, denying and asserting; but we typically ask "How?" and not "Why?" in probing attitudes of cognitive access such as finding, noticing and ascertaining [SAY97].

University of Cape Town

2.2 Other Epistemic Theories of Knowledge

There are numerous theories about knowledge that do not depend on truth as their primary condition for knowledge. Such theories are beyond the scope of this discussion and therefore only brief summaries of such theories are given below [GOL99]:

- There is no such thing as transcendent truth. What we call "true" is simply what we agree with. So-called truths or facts are merely negotiated beliefs, the products of social construction and fabrication, not 'objective' or 'external' features of the world.
- Knowledge, reality, and truth are the products of language. There is no language-independent reality that can make our thoughts true or false.
- There is no privileged epistemic position, and no certain foundations for belief. All claims are judged by conventions and language games, which have no deeper grounding. There are no neutral, trans-cultural standards for settling disagreements.
- Truth cannot be attained because all putative truth-oriented practices are corrupted and biased by politics or self-serving interest.
- Appeals to truth are merely instruments of domination or repression, which should be replaced by practices with progressive social value.

All the above variations have their support from certain philosophers, but none as yet has proved itself more viable than truth based epistemology. In [GOL99], Alvin Goldman presents substantial argument against each one of the above variations on truth-based epistemology to show how they have failed.

2.3 Sources of Justification and Knowledge

As one can see from the discussion in Section 2.1 (*Justification and Knowledge as Justified True Belief*), many of the theories of knowledge depend on a person's rating of reliability of the source of beliefs.

University of Cape Town

The two most well known theories about what humans are justified in believing are *rationalism* and *empiricism*. On the one hand are the empiricists that maintain that all human knowledge comes from sense experience or derived from the experiences attained from the senses. On the other hand are the rationalists that believe that only reasoning can provide man with certain knowledge.

Despite the two contrasting ideas about knowledge, one would struggle to find an empiricist that denies the importance of reason, and similarly, one would find it difficult to find rationalists that deny the usefulness of sense experience. The difference between rationalists and empiricists is one of emphasis: rationalists emphasise the importance of reasoning, while empiricists emphasise the importance of sense experience. It is important to ask whether or not a theory that blends empiricist and rationalist views produces knowledge that is certain to a high enough degree to be acceptable by both schools of thought.

Sources of knowledge and justification that will be discussed in this section include perception, reasoning, memory, introspection and testimony.

University of Cape Town

2.3.1 Perception

One has many beliefs based of what one sees, hears, feels, smells and tastes which seem to have no reason to be doubted. These perceptual beliefs are commonly justified, and are quite often true and constitute knowledge [AUD98]. Perception is a source of belief, justification and knowledge without dependence on other belief producing capacity, such as reasoning. It is comprised of four primary elements; a receiver, an object, a sensory experience and a relation between the object and the subject. [AUD98].

Perception is a type of causal relation. Even simple perception requires an object of perception and a sensory experience that in some way corresponds to that object and records an extensive amount of information about the object. "It [perception] implies that the perceiver at least normally has justification for certain beliefs about the object, and it normally produces both justified beliefs about that object and knowledge of it" [AUD98].

2.3.2 Memory

Without memory, perceptual knowledge could not be amassed and used to help one build theories or make decisions about day-to-day situations. Furthermore, memory supplies raw material for introspection and thought. "One would have vastly less to look in or think about if we did not remember sights and sounds, conversations and embraces, ideas and plans" [AUD98].

Memory preserves beliefs and allows one to use these stored beliefs to supply premises in reasoning. Memory is a source of justification by having beliefs reinforce other beliefs. It should be noted however that memory only retains beliefs and knowledge, it does not generate them, it can however be used during the process of reasoning to help generate new knowledge.

2.3.3 Introspection and Consciousness

Self-knowledge is an important source of information when questions arise about justification regarding external objects. "Confronted with a strange perception, one can carefully consider stability, coherence and variations of perceptual experiences of it in order to rule out hallucination" [AUD98]. Through introspection one acquires justified beliefs and knowledge only about one's internal world [AUD98]. Introspection is *active*, that is, it can occur all the time and is not initiated by an event. This can be compared to perception and memory that are *reactive* – occur at will or as a reaction to an event.

2.3.4 Reasoning

"Reasoning differs from introspection, memory and perception because it allows one to know truths that hold not only in one's experiences, but also in any circumstances whatever" [AUD98]. Logic allows one to maintain consistency in one's belief structure. Logic, however is not the only tool used by the human mind; probability plays a major role in evaluation of situations. "Every cognitive representation of the world is associated with some level of uncertainty" [KR95].

University of Cape Town

Deductive reasoning is geared towards validity of premises. In general, deductive reasoning is truth preserving; it allows one to start with a premise that has conclusive evidence and end up with conclusions for which one has conclusive evidence. It is important to keep in mind that each step of the deductive process needs to have conclusive evidence that substantiates the deductive step [STE96].

Inductive reasoning is reasoning that offers a very likely chance that a premise is true. By applying inductive reasoning from multiple sources of information to a single belief, the level of justification of that belief can increase or decrease appropriately [AUD98].

As one experiences the world, new beliefs arise, both from basic sources and inferentially. As rational beings, humans are almost constantly forming beliefs on the basis of other beliefs. "One can learn something new from many inferences, and in making inferences to propositions that we think best explain something we take ourselves to know already, we sometimes learn truths that are both new and important" [AUD98]. One enlarges and strengthens a body of knowledge and justified beliefs through inference.

University of Cape Town

2.3.5 Testimony

Testimony refers to reliance for knowledge and justification from other people. Unlike individual sources of knowledge and justification such as perception, memory, consciousness and reason, testimony is a primary social source of knowledge and justification. Many of the beliefs one gains through testimony are justified or constitute knowledge. Testimonial beliefs are not inferentially dependant on knowledge or justified beliefs or prior premises. Testimonial beliefs are therefore source dependant but not necessarily premise dependant [AUD98].

In the same way that memory does not produce beliefs, testimony cannot produce beliefs unless the person testifying has gained that belief via another source. In this way, testimony is based on the credibility of the source. The credibility might be built up by

the source by giving more information about the context of the belief and this information fits into one's personal belief structure.

2.4 Social Epistemology

Individual epistemology assumes that every person has the ability to represent reality in different ways. Everyone creates inner representations that partly or fully correspond to the external world. *Reality* is the point of reference for inner representations and *truth* is the degree to which inner representations correspond to the real world [KR95]. Many of the theories discussed thus far have primarily been focused on the mental operations of individuals and in isolation from other people.

Alvin Goldman differentiates social epistemology from individualistic epistemology, as "Individualistic Epistemology is a discipline that evaluates practices along truth-linked dimensions. Social epistemology evaluates specifically social practices along these dimensions" [GOL99]. Not only does knowledge reside in the individual, but it also resides at a social level; in the relations among people in a community [KR95]. It is human culture to communicate knowledge about discovered information. Communication is a quicker and easier means of furthering one's knowledge in a given topic than fresh discovery [GOL99].

Goldman described communication in four major steps: discovery, production and transmission, reception, and acceptance:

1. Discovery – discovery of new knowledge through processes that are deemed reliable to the individual
2. Production and Transmission – this step involves a number of decision making processes that result in communication of the discovered knowledge:
    a. Should we pass this information on?
    b. What part of the information should we pass on?
    c. How should the information be passed on?
    d. To whom shall we pass it on to?

University of Cape Town

3. Reception – At this stage, the appropriate recipient gets the appropriate knowledge
4. Acceptance – The recipient of the information adopts an attitude towards the knowledge. Attitudes would include believe, reject, withhold judgement, intermediate degree of belief. Message acceptance is highly dependant on the perceived competence of the sender and the perceived honesty of the sender.

The truth of knowledge obtained from other people can be estimated based on the trustworthiness of the person communicating the information. In this way, people are perceived as knowledge producing process that can be seen as reliable or unreliable in the same way as, for example, perception.

Despite this reliabilist view of people, acceptance of knowledge does not need to be based purely on the truth-value of the knowledge. *Consensus consequentialism* theory replaces the truth-value with agreement or disagreement. This theory assumes that the aim of intellectual practices is not to produce true belief, but to produce consensus among all parties involved. *Utility consequentialism* theory evaluates the value of a belief according to how much happiness, satisfaction, or value realisation it produces [GOL99].

2.5 Ontology

An ontology is the examination of *Being*, that is, it is an examination of what exists in the world. Philosophers propose that all human knowledge can be categorised into a predefined classification of knowledge types.

There have been numerous philosophers that have attempted to define ontologies. Such philosophers include Aristotle [BRE33], Husserl [HUS13], Whitehead [WHI29], Heidegger [HEI75] and Kant [KAN87]. In the context of this research, it is important to understand the concept of an ontology, rather than specific details about different ontologies that have been developed, therefore, a very brief explanation of Kant's ontology will be provided.

University of Cape Town

Quantity	Quality	Relation	Modality
Unity	Reality	Inherence	Possibility
Plurality	Negation	Causality	Existence
Totality	Limitation	Community	Necessity

Table 1: Kant's Ontology

Table 1 displays Kant's categories of understanding. According to Kant, the four sections define basic elements that constitute thinking:

- Quality – the extent of things
- Quantity – intension of things
- Relation – things in relation to one another
- Modality – the way in which things are related to us

Kant argues that these categories are innate, that people are never conscious of acquiring them. These categories are also never derived from the senses and are derived by reason alone.

University of Cape Town

2.6 Summary

In epistemology, a widely held definition of knowledge is *justified true belief*. This definition of knowledge stipulates that to know something, the belief must be true, and that one must be in possession of justification for its truth. The contention between different philosophers lies in different ideas as to what constitutes truth, and what forms of justification can be used to support that truth.

There exist beliefs that can be defined as *necessary truths*, for example, "all bachelors are unmarried". Self-evident truths can be called *epistemic immediacy* because they have no need for justification. Similarly, there exist *necessarily false* beliefs, for example, "I drive a colourless red car" that are obviously false without the need for proof of this. Beliefs that are neither necessarily true nor necessarily false are *contingent*.

One theory that attempts to alleviate the problem of justifying contingent beliefs is the theory of *reliabilism*. Reliabilism measures the truth of a belief according to the reliability of the process that either creates or justifies that belief. Reliabilism does not specify the processes that are reliable or unreliable, but instead defines a reliable process as one that "...leads to a sufficiently high preponderance of true beliefs over false beliefs" [SOS91].

There are five major processes that can produce and justify beliefs; *perception*, *memory*, *introspection*, *reasoning* and *testimony*. It is the reliability of these processes that is under dispute. For example, some philosophers yield to the fact that perception is a fallible processes but it would not be practical to ignore knowledge gained through this process. These philosophers therefore argue that knowledge can be possessed without complete certainty. On the other hand, deduction is a reasoning process that is widely accepted as being truth-preserving, but it has proved to be very difficult to justify all knowledge through deduction alone.

The fact that humans are social animals cannot be ignored in the context of knowledge acquisition. *Testimony* is the process through which one person gains knowledge from another. In the same way that the reliability of processes such as perception can be gauged, so too can the reliability of other people be gauged with respect to the quantity of reliable information these people produce.

University of Cape Town

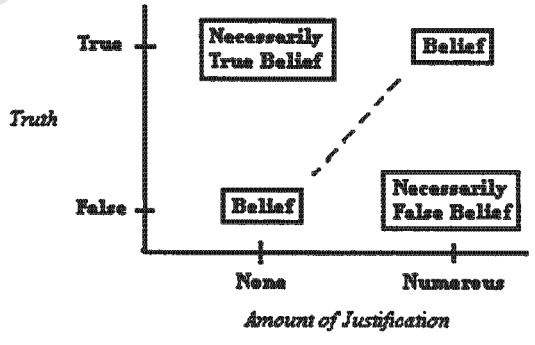


Figure 3: Truth as a function of the amount of justification it has received

Figure 3 presents how the truth level of a belief is dependant on the amount of justification the belief has received. Justification would be in the form of perception, memory, introspection, reasoning and testimony.

To conclude, it is important to emphasise that the objective of epistemology is not to understand how humans represent knowledge, nor is it to understand how humans re-use knowledge. The objective of epistemology is to investigate the essence of what a person can claim to know.

University of Cape Town

3 Knowledge Representations in Artificial Intelligence

Knowledge representation models in artificial intelligence are based around the knowledge representation hypothesis: "Any mechanically embodied intelligent process will be comprised of structural ingredients that: a) we as external observers naturally take to represent a propositional account of the knowledge that the overall process exhibits b) is independent of such external semantical attribution, play a formal but causal and essential role in engendering the behaviour that manifests that knowledge." [SMI82].

What the knowledge Representation hypothesis simply states is that there are two major properties that a knowledge structure needs to satisfy. Firstly, it must be comprised of propositions. That is, there must be a truth theory in the structure. The structure does not need to look like a proposition, but we need to be able to understand it in that way. Secondly, the symbolic structures within a knowledge-based system must play a causal role in the behaviour of the system [LB84].

Formal knowledge representations are primarily used for allowing automated reasoning processes to take place on information. It is possible for humans to make use of formal representations for the purposes of knowledge re-use, but they are primarily geared to towards re-use by computers.

The objective of this chapter is to present the most common formal knowledge representations that can be found in artificial intelligence, as well as to discuss pertinent issues in formal knowledge representations such as obtaining a balance between expressive power and computation efficiency, and discussing whether all formal representations are equivalent to each other.

Section 3.1will present the major formal knowledge representations. Section 3.2 discusses equivalences between the knowledge representations. Section 3.3 briefly describes some reasoning processes that can executed on the formal knowledge

University of Cape Town

representations discussed. Section 3.4 shows how meta-knowledge can be used to make formal knowledge representations computationally efficient. Section 3.5 concludes by comparing concepts found in epistemology with ideas in formal knowledge representations.

3.1 Main Approaches to Knowledge Representation

3.1.1 Logic Systems

Logic is widely accepted as the most expressive means of representing knowledge. Logic describes the world in terms of objects and their properties and relations. No limitations are placed on what can be regarded as an object; times, events, kinds, organizations, worlds, and sentences - not just concrete physical objects.

In logic based system, first order or predicate logic is used to represent knowledge. As a very simple example, if one wants to represent the statement "John knows Mary", one could use the logical expression *knows(John, Mary)*.

Logic has well defined semantics and proof mechanisms – one can understand what things mean and can make sound deductions from expressions. Logic is widely accepted as the most expressive form of knowledge representation, but is also understood to be very poor representation for computationally difficult problems [MOO82].

One of the most powerful ideas in logic is the ability to separate axioms from rules of inference. One can understand axioms without any need to know what type of inference will take place using these axioms, and similarly, one can understand an algorithm (or proof) without understanding the axioms on which it operates [WIN75].

3.1.2 Semantic Nets

A semantic net is a structure for representing knowledge as a pattern of interconnected nodes and arcs. Nodes represent concepts of entities, attributes, events and states, while arcs represent concept relations. Figure 4 illustrates a very simple example of a semantic network. Some relations represent linguistic constructs such as agent, patient, recipient,

University of Cape Town

or instrument, others represent spatial, temporal, causal and logical connectives, or even roles that one concept plays with another, for example mother, owner, residence [WOO91].

Some networks emphasise the ability to define new concepts, some are designed for representing natural language, some for artificial intelligence applications and others are designed to assert propositions and reason with the propositions, but in general, semantic networks are multi-purpose graph structures.

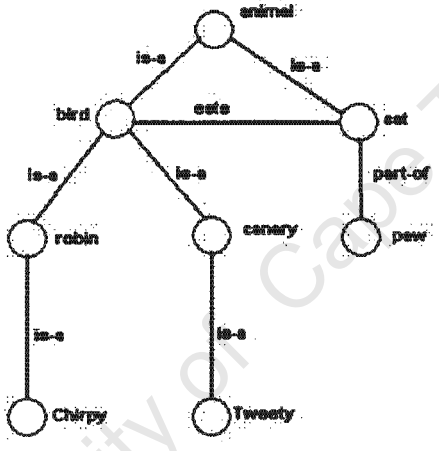


Figure 4: Example of a semantic network

In semantic networks, links are used for both recording of information and to serve as access paths by which facts are accessible to each other. Links are also used to connect individual elements that make up complex piece of information [WOO91].

Conceptual graphs (CGs) are graphical logic representation system that is based on semantic networks [SOW01]. Because of the ability to translate CGs to language, they serve as an intermediate language for translating computer formalisms to and from natural language. CGs have a form that is logically precise, humanly readable, and computationally tractable. With their graphic representation, they serve as a readable, but formal design and specification language.

University of Cape Town

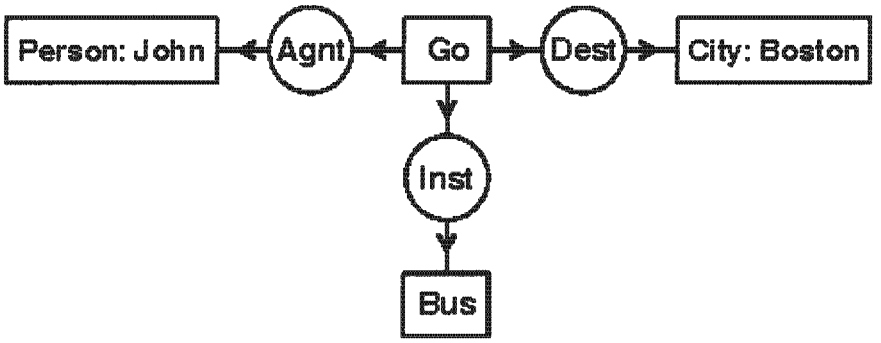


Figure 1: Concept graph for "John is going to Boston by bus"

Figure 1 shows a conceptual graph with four concepts: [Go], [Person: John], [City: Boston], and [Bus]. It has three conceptual relations: (Agnt) relates [Go] to the agent John, (Dest) relates [Go] to the destination Boston, and (Inst) relates [Go] to the instrument bus. In this way, this graph represents "John is going to Boston by bus". It can also be translated to the following formula in predicate calculus:

```
(?x:Go)(?y:Person)(?z:City)(?w:Bus)
   (name(y,'John') ? name(z,'Boston') ?
      agnt(x,y) ? dest(x,z) ? inst(x,w))
```

3.1.3 Rule Based Systems

Rule based systems emerged as one of the most practical methods of implementing knowledge based systems. Productions have two components, a premise ('if' part) and conclusion ('then' part). Given the rule "A implies B", if A is true then B is deduced to be true. Rule based systems are very powerful in exploiting modularity of knowledge and assuring that the correct knowledge is used in appropriate situations [SZO86].

The success of rule based system has been exemplified by systems such as MYCIN [DBS77]. MYCIN provides medical diagnosis of (and therapies for) infectious diseases. Davies, Buchanon and Shortlife point out that the domain for which the rule-based system is being created needs to have reached a certain level of understanding of basic principles and a recognized set of primitives. If there is no uncertainty in the knowledge then a different structure would be more efficient [DBS77].

University of Cape Town

3.1.4 Frames

Minsky strongly criticises purely logical approaches to modelling knowledge. He argues that "stating the facts" in terms of always correct, logical assumptions is very much harder than generally believed. "No logistic approach has had success with any type of large scale modelling task. Logicians are preoccupied with completeness and consistency, this is a major stumbling block for logistic approach" [MIN74]. Minsky believes that "thinking begins first with suggestive but defective plans and images that are slowly (if ever) refined and replaced by better ones" [MIN74].

This method of working has been substantiated by research such as [MEN99] by making observations about how experts work. Experts partially match current problems to libraries of accepted practices, implement an accepted test for adaptation of the libraries and modify the accepted practices accordingly [MEN99].

A frame is a structure for representing a stereotyped situation. Attached to the frame are several kinds of information; some about the use of the frame, some about what one can expect to happen next and some about what to do if these expectations are not confirmed. A frame is a network of nodes and relations. The top levels of nodes are fixed and represent things that are always true about a certain situation. The lower levels may have slots that must be filled with specific instances. Assigned to each terminal value could be conditions that must be met by that value. Each assignment is either a value or a pointer to another sub-frame. More complex conditions can specify relations among the things assigned to several terminals. Frames have thus successfully managed to mix procedural and declarative knowledge [WIN75].

Terminals are usually filled with default values when no better value can be found (or is available). Default assignments try to model the way a human generalises, creates analogies and makes judgements.

Collections of related frames are linked together into frame-systems. Different frames of a frame system may share terminals. This is a critical characteristic of frames that is vital

University of Cape Town

to the co-ordination of information. Frame systems are linked to each other by an information retrieval network. When a proposed frame cannot be made to fit reality, then this information retrieval network provides a replacement frame. Once a frame is proposed to represent a situation, a matching process tries to assign values to each frame's terminals. This matching process is partly controlled by information contained in the frame and partly by knowledge of the system's goals.

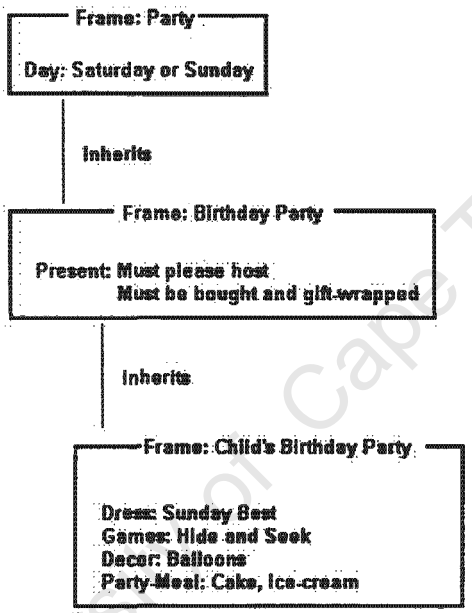


Figure 6: Example of Frame hierarchy for a stereotypical party

RX is an implementation of a clinical database that makes use of frames to implement an inheritance hierarchy [WBW86]. RX contains four different types of knowledge; rules, categorical, definitional and causal. The database is time based and thereby shows patient state over time. Categorical knowledge is implicit in the data model due to the position in the hierarchy where a piece of information can be found. Definitional knowledge provides deductions going up the levels of hierarchy and are used as an alternative means of computation for dealing with missing values. Causal knowledge are links and dependencies among nodes in the hierarchy. The essence of the RX system is to keep track of causal effects of medication being prescribed to patients and the effects of this

University of Cape Town

medication. The causal effects are not single values, but are made up of multiple essential values as well as other optional values. Examples of the types of values used in RX are frequency, intensity, direction, setting, distribution and evidence. Causal effect links can form very long chains that describe a patient's state over time.

3.1.5 Connectionist Theory and Neural Nets

Connectionist theory makes use of fundamental processing units (such as abstract neurones) that are connected to many others via channels. These channels have weights that relate to how much influence one unit has over another. Connectionist theory assumes that all knowledge is stored in the connections between the neurones. In this way, knowledge is implicit in the structure, rather than explicit in the states of the units themselves [RUM89].

There are seven basic components of a connectionist framework [RUM89]:

- set of processing units.
- state of activation over time - vector describing the state of each unit in the whole system
- output function
- pattern of connectivity
- activation rules – which channels to activate at what time
- learning rule – describes new, old or modification of connections
- environment - stochastic model over time and space of possible input patterns

Connectionist models are efficient at solving constraint-satisfaction problems, in best-match search and pattern recognition. They are a general mechanism for adaptive learning and exhibit graceful degradation with damage or information overload.

3.2 Equivalence in Knowledge Representations

It can be argued that if a formal knowledge representation can be used to allow reasoning processes to take place on the information contained inside it, then it is in some way equivalent to other formal knowledge representations. It would be difficult to understand

University of Cape Town

how a non-logical system could make assertions , as any system that makes deductions of some kind is logical in one way or another [MOO82].

The arguments that logic cannot successfully represent human working processes [MIN74] has been counter-argued by work such as [HAY79], that shows how models such as frames are equivalent to logic representations. They translate KLR-0 [BW77] into predicate logic and show how logic can be used to represent default reasoning as well as heuristic information. [HAY79] argues that frames are a suggestion about how to organize large amounts of data rather than a new representational model. We should note however that comparing frames and logic in terms of representational adequacy is not sufficient in arguing against the usefulness of frames. Unlike logical expressions, frames are easier to understand and therefore easier to use and are more inferentially efficient than logical expressions.

"Representational and computational strategies employed in semantic nets are abstractly equivalent to those employed in all state-of-the-art systems incorporating a substantial propositional knowledge base, whether they are described as logic based, frame based, rule based, or something else" [SCH91]. Schubert argues that graph theoretic propositional representation does not distinguish semantic nets from other representations. Predicate calculus, "proximity-based" inference and inheritance are no longer unique to semantic nets due to their existence in logic, rule and frame based systems [SCH91]. Clustering of properties and inheritance were once the most significant features of semantic nets, now they have been "incorporated into *all* knowledge representation formalisms that are designed to cope with substantial bodies of propositional knowledge" [SCH91]. Frames encode clusters of information about entities and are organized into hierarchies. Hierarchies have also been integrated into databases, expert system shells and PROLOG extensions. Clustering has been implemented by systems that would not be called frame-like or net-like systems.

It could be argued that semantic nets are unique not because they implement characteristics such as inheritance or clustering, but rather in the way they promote their

University of Cape Town

use to control inference within the net. Schubert argues that this sort of proximity reasoning has been implemented by systems that are not semantic nets; that any reasonable inference control mechanism will inevitably be proximity-based [SCH91].

The fact remains that despite all the work done to show how formal knowledge representations are in many ways equivalent to each other, real world problems will inevitably be solved by representations that are most efficient and easily understood. Seldom will one representation technique be chosen over another because it can represent more, it is often the case that a representation technique that is capable of representing the minimum amount needed to solve the problem and is efficient is chosen over a very representationally powerful one.

3.3 Reasoning with Knowledge Representations

Representations and reasoning techniques are both necessary in artificial intelligence; if one only deals with representation, then the system will not have the ability to reason. Epistemology deals primarily with "what is a person justified in believing?" rather than "how does a person reason?", therefore reasoning is not central to this dissertation and this brief outline is included for completeness only.

A typical problem space that a reasoning system would deal with would be a parametric problem. Parametric problems are problems where solutions can be expressed as values for a given set of parameters. The solutions usually need to conform to a number of constraints and requirements, one of these being an optimisation requirement [ZM96]. Parametric problems can exploit preferences for certain parameters. These preferences can have an effect on local and global optimums of solutions.

Reasoning in logic based systems - Logic has well defined semantics and proofs; one can understand what things mean and can make sound deductions from the information. Non-monotonic inference allows one to make inferences based on the lack of a piece of information. Some logics that provide a basis for non-monotonic reasoning are default

University of Cape Town

logic, abduction and inheritance [RK91]. Default logic allows one to reason about information where a general case exists.

Reasoning in rule-based systems - There are two basic types of deduction techniques that can be used on rule bases: forward chaining and backward chaining. Forward chaining is driven by the input (or known facts) and rules, while backward chaining is driven by an expected conclusion (or goal) [SZO86].

Probabilistic reasoning - The main idea of probabilistic reasoning can be summed up as "... the likelihood of classes of objects or events is inferred by propagation of belief values in the light of changing evidence" [BUX97]. Probabilistic reasoning has a number of uses; integration of multiple viewpoints, allowing representations of contexts, for control, and for learning. It is useful for representing answers where it is not computationally feasible to calculate a definite/exact answer.

There are a number of different probabilistic theories that have been proposed. These include Bayesian theory, Certainty Factors, Bayesian Networks, and Dempster-Shafer Theory. Bayesian theory is intractable because too many probabilities need to be provided and people are poor estimators of probability, and computational time is too large. Bayesian networks assume that it not necessary to take into account all possible combination of events. Most events are conditionally independent of each other, so those relationships do not need to be computed. The theory only computes "clusters" of events that interact with each other [RK91. In Dempster-Shafer theory, instead of assigning each individual element a probability, we assign probabilities only to possible conclusions and analyse how much information there is for one to disbelieve a certain hypothesis [RK91].

Case based reasoning - Researchers in Case Based Reasoning (CBR) maintain that people rarely solve problems from "scratch", rather, people tend to use a "reflective" approach by continually testing and changing accepted problem solving practices [MAKM99]. In CBR, when a certain situation or problem arises, one chooses the most similar historical case and one makes a decision according to this historical case.

University of Cape Town

Circumscription - Circumscription is a method of reasoning that "jumps to conclusions". It simply produces conjectures that may be wrong or even produce inconsistency. It is argued in that humans use circumscription and therefore this is why humans are susceptible to errors. An intuitive idea of circumscription is as follows: We know some objects in a given class and we have some ways of generating more. We jump to conclusions that this gives all the objects in the class. Thus we circumscribe the class to the objects we know how to generate [McC77].

3.4 Inferential Efficiency and Meta-Knowledge

Not only should the representation allow for inference, but the inference mechanism should be focussed in the most promising directions. Computational tractability can be improved by using control knowledge [MYL86]. That is, knowledge that helps a reasoning process decide on what strategy to take when trying to solve the problem. This type of knowledge would therefore limit the number of choices in a large solution space and make difficult problems more solvable.

[DSS93] and [LB84] argue that computational efficiency is a trade-off with expressive power. First order logic is very expressive but efficient theorem provers do not exist due to the lack of control structures for "next assertions" [MOO82]. Databases are a type of limited knowledge representation. They are limited in terms of what they can express but they have great computational power. Semantic nets allow computation with the use of graph searching algorithms because they can be represented by directed graphs. Frames have similar computational characteristics as semantic nets but offer some more constraints on the values: restrictions (constraints to the attributes) and attached procedures (if-needed and if-added) that make computation more efficient.

It is generally accepted that it is computationally intractable to search a large knowledge space unless some degree of knowledge about the world and/or problem domain can limit the search space [RK91] [dKDSS77]. There are a number of search criteria that could potentially alleviate this search space problem. Such search criteria include knowledge

University of Cape Town

about which states are preferable to others, knowledge about what rule/action to apply in certain circumstances and situations, knowledge about the order in which to pursue sub-goals and knowledge about useful sequences in which to apply rules [RK91]. This type of knowledge is called *meta-knowledge*. For a system to be able to reason about its goals, actions and reasons for belief, these must be explicitly represented in a form that can be manipulated by the deductive process of the system [dKDSS77]. This formal "internal" control is a problem domain formalized using assertions and rules just like an "external" domain [dKDSS77].

[dKDSS77] presents a language AMORD that contains control structures for such "internal" domains. Instead of forward and backward chaining, AMORD uses only forward chaining over the states of the problem solver which are explicitly represented over and above the information about the domain. This explicit representation of problem solver states gives the system basic reflective properties. Based on currently asserted goals and antecedent rules that state how information should be used (or how goals can be achieved), the system can decide on what new goals to pursue. By maintaining justification among facts and among goals, it can determine which assumptions have lead to contradictions. The AMORD system makes occasional assumptions that may prove to be false later; these assumptions are seen to be necessary for efficient search.

MYCIN [DBS77] also contains meta-rules for speeding up processing. These meta-rules tell the system which rules to try first, last, before others or not at all. The authors note that there is no need to stop at a single level of meta-rules and go on to define the type of control knowledge that would be contained in higher levels of knowledge: 2nd level meta-rules would contain information about what strategies to try in certain cases, 3rd level meta-rules would contain information about how to select a strategy to try, etc. The authors do note however, that in the implementation of MYCIN, these further meta-levels where not implemented because there was no need to improve performance any more. This could be because the system contained only a small rule set. For larger rule-sets it may be necessary to implement such meta-knowledge.

University of Cape Town

Meta-knowledge can also be used as a means of representing *stable knowledge*. Stable knowledge is knowledge that will change infrequently and does not require much maintenance. Such knowledge would act as a framework for other less stable knowledge. Some people believe that meta-knowledge complicates knowledge maintenance because they believe that there is no such thing as stable knowledge. Therefore, one ends up having to maintain not only the actual knowledge, but also the meta-knowledge. Menzies argues that meta-knowledge is not useful if it helps in system construction but complicates system maintenance [MAKM99]. The XCON system that is used to configure DEC computers has thousands of rules of which more than half need to be changed annually. Garvin ES-1 is a system used to interpret biochemical results and needs large-scale maintenance every year. Using these examples, Menzies argues that there does not seem to be anything that can be regarded as stable knowledge and therefore nothing like reusable meta-knowledge can exist [MEN99].

3.5 Comparing Epistemic and Formal Representations of Knowledge

Expert systems are concerned with automated problem solving. Research in the field has primarily focussed on representing experts' knowledge in the most efficient manner with which a computer can solve domain-specific problems. Historically, expert systems have been more concerned with questions like "How would an expert solve this problems?" and less with questions like "What does an expert really *know*?". These are fundamentally two different questions; the first is an investigation of human thought processes, while the second is an analysis of what knowledge an expert has. So despite the fact that representations such as logical expressions, semantic networks, frames and connectionist models have all proved to be very useful models for solving different types of problems, none of them attempts to answer the philosophical question of what a person justified in believing.

In the same way that epistemology was historically grounded in rationalist views, so too were artificial intelligence models initially concerned with representing *true* beliefs in structures such as logical statements. Over time, theories of knowledge that do not

University of Cape Town

require *completely certain* beliefs have proved themselves to be practical alternatives to the extreme rationalist standpoint on the nature of knowledge. In the same light, some knowledge modellers in artificial intelligence systems have started to see the usefulness of representing uncertain knowledge, while others choose to represent knowledge in less expressive formalisations that are computationally more efficient.

Humans' knowledge of nature is constantly evolving and refining itself, there are very few (if any) domains of expertise that can claim to have complete knowledge of their field. This has forced representations such as rule-based systems to allow for uncertain knowledge, and has forced the development of models such as frames and neural networks that have uncertainty inherent in the models themselves.

MYCIN [DBS77] uses the idea of Certainty Factors (CF) in a rule-based system. Evidence confirming a hypothesis is collected separately from evidence that "refutes" it. The CF of the hypothesis is the sum of the current evidence *for* and *against* it. If a CF=1, the information is definitional. Therefore, the MYCIN system can offers its user a value describing the probable truth of the medical diagnosis it has just made.

Frames represent stereotyped situations that evolve over time. From an epistemic point of view, frames represent knowledge that is at first "unclear", and becomes clearer and clearer as more evidence for a specific situation is obtained. Therefore frames represent a type of certainty without the need to represent actual probability values. The certainty (or accuracy with which the frame depicts a real world situation) can be derived from its position in the inheritance hierarchy.

Neural networks are structures that depend on probability theory to represent knowledge to such an extent that they contain no propositional beliefs at all. All knowledge is implicit in the strength of the arcs connecting the nodes. Results of certain types of neural networks are a set of probabilities that represent possible values as a result for a problem.

University of Cape Town

Even the most rigid of all knowledge representations, logic, has had numerous extensions applied to it over the years. Most interesting is the development of fuzzy logic. "Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth -- truth-values between *completely true* and *completely false*" [KHJ00].

It is not only truth-based epistemology that has had an influence over expert systems. The RX system [WBW86] represents a *causal theory of knowledge* by being able to show the causal effects of medication taken by patients over a period of time.

Knowledge contained inside an expert system represents the knowledge as that of one expert in a specific domain. In many cases, however, this single body of knowledge has been formalized after a lengthy process of formalising consensus knowledge from many different experts. Because of one's perception of the levels of expertise of different people, it is clear to see how allowing problem-solving heuristics the ability to combine different schools of thought can play a vital role in producing results that are a combination of trustworthy sources. Probability theory has not only allowed artificial intelligence to represent uncertainty of a result, but has also allowed the representation of information that is both justifying and defeating for a proposition, and thereby forming deductions based on different viewpoints and opinions.

University of Cape Town

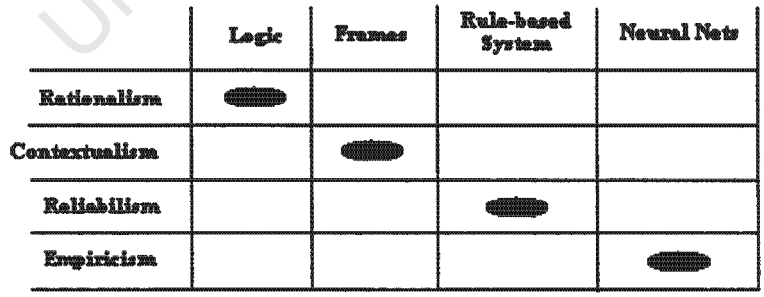


Figure 7: Epistemic view of formal knowledge representations

One can argue that many of the formal knowledge representations discussed in this chapter could be classified under numerous epistemological theories, however, Figure 7 attempts to classify these knowledge representations under the epistemological theories they most represent:

- *Logic* represents rationalism through the dependence on deductive reasoning to produce new knowledge.
- *Neural Networks* are a typical example of empiricist views because they are trained through input-output observations of functions
- *Rule-based systems* could be viewed as a type of reliabilism because the individual creating the rules of the system is viewed as a reliable source of knowledge in the relevant field.
- *Frames* contain beliefs that can be held as *true* within a context of the frame.

So with the development of formal knowledge representations that have the ability to represent uncertainty of knowledge, artificial intelligence research has in many ways followed the lead of epistemology by allowing for a more pragmatic view of an ever-changing and uncertain world.

University of Cape Town

4 Corporate Memory Models

"A corporate memory is an explicit, disembodied persistent representation of the knowledge and information in an organization" [vHvdSK96]. Therefore, to improve the competitiveness of an organization, any piece of information that contributes to the performance of an organization should be stored in the corporate memory. "This includes knowledge about products, production processes, customers, marketing strategies, financial results, strategic planning and goals, etc." [vHvdSK96].

This chapter is focused on *Corporate Memories*, the mechanisms used for storage of information during knowledge management processes. Section 4.1 discusses the need for representing knowledge in free-form manner to allow for more effective re-use by humans. Section 4.2 outlines the advantages and disadvantages of representing knowledge declaratively or procedurally. Section 4.3 presents a list of requirements to which corporate memory systems much comply. Among the most important requirements that will be discussed in detail are: the need for corporate memory systems to integrate into the workplace, the need for representing multiple viewpoints held by different people, the need for meta-knowledge and the need for information classification.

Section 4.4 presents some general techniques used to facilitate information retrieval, and Section 4.5 outlines methodologies that can be used for maintaining information. Because the approach taken in this chapter is to outline the major issues involved in representing knowledge of an organization and show how systems have used different approaches in dealing with these issues, Section 4.6 presents a specific example of a corporate memory system to illustrate how such systems typically look to the user. Section 4.7 concludes by comparing work done in corporate memory systems to theories of knowledge found in epistemology.

University of Cape Town

4.1 Free-Form Knowledge Representations

Free-form knowledge representations are primarily geared towards storage and retrieval of knowledge for human re-use. They permit items of knowledge to be of any size and any format. This is unlike the formal knowledge representations discussed in Chapter 3 (*Knowledge Representations in Artificial Intelligence*) that store small units of knowledge for use by automated problems solving processes.

Much of the work done in the area of free-form knowledge representation is done in the context of human re-use of knowledge. The reason for this is understandable, humans already have an ability to comprehend information and use information to solve problems without having the information represented in a formal way. Therefore, is seems that most systems that have been built using informal knowledge representation have ended up being document management systems that facilitate human learning, but are not at all suited for use by deductive algorithms for problem solving.

[EUZ96] highlights the importance of having the ability to model both formalized information as well as free-form information in a single system. It is pointed out that in a corporate memory environment it is difficult to consider formal knowledge in isolation from other sources of knowledge such as bibliographic references, research papers and experimental data. [EUZ96] suggests that free-form knowledge should be used to annotate formalized knowledge to explain (for example) why new or changed information has been accepted or rejected by the knowledge base.

A number of problems with formalising a knowledge base are discussed in [BUC97b]. If a formalized approach is taken too soon in the development of a knowledge-based system, some problems may occur when more data is inserted in the structure and the structure evolves and becomes "less tolerant to change" and therefore less flexible. Furthermore, an incorrect model is less likely to be corrected due to the complexity of the task. [BUC97b] argues that free-form information storage is very important for modelling knowledge, and that only stable, sanctioned knowledge should be formalized into more rigid structures. Formal representations require knowledge to be filtered so as to "make it

University of Cape Town

fit" into a formal structure. In this filtering process much of the valuable difficult-to-formalize knowledge could potentially be lost [BUC97b] [ZM95]. It is important to allow easy addition of information into a knowledge base; formal information can only comfortably be accessed by individuals trained to manipulate formal constructs, and therefore representations can become less comprehensible. Automated analysis of formalized data could be detrimental to the system because valuable information could be stripped out by automated agents that have no understanding of the information or the context in which information is being collected.

It does seem however, that a certain degree of formality is necessary for effective use of free-form knowledge. Despite the argument of the most avid opponents to formal representation of knowledge, most agree that important pieces of information should be formalized after a significant amount of experience has been gained through its usage [MAKM99]. Section 4.6 (*An Example Corporate Memory System*) below describes, in greater detail, a system which makes use of documents for knowledge storage, but makes use of a formal framework around which to describe relationships between items of knowledge.

University of Cape Town

4.2 Declarative versus Procedural Knowledge

One of the more common distinctions made in knowledge representations is between representation as a process and representation as a product of a process [SR96]. Processes concern themselves with the transformations of ideas and preservation of these transformations in the representation. Descriptions of a product of a process are more concerned with structural characteristics of the representation. These two different types of representations are commonly called procedural and declarative representations respectively.

Proceduralists model knowledge by modelling *how* knowledge is gained, while declarativists model the actual information (knowledge *that*) [WIN75]. With declaratives, one knows different things by knowing more things, whereas with procedural the knowledge is embedded in the procedure itself.

Some of the benefits of representing knowledge declaratively are [WIN75]:

- there might be many possible uses for a piece of information. This is easier to represent as a declaration than as many procedures.
- if the knowledge is a set of independent facts then any change to a fact only has implications on that fact, and nothing around it. That is, information is self-contained.
- much of what we know can be stated simply in declarative form. Therefore, when communicating information one uses a declarative form.

Some of the benefits of representing knowledge procedurally are [WIN75]:

- in using procedures, one trades off a degree of flexibility for a gain in ease of representation. Procedures make representation of decision making processes much easier and more compact.
- using procedures allows one to represent second-order knowledge. This would be very difficult to do using declaratives unless one limits oneself to a very specific domain.
- much of what one knows about certain domains is heuristic, not declarative. Heuristics are easier represented using procedures.

[MH99] notes two advantages of procedural knowledge modelling in corporate memories. Firstly, office workers are accustomed to working with processes in their day-to-day tasks. Knowledge is needed for task execution rather than just knowledge's sake. The second advantage is that managers and team members are mostly very comfortable using processes such as project plans for task execution.

The EULE2 system [REI98] is an example of a corporate memory that models working procedures as processes. Office processes are represented as graphs. Each node represents a task and arcs are choices that need to be made by the user. The system contains representations of office tasks consisting of a set of actions and effects of those actions, instances that are manipulated by the tasks, and the business concepts they belong to as well as the laws and regulations that must be obeyed by the office tasks.

University of Cape Town

One cannot say that representing knowledge as an object is more correct than representing knowledge as a process, nor vice-versa. In different domains one might find different techniques more useful than others. In fact, it would be difficult to find a domain of expertise where only one representational technique is needed. In most cases, claiming to have expertise in a subject means that one knows not only about the characteristics of objects in that field, but also about processes that create and manipulate such objects.

4.3 Corporate Memory Requirements

Like other software systems, corporate memory systems have their own unique set of requirements. However, because corporate memory systems deal with a raw material that is very difficult to define and manage, the requirements outlined for corporate memory systems vary substantially from person to person. Below are a number of different sets of requirements that have been described in corporate memory literature, followed by a new list that attempts to consolidate all of these.

University of Cape Town

[ABH+98] outlines five requirements for corporate memories:

- Ability to collect information from various sources
- Need for minimum of up-front knowledge maintenance
- Ability to exploit user feedback for maintenance and evolution
- Effective integration into existing work environment
- Active presentation of relevant information

[BUC97a] outlines four requirements for a knowledge representational tool:

- The representation must be capable of representing multiple viewpoints.
- It must represent changes in priorities.
- It should be able to communicate changes to users
- It should be able to retrieve previous work done.

According to [SC00], a corporate memory system should have the ability to:

- Capture knowledge,
- Communicate knowledge
- Refine knowledge
- Annotate knowledge
- Classify knowledge
- Retrieve knowledge
- Navigate through knowledge
- Be multi-modal - knowledge should be available through many different sources having many different medias.

[vHvdSK96] examines yet a different set of requirements for knowledge systems:

- It should be easy for workers to decide which other workers could have knowledge about a certain domain.
- It should be easy and rewarding for workers to submit lessons learnt into the knowledge base
- It should be easy for workers to access the knowledge for learning by combining information from different parts of the knowledge base
- There should exist well defined criteria for deciding if the information is useful
- There should exist well defined criteria for deciding how knowledge should be formulated
- There should exist well defined criteria for deciding where knowledge should be stored

Other features that have come to the forefront of knowledge management in recent years include:

- Ability to function in environments that are comprised of geographically dispersed workers, distinct knowledge bases and knowledge at different levels of formalisation [SIM96].
- Facilitate both synchronous and asynchronous modes of communication [SC00].

University of Cape Town

- Ability to store contextual and cultural information about knowledge [WF87] [CSSvH96] [EAS93].
- Store user profiles [vHvdSK9696].

Below is a list that attempts to combine all the requirements outlined above into one coherent set of requirements for a corporate memory system. The corporate memory system should allow for:

- Knowledge capture from multiple knowledge sources, at different levels of formalisation, using both declarative and procedural techniques
- Storage of user profiles that describe an individual's expertise, interests and goals
- Communication of knowledge both synchronously and asynchronously to relevant users according to the user profiles
- Refinement of knowledge by exploiting user feedback that captures accuracy and usefulness of information
- Annotation of knowledge
- Storage of contextual and cultural information
- Storage of higher levels of knowledge – meta-knowledge
- Classification of knowledge into categories to simplify knowledge retrieval
- Retrieval of knowledge through search mechanisms or through a navigational approach.
- Interaction with users through a variety of different sources, for example, internet, mobile phones, etc.
- Representation of multiple viewpoints
- Integration into the workplace, including the ability to function in a geographically and temporally disparate organization

The remainder of this section looks at some of these requirements in more detail to show problems and solutions adopted by corporate memory systems.

University of Cape Town

- Ability to store contextual and cultural information about knowledge [WF87] [CSSvH96] [EAS93].
- Store user profiles [vHvdSK9696].

Below is a list that attempts to combine all the requirements outlined above into one coherent set of requirements for a corporate memory system. The corporate memory system should allow for:

- Knowledge capture from multiple knowledge sources, at different levels of formalisation, using both declarative and procedural techniques
- Storage of user profiles that describe an individual's expertise, interests and goals
- Communication of knowledge both synchronously and asynchronously to relevant users according to the user profiles
- Refinement of knowledge by exploiting user feedback that captures accuracy and usefulness of information
- Annotation of knowledge
- Storage of contextual and cultural information
- Storage of higher levels of knowledge – meta-knowledge
- Classification of knowledge into categories to simplify knowledge retrieval
- Retrieval of knowledge through search mechanisms or through a navigational approach.
- Interaction with users through a variety of different sources, for example, internet, mobile phones, etc.
- Representation of multiple viewpoints
- Integration into the workplace, including the ability to function in a geographically and temporally disparate organization

The remainder of this section looks at some of these requirements in more detail to show problems and solutions adopted by corporate memory systems.

University of Cape Town

4.3.1 Representing Different Types of Knowledge

One can differentiate between two types of knowledge that exist in organizations; *Explicit knowledge* is in the form of books, documents, emails and other media, and *Tacit knowledge* as the accumulated practical skill and expertise of an individual [vHIP88]. Explicit knowledge can be stored in system, but tacit knowledge is less tangible.

Knowledge can be found in various forms: *non-computational* – knowledge comprised of paper based documentation, *document based* – electronic documents that may or may not be indexed, *knowledge-based* – knowledge made up of formal representation of an expert's knowledge, *case based* – knowledge made up of historical cases of successes and failures to enable users to exploit past experiences [DCGR98].

	Kinds of mental material				
	Semantic	Episodic	Skilled	Tacit	Unconscious
Representations that highlight abstract characteristics along specific dimensions					
Charts	*		*	*	
Graphs (i.e. Scaled consecutive values)					
- Two dimensional graphs	**		*	*	
- Three dimensional graphs	**		*	*	
- Radial graphs	**		*	*	
Graphical abstract representations of complex relationships					
Process diagrams					
- Flowcharts	**	**	**	*	
- Critical path diagrams	**	**	*	*	
- Soft systems methodology root definitions	***	*	**	**	
Network diagrams					
- Informal					
• Mind maps	*	*	**	**	*
• Rich pictures	*	*	**	**	*
- Formal					
• Semantic Networks	**	*	**	*	
• Tree Chart	**	*			
• Cognitive Maps	***	**	**	*	
Venn Diagrams	**		*	*	
Representations that emphasise wholeness and integrity					
Visual analogies/metaphors	**	**	**	***	**
Static images					
- Geographical maps	**	**			
- Diagrams, imaging, drawing and photographs	***	***	**	***	***
- Physical models	***	***	*	**	*
Dynamic Representations					
- Animations	***	**	***	**	*
- Film/video	***	**	***	**	

Table 2: Some postulated strengths of particular visual representations for different kinds of knowledge [SPA98]

University of Cape Town

[SPA98] defines knowledge as a combination of *mental material*, *forms of thought* and *types of thought*. Mental material is comprised of *semantic understanding* – common sense, *episodic* – memories one holds, *skills* – skills that form the essence of one's job, *tacit feel* (cognitive stimuli that are difficult to express) and *unconscious interpretation* – interpretation at a non-conscious level. Forms of thought include: *propositional* – seeking for propositional nodes used during experiences, and *imaginistic* – mental imagery. Types of thought include: *autistic* – non-rational view of events, *reasoning* – logical reasoning of a problem, and *mood* – states of mind that affect interpretation of events.

Furthermore, [SPA98] postulates strengths of different representations according to the types of mental material they represent. From Table 2 above, it is clear how different representations have strengths and weaknesses for representing different forms of information.

4.3.2 Workplace integration

Knowledge can reside in many different places; bookshelves, filing cabinets, desks, hard drives, etc., but from a user's perspective the information needs to be accessible from a single point of entry, independent of user location [TAHZ00]. In a geographically dispersed workforce, workers do not have the opportunity to share and compare knowledge. This may result in multiple knowledge bases that may need to be integrated at a later stage to enable an organization to obtain a global shared knowledge environment.

Knowledge Management tools need to interface directly with tools being used in the workplace [KA97]. They must also interface with existing workflows in the organization [CON96] [SC00]. The need for workplace integration of corporate memory tools is fueled by the fact that employees' satisfaction of a corporate memory system is very much dependant on the ease of information retrieval, the quality of results obtained, and the applicability of the result to the user's needs [DCGR98].

University of Cape Town

The EULE2 system [REI98] mentioned in Section 4.2 is designed to integrate into the workflow environment of an organisation. The workflow software would make calls to the EULE2 system as illustrated in Figure 8 below.

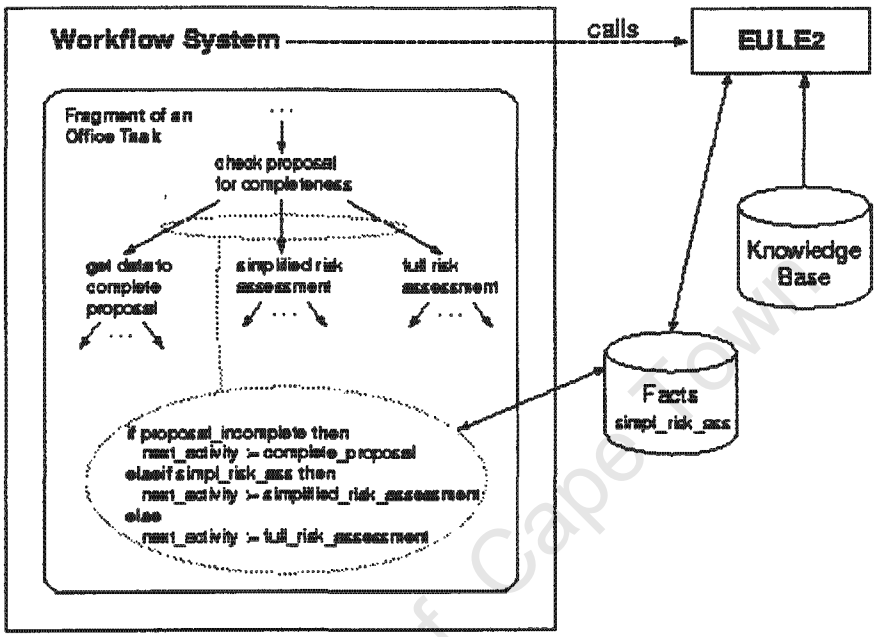


Figure 8: EULE2 system integration with workflow systems

Because EULE2 plays a key role in the workflow of information, the knowledge based would be updated through the normal usage of the workflow environment.

4.3.3 User Profiling

[vHvdSK96] emphasises the need to maintain user profiles. Such information will help the system to understand what types of information certain employees are interested in and permit workers themselves to be able to control the type of information they receive based on their own interests and goals. Responses given by the corporate memory to questions about certain topics should be given in the context of such goals [WEB86].

[SC00] describes a method of user profiling in which departments in the organization would be responsible for defining their own user and departmental processes and thereby defining email lists for specific interest groups. Each department would have a single

University of Cape Town

[EFKN94]. Conflict should be treated as an important part of a domain and needs to be represented and understood [EAS93].

[SG88] identifies four types of comparisons between expert concepts:

- Consensus – experts use the same terminology to describe the same concept,
- Correspondence – experts use different terminology to describe the same concept,
- Conflict – experts use the same terminology to describe different concepts,
- Contrast – experts use different terminology to describe different concepts.

The above conflict types can be resolved by the following types of conflict resolution strategies [EAS91]:

- Cooperative methods – include negotiation and education
- Competitive methods – include combat, coercion and competition
- Third party methods – include arbitration, appeals to authority.

One of the primary reasons for conflict has been found to be due to *ontological drift*. As ideas pass through different subgroups they are interpreted in terms of that particular group's set of meanings, which more often then not, do not correctly map onto the meanings used by other groups [EAS93]. Ontological drift can also be resolved by collaboration between communities. Other reasons why conflict could arise would be insufficient exchange of information, differing goals and requirements of different members of a group and personal factors such as differing value systems and personality characteristics [ROB89], [EAS91].

In the software development field, a certain level of understanding is established through the production of key documents during the software process, which define concepts and tasks through a series of publicly examinable specifications. Such documents prevent ontological drift. Therefore, by externalising ideas, participants in a discussion can share views with one another and develop understanding of each other's perspectives [EAS93].

University of Cape Town

[EFKN94] presents a knowledge based system that is implemented in a decentralised environment. This type of decentralised environment is seen to be problematic because of the difficulties of "maintaining consistency between large collections of agents" [EFKN94]. The system overcomes this problem by recognising that it is desirable to tolerate (and sometimes even encourage) inconsistency to maximise freedom of expression and to prevent premature commitment to certain decisions.

In [EAS91] a free-form, document-centric environment uses free-form annotations to allow experts to express their view about information that others have presented. Four types of annotations may be attached to items: comments – general purpose annotation, assumptions – note where a description appears to have made some unstated assumption, issues – points that need to be addressed, justification – added support to a particular viewpoint or proposal.

An example of a tool that implements a co-operative method of conflict management is QuestMap [BUC96]. This tool stores many different types of media either *for* or *against* an argument. These pros and cons then lead to further pros and cons until many of the arguments are negated and an agreement or decision is reached. An example screen-shot of the QuestMap system can be seen in Figure 9 below.

[ROB90] makes use of an arbitrator to evaluate preferences expressed by various perspectives and to guide the search for new solutions. The arbitrator needs to search for combinations of proposals that increase the satisfaction of all concerned. This process is used to identify an ideal, but usually unattainable goal. Such a goal could stimulate other combinations that come close to the necessary goal.

University of Cape Town

the Co4 system, [EUZ96] also presents a knowledge base maintenance methodology / protocol that implements the idea of editorial boards that need to approve information prior to it being added to a group base. In the Co4 system, knowledge is only accepted into a group base if the information is accepted by all the submitter's peers. It is pointed out that this relatively extreme idea of "consensus" can be replaced by other more lenient criteria such as "majority" or "intersection".

4.3.5 Meta-Knowledge

Meta-knowledge is *knowledge about knowledge*. Meta-Knowledge can be divided into two categories, *contextual* and *semantic* [SDB00]. Contextual meta-knowledge is information that creates a context for the knowledge, for example: creator, title, location, modification date, and history. Semantic meta-knowledge is information about the knowledge itself, for example keywords, abstracts, and text annotations.

Among the more concrete types of information that should be maintained about knowledge are time, cause and action. The representation should furthermore have the ability to inform about [MIN74]:

- what caused certain knowledge (agent),
- what the purpose of the knowledge was (intention),
- what the consequences are (side-effects),
- who the knowledge affects (recipient),
- how the knowledge is implemented (instrument).
- why the system came to know this information (intentionality),
- where to find more information about an idea (relational information),
- hiding of one fact by another fact (occlusion),
- explanations of objects by their function, not their physical structure (functional variants)
- objects are described in contexts that should automatically change the perception of the object (parasitic contexts)

University of Cape Town

The need for the above information is substantiated by [BUC96], who also adds the following types of information that would be useful in a decision making environment:

- have similar problems been faced?
- who suggested the solution?
- what solutions were considered and why were they rejected?
- what criteria were taken into account in a certain decision?

Many corporate memory systems resort to modelling meta-knowledge as a means of describing the information in its context. CommonKADS [dHBM+94] uses the following knowledge items in its description of information:

- General
    - Name - general name of this piece of knowledge
    - Role Description - the role the knowledge is associated with
    - Activity - the related organizational task
    - Domain(s) - references to organizational areas/objects/procedures
- Content
    - Generic task type – generic task from CommonKADS library
    - Nature – heuristic, formal, uncertain
    - Products/Services – marketable products of the organization
    - Functions – organizational functions involved
- Availability
    - Time – when the knowledge is available
    - Location – where the knowledge will be available from
    - Form – form of the information (paper, electronic, mind, collective)

[vHvdSK96] does not agree that all the above information needs to be maintained about a certain piece of knowledge. He suggests a simpler set of information that needs to be kept:

- Activity – the organizational activity that a piece of information refers to.
- Domains – the subject of the knowledge item
- Form – the form in which the knowledge resides (paper, electronic, mind)

University of Cape Town

- Type – the type of knowledge, for example, procedure, handbook, manual, best practices, worst practices, progress report, failure report, comment, etc.
- Products / Services – The product or service this knowledge refers to
- Time / Location – Time and Location that the author of this knowledge will be available.

Some example attributes for meta-data taken from the KARAT [TAS97] approach for representation of software requirements are: title, description, explanation, link to source, creator, last modifier, priority, status, links to other knowledge items, links to structuring models.

A further type of information that could be tagged to a piece of knowledge is an *issue*. The idea of issues is used by gIBIS [CON89] and Argnoter [SFB+97] to highlight points in a design that need to be addressed.

[VNJ00] makes use of meta-knowledge in a slightly different way. The knowledge used in the system can be filtered according to criteria such as concepts connected to a piece of text, frequency of occurrences of a piece of text, or complexity of the concept. This type of meta-knowledge is obtained implicitly from the model by analysing its structure rather than extracting explicitly stored information from the model.

4.3.6 Classification & Ontology

Many corporate memory systems are limited by the fact that documents can only be stored in a single place in a hierarchy of topics. Documents cannot be assigned to many topics without needing to create links or creating multiple copies. Modern document management tools have the ability to classify documents according to many different classification schemes. Such tools are aware of document dependencies, versions and representations, although "...integration of supporting techniques, e.g., the automatic classification of knowledge items to various classification schemes, has hardly been considered" [TAHZ00].

University of Cape Town

Commonly, a single classification scheme is used to structure information. For real domains such a scheme is very complex to maintain. [TAHZ00] presents a means of having multiple classifications, one for each view of the information. The system described allows users to manually classify documents, or to let the system automatically classify documents according to automatic text classification algorithms.

The knowledge categories that organizations define for their corporate memory system is called an *Ontology*. The Oxford English dictionary defines Ontology as the study of the nature of existence. Therefore, the fundamental question of ontology is simply: "What is there?", that is, what things exists in the knowledge one is attempting to model?

Classification can be seen in various forms; for example in some web search engines, one has the ability to search from a very general category to more specific domains of interest. However, the question remains about who developed such hierarchies in the first place. This type of classification might be well accepted in highly formalized environments for very specific domains, but could not be very well accepted in "...more loosely organized community settings such as dynamically emerging community of individuals with different backgrounds" [VNJ00]. In these environments, categories should reflect the interests of the users.

The top levels of the ontology of the *Cyc* project [LG90] are illustrated in Figure 10. The Cyc project is a massive knowledge management project whose goal it is to map all of human knowledge. Figure 10 represents an early stage of the Cyc upper level ontology; since then the ontology has been extended to more than 100,000 concept types.

Ontology creation is widely accepted as one of the most important stages in corporate memory development. Ontologies help in organization and retrieval of knowledge. [REI98] proposes that the problem of integration of multiple knowledge bases can be solved by using some type of ontology or mapping of similar terms between knowledge bases.

University of Cape Town

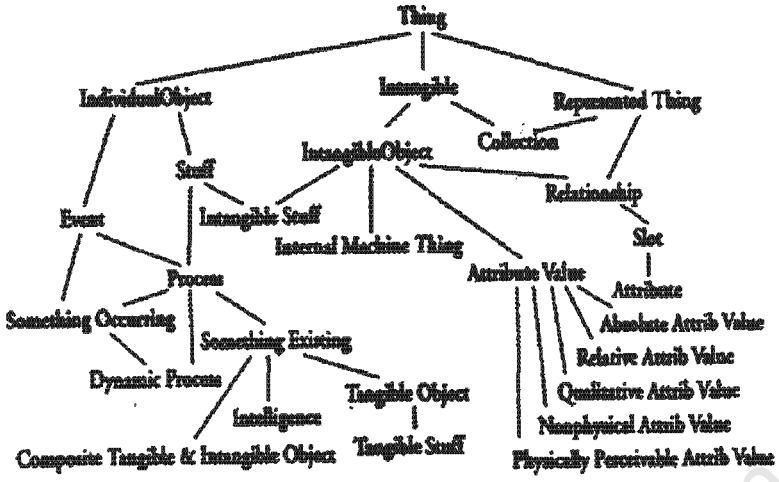


Figure 10: Top level categories in an early version of the Cyc ontology

Minimum ontological commitment is a term used to describe an ontology that is minimal to be able to explain a specific domain [ZM95]. A minimum ontology can suffer from a problem whereby certain information does not fit into the ontology and this possibly vital information could be discarded.

The QBIQ (Quality Integration Application for Operational Documents) system allows four different levels of classification for documents: Top Level - applicable for all departments in an organization, Cross functional level - high level documents applicable to many departments or processes, Departmental level - documents for specific departments, Item specific level - documents relating to specific item numbers or analysis numbers.

[MBD99] describes an ontology-driven environment that enriches documents by providing information about the context in which a document is created, and linking it to related information. A top-down approach is used where an ontology of necessary information is constructed first by defining concepts, relations, and axioms and thereby defining the type of information one wishes to extract from the documents. This is followed by the enrichment of documents with this ontology. This can be contrasted with a bottom up approach where documents would be enriched according to information

University of Cape Town

specifically contained in them (no prior ontology development). A screen-shot of the ScholOnto ontology presented in [MBD99] is provided in Figure 11 below.

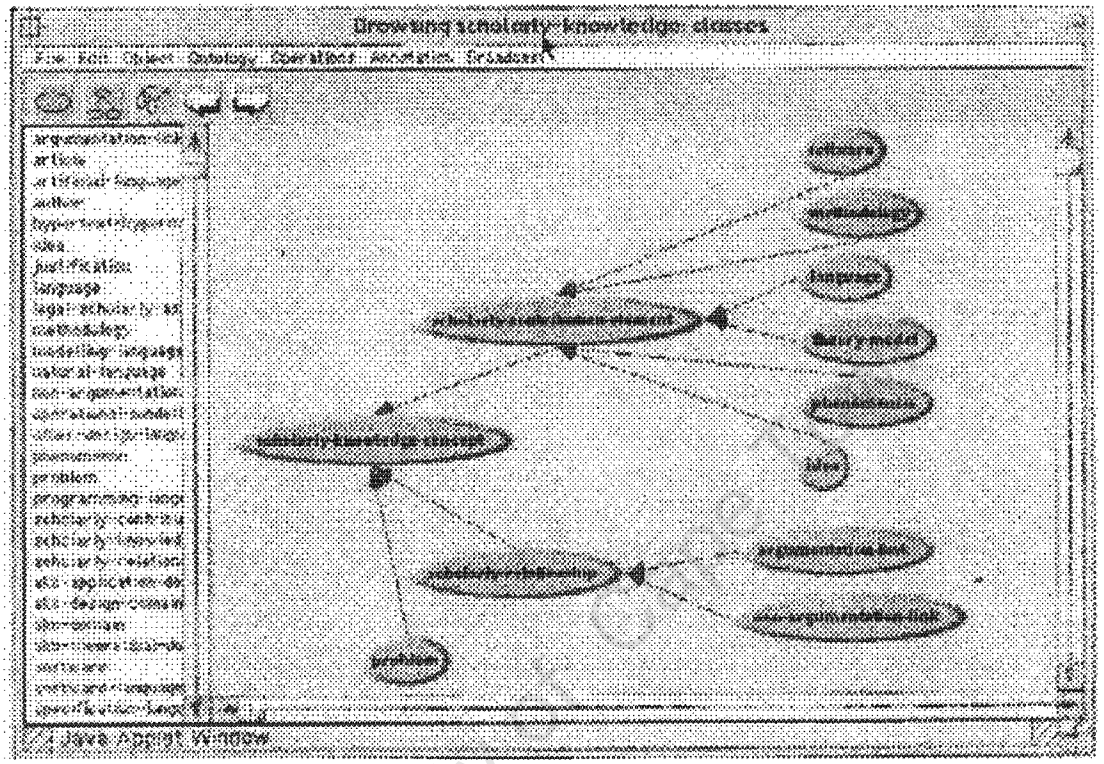


Figure 11: Screen-shot of ScholOnto ontology browsing tool

There are a number of commercial products and projects that are currently being developed that deal with creating ontologies and automatic classification of documents. The reader is referred to OntoBroker [FES98], WebKB [ME98], IBM's text analysis tools [IBM00], Oracle's ConText Cartridge [ORA00], Semino's text mining tool [SEM00] and WebMind's assisted learning technology [WEB00].

4.3.7 Summary

Although a relatively complete list of high-level corporate memory requirements has been outlined in Section 4.3, Subsections 4.3.1 through 4.3.6 have elaborated on the most important requirements. These are seen as the most important requirements because they are fundamental to the success of any corporate memory. At a minimum, a corporate memory should allow for free-form capturing of knowledge, storage of user profiles,

storage of meta-knowledge, multiple viewpoints, classification, and integration into the workplace. Adherence to these requirements will ensure that the system allows for storage of all types of information, that information can be directed to relevant individuals, information can be easily retrieved, and the system is effective in its usability.

Inclusion of other requirements such as communication facilities, knowledge annotation and interaction through different sources, are all features that would make a corporate memory more powerful, but in many cases could be seen as application specific and therefore not forming part of the fundamental requirements.

4.4 Learning with corporate memory systems

As pointed out at the beginning of this chapter, the main purpose of a corporate memory is to enable re-use of knowledge. Learning can happen in one of two ways, either using a push model – by depending on the system to distribute knowledge to relevant users based on user profiles [vHvdSK96] [DCGR98], or a searching model that requires the user to actively find knowledge.

The learning mechanism of the corporate memory system should allow the user to learn by example, by reflection, using case based learning, using goal directed learning, exploratory learning or learning by communication [vHvdSK96].

Usage of the knowledge in the corporate memory would happen through some type of groupware or querying devices. [vHvdSK96] points out a number of different ways of querying a corporate memory:

- Hierarchical searching – knowledge is searched by traversing the knowledge structure
- Attribute searching – the knowledge is searched by specifying values for attributes
- Content searching – searching happens according to arbitrary search terms specified by the user of the system.

University of Cape Town

Each method has advantages and disadvantages. For example, attribute searching is more flexible than hierarchical searching but requires a set of predefined attributes. Content searching is even more general than attribute searching but requires users to formulate appropriate searching conditions and is not as computationally efficient as hierarchical searching, although, navigating through a structure cannot really be called an "automated searching facility" like content and attribute searching.

[MW99] and [MBD99] advocate the use of agents as an effective means of achieving interaction with knowledge based systems and its environment and thereby minimising the need for a user to actively search for needed information. [MW99] offers an example of British Telecom's Knowledge Sharing Environment that is based around agents that work in four areas; email notification, search engine, what's new and interest groups. Using these four areas, users are notified by agents of what that agent perceives to be relevant knowledge, which he/she could potentially find useful.

Despite all the sophisticated tools that can be implemented by corporate memory system designers, it remains a fact that in many cases, one of the quickest way of finding a solution to a problem would be to ask someone who knows the answer, especially if that person is in the same location [DP98]. The next best option would be to ask another colleague in another part of the organization by email or bulletin board. If no help was found immediately, then one would attempt to find the answer in an explicit form such as a manual or reference book. It is therefore essential for a corporate memory system to enable users to interact with each other for the purposes of finding relevant information.

4.5 Maintaining the Corporate Memory

Maintenance and evolution of the corporate memory should be done in order to better achieve the goals of the system [DCGR98]. [FB99] identifies knowledge maintenance as an important characteristic of knowledge systems to capture new know-how that experts accumulate as a result of the knowledge system but also highlights that "Maintaining the corporate memory is the hardest concern in a corporate memory project" [FB99].

University of Cape Town

[FB99] notes two lessons learnt from knowledge maintenance. The first being that the need for an *experience feedback* mechanism is becoming important – it is realized that any know-how needs to be constantly maintained by having users capturing their experiences of using certain items of knowledge to enable more effective re-use by others. The second lesson is that the number of participants involved in corporate memories is on the increase and therefore maintenance by traditional methods such as interviews is becoming impractical – the experience feedback loop needs to be reduced or made more efficient.

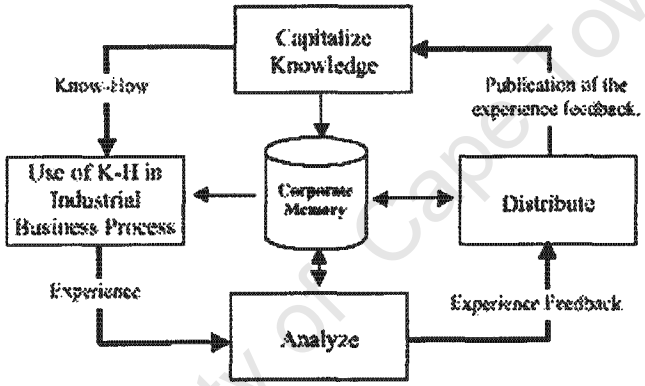


Figure 12: Experience Feedback loop [FB99]

Figure 12 presents an example of a formalized analysis and feedback cycle discussed in [FB99]. *Knowledge capitalization* refers to storage of information and placing that information in its professional context. In the *analysis stage*, an attempt is made to extract useful information from experiences gained through usage of an item of knowledge, and thereby generating feedback about the experience to others.

A very important user feedback mechanism that is becoming predominant in corporate memory systems is *ranking* of knowledge. "Ranking is important since there will always be multiple hits on any knowledge retrieval request and they must be ranked intelligently so that the user is presented with the best match first" [SDB00]. For example, a company called Hotdispatch.com [HOT00] has developed helpdesk technology that enables a

University of Cape Town

person to ask a community a question. Anyone in the community has the right to respond to the problem. When the enquirer receives an answer, the answer is rated according to clarity and usefulness, and one of the responders is chooses to receive payment for the answer. In this way, competencies and reliability ratings are assigned to people thereby allowing the system to recommend knowledgeable people in specific domains of expertise.

Connection between group members and contributions can be made not only to enable people to form mental pictures of mutual interests, backgrounds and understandings, but also to form mental pictures of competencies, thus enabling users to estimate the value of a certain piece of knowledge. Reputation is often used as a proxy for quality; knowledge quality can therefore be quickly assessed by the trustworthiness of the knowledge originator [BLD00].

Versioning is another mechanism that can be used during the maintenance cycle. Versioning allows systems to store multiple versions of the same information for purposes of comparison and the ability to monitor change of ideas over a period of time. "Objects" are defined as a collection of related information, "Versions" are a description of the object implementation, and "Configurations" refer to a collection of related versioned objects [BK86]. In the PIE [BG80] system, the designers have implemented the idea of layers where each layer would contain new versions of objects. One would have the ability to navigate between these layers to obtain different versions of objects and therefore trace their evolution over time.

[KAC86] discusses a database to store design information about objects. The database consists of a number of interrelated design objects. The representation mechanism uses a three-dimensional space made up of a *version plane*, *configuration plane* and an *equivalence plane*. The version plane organizes objects over their lifetime in a version history. The version plane could be linear or hierarchical if more than one version can be valid at any time. The configuration plane combines versions of different objects to form composites. The equivalence plane correlates equivalent objects across configurations.

University of Cape Town

4.6 An Example Corporate Memory System

Now that the requirements and techniques of corporate memory systems have been discussed, an example corporate memory system will be presented to show how these are drawn together in a system. The system in [BMD99] presents a knowledge-based web environment that supports interpretation of ideas in relation to published literature.

Text is a very rich medium in which to present ideas; the only problem is that it takes too long to read and is too difficult to analyse computationally. The goal of this system is to provide users of the system summary representations of ideas being stored in the system and how these ideas relate to other literature.

The system enables tracing of ideas, what impact they have made, how they have been supported and challenged. This method of commentary on peers' ideas can answer a number of important scholarly questions such as: has anyone built on these ideas? Has anyone challenged this paper? Has a different theoretical perspective been proposed? Are there any document building on theory but contradict predictions? Are the papers applying methods to certain domains? Does a tool exist that tackled a certain problem? What impact did a certain theory have? Are there any distinctive theories on a similar problem?

This system uses documents to represent the content of knowledge. However, despite this free-form representation of knowledge *content*, the system makes use of a semantic network to represent *relationships* between knowledge. This semantic network makes use of predefined labels for its nodes and arcs. Users of the system are therefore limited to using the specified knowledge and relationship types when capturing information and relating pieces of information to each other. Figure 13 below shows the elements used for summarizing the key contributions of publications, and relationships to other literature.

University of Cape Town

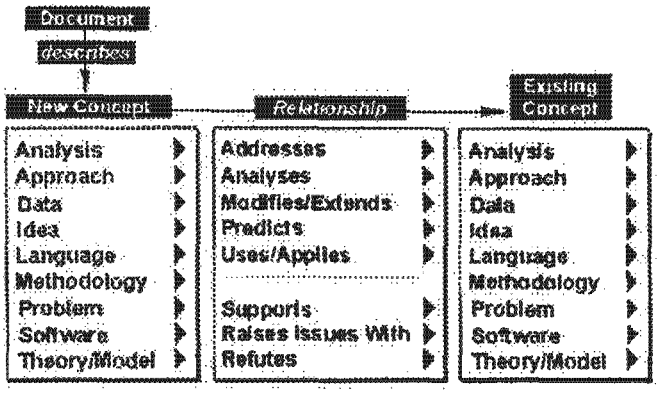


Figure 13: Representational elements for summarizing the key contributions of a publication, and their relationships to other concepts [BMD99]

Using the ontology displayed in Figure 13 above, a network of literature and relationships is constructed. There are a number of interesting possibilities that can be achieved with this type of knowledge network. For example: if "impact" has been defined as number of documents using a theory, then we can answer questions like "What impact did the theory have?". One also has the ability to create complex rules about interrelated documents and links to see if there are two schools of thought about a certain theory (for example). Intelligent agents can be used to answer more complex questions that cannot be solved with simple search criteria; for instance getting a geographical perspective of ideas.

Figure 14 displays an example of how the structure of nodes and relations can be used to visualise the relationships between documents. In this example, an author has described the key contribution of an article with respect to other work done (links to the left). The links to the right represent how other researchers have modified or extended the work.

University of Cape Town

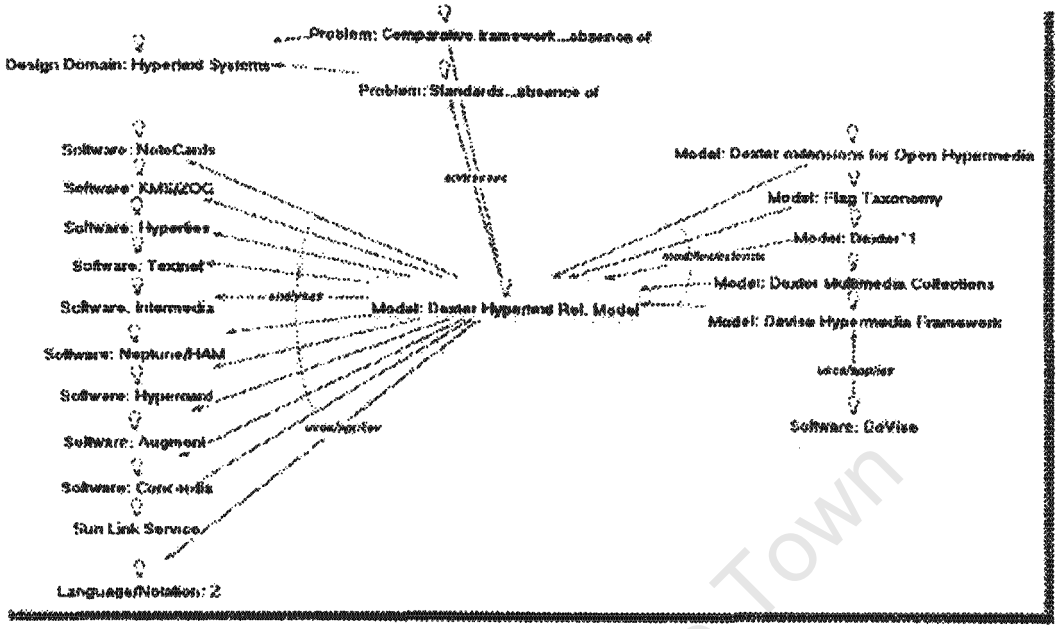


Figure 14: Example of a how a document has been related to other documents in the system [BMD99]

This system adheres to some of the most important requirements of a corporate memory system.

- Work-place integration through the use of a web-based interface
- Freedom of knowledge representation through the use of documents, and therefore no limitations on formal knowledge representations.
- Ability to represent multiple viewpoints by maintaining information about the authors of different pieces of knowledge
- Classification of knowledge through the use of an Ontology
- Knowledge retrieval through searching mechanisms as well as navigation.

The system does not adhere to the following requirements (outlined in Section 4.3):

- Maintenance of user profiles
- Communication mechanisms
- Facilities for user feedback
- Knowledge annotations
- Storage of cultural and contextual information

University of Cape Town

- Storage of meta-knowledge
- Facilities for knowledge retrieval through a variety of media

This system is a very good example of how a formal knowledge framework has been used to relate free-form knowledge to enable visualisation of how an item of knowledge relates to other information contained within the system.

4.7 Comparing Epistemic and Corporate Memory Representations of Knowledge

Research into corporate memory systems has primarily focused on how to represent knowledge most efficiently for effective retrieval by people. Systems such as document repositories deal with questions like "What is the most relevant knowledge for the user?" rather than with the question "What is the user justified in believing?".

Modern corporate memory systems have started to understand the need to represent people in an organization as individuals with their own belief systems and unique ways of structuring and retrieving knowledge. This progression has been characterised by the increase in corporate memory systems that have the ability to represent individual ideas as well as consensus ideas. Such conflict management systems can also be seen as realisations of *epistemic consensus consequentialism*, where, if the knowledge has produced consensus among a team, it is proved to be very useful for many (if not all) of the team members.

Systems that depend on "pre-analysis" of knowledge result in knowledge that is very probably true, or is a viewpoint held by most experts in a field and has therefore proved itself to be very useful. Editorial boards are examples of *Annis's contextualism* (See Section 2.1.4, *Contextualism*), where the objector group would be the editorial board that needs to agree that the new knowledge being presented is contextually basic with respect to that objector group, or there is sufficient evidence to show that this new knowledge is justified.

University of Cape Town

In the same way that social epistemology uses the reliability of a person as a measure of the truth of a certain item of knowledge, so too have certain progressive knowledge management companies. As mentioned in Section 4.5 (*Maintaining the Corporate Memory*), Hotdispatch.com [HOT00] is a good example of a company that makes use of user feedback to rate the competence of individuals and thereby enabling people to find the skills needed to solve specific problems.

In epistemology it is not only people that need to be "ranked" according to the reliability with which they produce knowledge. Reliability of belief producing processes plays a vitally important role in assessing the truth or usefulness of knowledge. Some organizations may choose to capture their knowledge using procedural techniques, while others make use of declarative representations for representing results of the procedures. In general though, corporate memory systems do not rate the trustworthiness of processes.

Coherence of new information to one's existing belief system plays an important role in terms of how a person accepts this information. Some of the types of meta-knowledge that have been presented, such as *cause*, *intention* and *side effects*, strengthen the coherence an item of knowledge has with its surrounding beliefs.

To enable effective retrieval of knowledge, there has been a need to add some type of formalisms to corporate memory systems. One of the most popular formalisms that can be found in such systems are classification hierarchies, or as some prefer to call them, Ontologies. Ontologies store the topics about which knowledge exists in an organization. Ontologies found in corporate memories contain entries that are useful for effective knowledge retrieval, whereas epistemology deals with very high level topics that for the most part would not be very useful for corporate memories. It is however very important to note that both fields have expressed a need to represent the existence of things around which knowledge is created.

University of Cape Town

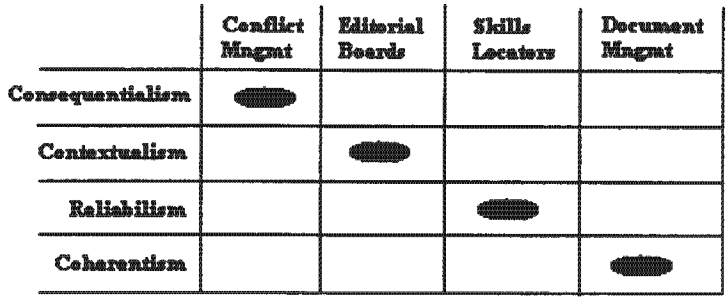


Figure 15: Corporate memory structures from an epistemological perspective

Figure 15 classifies some of the corporate memory structures discussed in this chapter under the relevant epistemic theories:

- *Conflict Management* systems manage different views and are therefore implementations of consensus consequentailaism
- *Editorial Boards* ensure that information is relevant to the context of the publication and therefore represent contextualism.
- *Skill locator* systems (such as Hotdispatch.com [HOT00]) ensure that one can obtain reliable skills in a specific domain of expertise
- *Document Management* systems contain documents that (primarily) contain holistic information and specific topics, and therefore represent coherent pieces of information.

Despite the fact that many corporate memory system designers are oblivious to epistemology, it seems that the trend in modern corporate memory models is to represent the reliability of information based on many factors such as reliability of belief producing agents, the coherence of the belief within one's belief system, and consensus among experts as to the usefulness and relevance of new information. Therefore, by designing systems that give users greater freedom of expression and ability to judge knowledge obtained from others in terms of how truthfully it represents the world, corporate memory systems are gradually starting to represent the characteristics of social and individualistic epistemology.

University of Cape Town

5 An Epistemically Grounded Corporate Memory Model

Although many of the systems mentioned in Chapter 4 (*Corporate Memory Models*) have made use of some ideas from epistemology, these overlaps are primarily coincidental; no model is explicitly based on epistemology. The corporate memory model presented in this chapter, called *EpiK*, has been explicitly based on epistemology and proves itself to be as useful as existing models for facilitating information retrieval and more useful than existing models for understanding the dynamics of information and people in an organization.

EpiK is based on two major principles. Firstly, different people may believe things in different ways. Modeling knowledge as a shared memory does not offer individuals the opportunity to represent thoughts and ideas as they see fit. Secondly, different people have different rankings of the truth and usefulness of beliefs.

Section 5.1 will provide a brief overview of EpiK, section 5.2 describes its core components, and section 5.3 will describe higher-level derived constructs. Section 5.4 shows how EpiK can be used for the purposes of learning. Section 5.5 examines the ways in which it can help to understand the social dynamics of the organization as well as the dynamics of the knowledge contained inside the corporate memory. Section 5.6 will compare EpiK with other work done in the field of corporate memory systems. Section 5.7 concludes by describing the way in which EpiK is representative of epistemic concepts.

University of Cape Town

5.1 Overview of the Corporate Memory Model

The knowledge model described in this chapter was born out of an attempt to represent epistemic concepts found in philosophy in a form suitable for use as basis of a corporate memory system. The model, called EpiK, makes use of concepts such as *truth values*, *usefulness values*, *communication channels*, and *individual belief systems*. It does not rely on shared knowledge repositories and allows a person as much freedom as possible in

relating beliefs to each other. It assumes that all information is contained in peoples' minds and never as shared concepts. Every person has his/her own set of beliefs, methods of rating the truth and usefulness of beliefs, view of the external world, set of goals, and set of processes and people that are seen as sources of knowledge. To model knowledge sharing, there are links between beliefs in different minds that represent communications between people.

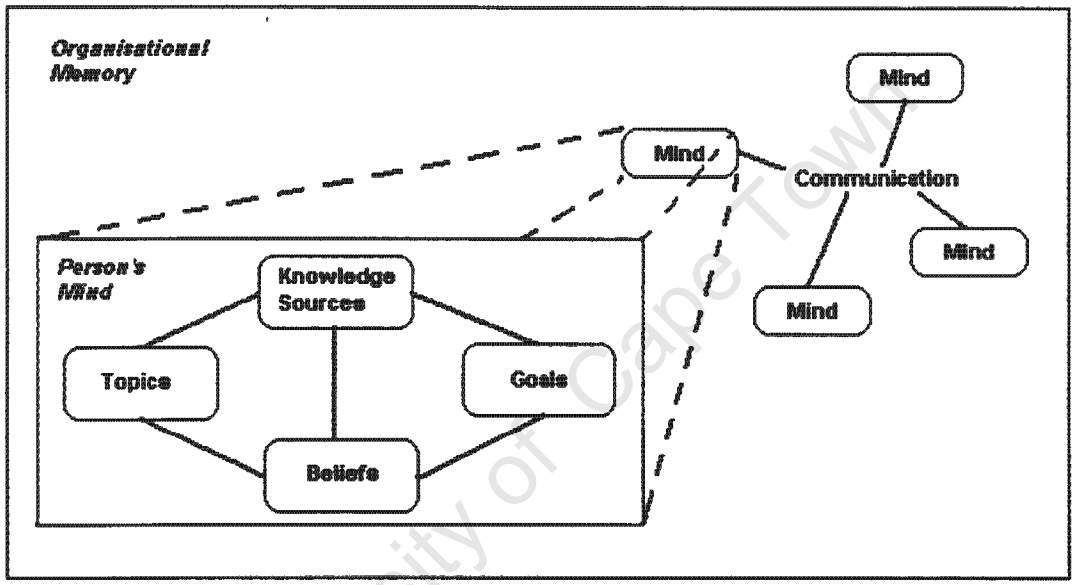


Figure 16: Overview of corporate memory model

University of Cape Town

Despite the fact that EpiK is aimed at being representative of real human knowledge structures and interaction, at an organizational level we are not limited in making use of it in a strictly epistemic fashion. We can, for example, extract information such as which beliefs are most widely known (more examples can be found in Section 5.5 (*Evaluating People and Knowledge from EpiK*). Answering such questions requires one to use the knowledge model in a manner that in some senses can be seen as going *beyond reality* by allowing the organization the ability to have a global view of the knowledge contained in every person's mind. This is acceptable when keeping in mind the final objectives of a corporate memory system, which is to enable people to learn quickly and effectively and to help management capitalize on the intellectual capital in their organization.

The model's core is comprised of two major components: *minds* and *communication channels*. A mind is a representation of a single individual's beliefs and communication channels represent exchange of knowledge between individuals. Every mind is comprised of beliefs, belief relations, topics, goals, knowledge sources, meta-knowledge and belief ratings. *Beliefs* contain the knowledge content, *belief relations* relate beliefs to create a belief system. Every belief originates from a *knowledge source*, contains knowledge about a *topic* and could also be associated with a *goal*. Every belief has meta-knowledge associated with it, as well as truth-values, and usefulness ratings that represent how useful an item of knowledge is in attaining certain goals.

Over and above its major components, EpiK has further components that exist to make information retrieval more effective. Such components include equivalence mappings between topics and goals in different minds. These components allow one to understand the different ways in which people classify and make use of similar items of knowledge and thereby derive a kind of organization wide ontology of goals and topics.

University of Cape Town

5.2 Core Components

This section will provide an overview of the core components that make up EpiK. The core components are beliefs, belief relations, topics, goals, knowledge sources, communication channels and meta-knowledge.

5.2.1 Beliefs and Belief Relations

A belief represents a piece of knowledge that could be in the form of production rules, semantic nets, frame systems, logical expressions, free-form text, or any other representation scheme that is necessary to represent that piece of information effectively.

Relations represent links between beliefs. Relations do not actually contain any information, so the relation *label* describes the way two beliefs are related to each other. For example, if an observation support a hypothesis then the relation between the hypothesis and observation could be the relation "supports" (for example).

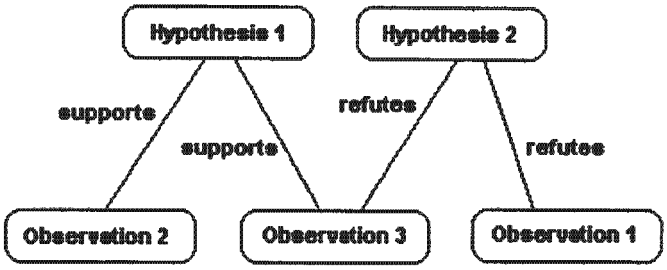


Figure 17: Example of Beliefs and Relations

Together, beliefs and relations make up a network of interrelated beliefs that represent a single individual's belief system. This lattice of interrelated beliefs is called a *mind*. A mind could represent not only a human's belief system, but also anything else that produces beliefs and communicates those beliefs to others; for example, a mind could represent the knowledge of an automated agent.

The user is not limited in the types of relationships that can be created. That is, EpiK does not specify a limited set of relationships that can relate beliefs. In this way, every person uses his or her own unique way of structuring information. Furthermore, one is not limited to representing the beliefs of people or agents *inside* an organization; minds can also represented the beliefs of people or agents external to an organization. These minds are called *virtual minds*.

University of Cape Town

5.2.2 Topics, Goals and Ratings

In the same way as everything a person believes is about something that exists in the world around us, so too is every belief related to one or more *topics* and *goals*. A topic could represent either a concrete or abstract thing about which knowledge could be produced. A goal represents the objective a person is trying to achieve.

A belief could contain knowledge about more than one topic and could be useful in attaining more than one goal. For every topic to which a belief is related, the belief is assigned a *truth-value* that represents how truthfully the person feels the belief describes

the related topic. Similarly, for every goal to which a belief is related, the belief is assigned a *usefulness-value* that represents how useful the belief is in attaining certain objectives.

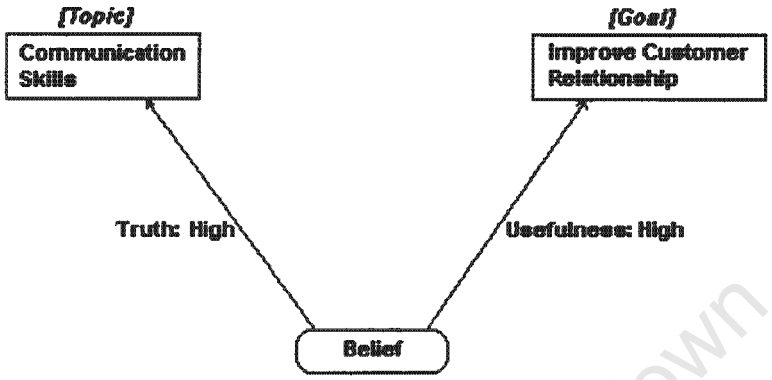


Figure 18: Example of a belief about a topic that helps to achieve a goal

Figure 18 presents an example of an item of knowledge that is very truthful in the way is represents information about communication skills and was found very useful in improving customer relationships. This relationship can be interpreted as "It is very true and very useful to know that communication skills improve customer relations".

5.2.3 Knowledge Sources and Derived Ratings

Every piece of information in a person's mind owes its existence to a knowledge source. Examples of knowledge sources are: other people, agents, observations, or a reasoning process.

Knowledge is generated by knowledge sources and related to topics and goals (See Figure 19). At the same time it is given truth and usefulness values. Using all this information, the trustworthiness and usefulness of the knowledge source can be derived with respect to the knowledge they produce about certain topics or goals.

University of Cape Town

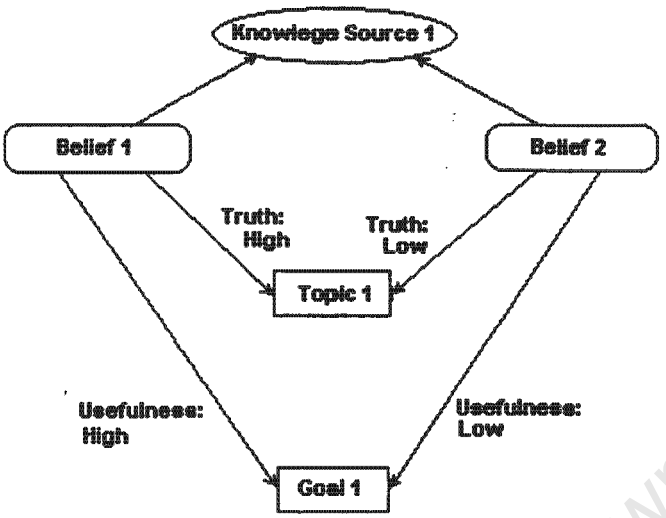


Figure 19: Two pieces of knowledge produced by one knowledge source

From the structure depicted in Figure 19, by taking the average of the truth and usefulness values, one can derive that an average truth rating for *Knowledge Source 1* when producing knowledge about *Topic 1*, and an average usefulness rating when producing knowledge that helps to achieve *Goal 1*.

As a knowledge source produces more truthful or useful information, the rating of the knowledge source will reflect these changes. In this way, a person is not forced to take an epistemic stance as to which knowledge producing processes or persons are reliable. The derived truth and usefulness values can then be used to quickly and easily evaluate the possible truth and/or usefulness of a new item of information that has been acquired.

A mechanism called *default rating rules* is provided as a means of making use of existing ratings of knowledge sources to automatically rate new knowledge. Default rating rules are based on the knowledge source, the related topics and the related goals of a belief. If a belief matches the criteria specified in the default rating rule, then the rule can be used to automatically rate the truth and usefulness of the belief. These rules are an attempt to alleviate the responsibility of the user to rate every item of knowledge obtained.

University of Cape Town

5.2.4 Communication Channels

Communication channels represent the exchange of beliefs between minds and are described by a source belief and a destination belief in another person's mind. Communication lines arise through the communication of information from one person to another and therefore act as relations between beliefs in different minds.

By relating beliefs using communication channels, one is able to recognize that two beliefs held in two or more different individuals' minds are actually the same. In this way communication paths of knowledge can be traced, thereby deriving the extent to which a piece of knowledge is known within the organization.

An example of modeling communication can be seen in our everyday interaction with electronic mail. People are often the recipients of messages that contain the words "fw:" or "(fwd)" in the message's subject. This tells the recipient of the message that this message has been forwarded from another individual. Furthermore, the path of the message is often attached to the top of the email message along with a list of recipients.

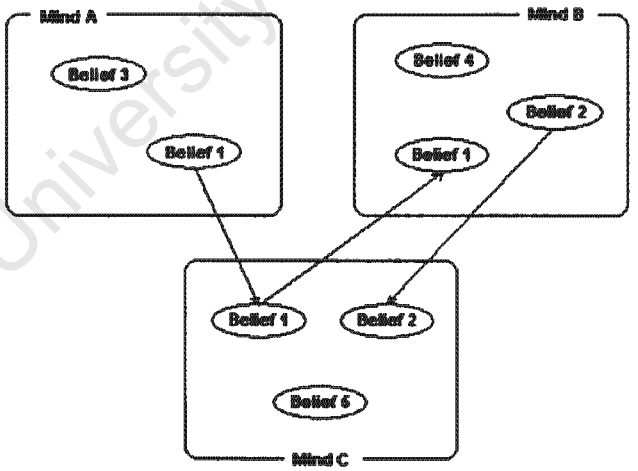


Figure 20: Communication of beliefs between minds

In Figure 20, one can notice that *Belief 1* is widely known, while *Belief 2* is known by two people, and *Beliefs 3*, *4* and *5* are only known by single individuals. In certain cases

University of Cape Town

it may be necessary to base the truth and usefulness ratings of a belief on what other people think of the belief. This is done by following the path of the communication of this information and analysing what other individuals think about the trustworthiness of the author.

By making use of communication channels, EpiK is able to identify equivalent beliefs across many minds, and thereby propose topics and goals that are similar across minds.

5.2.5 Meta-Knowledge

In EpiK, the concept of *meta-knowledge* does not refer to information that is used as control knowledge for problem solving, nor as information that is used in understanding internal knowledge representation structure. Meta-knowledge is defined as any information that is *automatically accessible* to a system implementing such a model. That is, the information a system can obtain without the need for human interaction. Such information is *factual* and not open to argument. The reasons for this definition of meta-knowledge is to avoid meta-knowledge maintenance issues, and to offer information retrieval heuristics a framework of constraints that can be assumed to be true.

The type of information that could be stored as meta-knowledge cannot be stipulated. The reason for this is because it is limited only by the type of interaction an implementation of EpiK would have with the environment. For example, if a reliable source of time was available to a system, then a possible item of meta-knowledge that could be stored with every belief is the time of belief creation.

Critics of meta-knowledge question the very existence of non-arguable facts. In the example above it could be argued that even reliable time sources may be incorrect by fractions of seconds and therefore the time of entry of knowledge can also be refuted. This is a valid argument, however philosophy does introduce a concept of *epistemic immediacy* in the form of *a priori* knowledge, which simply means that there are certain pieces of information that one assumes to be true without the need to contemplate truth-values.

University of Cape Town

5.3 Derived Constructs

By making use of the core components of EpiK, it is possible to automatically extract information about the social dynamics within an organization, and discover widely held, highly truthful and very useful beliefs.

5.3.1 Groups

A natural side-effect of communication is the formation of *groups*. Groups of people could represent departments, research teams or even social groups of people that interact with each other. By making use of communication channels the system is able to automatically identify individuals that communicate often with each other without the need to explicitly define such groups.

This automatic derivation of groups eliminates the need for creating "artificial" groups of individuals that somebody thinks may, or should, be communicating. Therefore, groups can be created, destroyed or changed during the lifetime of the knowledge model without having to explicitly make such changes.

5.3.2 The Consensus Mind

It is one thing to give users a great amount of flexibility to represent and relate knowledge, but it is another to offer the same users the ability to retrieve information effectively from such a free-form knowledge representation.

When attempting to learn about a new subject, or trying to solve a problem, people are in need for accurate, useful information. There is therefore a need to be able to offer users such information from the knowledge model. By making use of truth values, usefulness values and communication, one is able to derive information that is widely regarded to be valuable.

By retrieving beliefs that have high average truth-values and are known by a large percentage of a community, one has access to knowledge that is potentially very important to the organization. In the same light, retrieving beliefs that have high average

University of Cape Town

usefulness-values and are known by a large percentage of a community, one is tapping into a source of information that has been found to be very useful by other people.

Such *widely held*, *highly certain* and *highly useful* knowledge is maintained automatically by the system inside a *Consensus Mind*. The consensus mind has the same structure as the mind of a person, but it represents knowledge that is highly regarded throughout an organization. The consensus mind should be the starting point of any search for information within an organization.

Therefore, by monitoring the truth and usefulness values of beliefs, EpiK is able to dynamically and automatically maintain the consensus mind. As information is believed by a wider group of individuals and to a greater extent, it would be added to the consensus mind. Similarly, as information is believed to a lesser extent, it would be removed from the consensus mind.

In the same way as beliefs are maintained in the consensus mind, so too would the topics, goals and knowledge sources. Maintaining such information in the consensus mind would be essential to the usefulness of the consensus mind because it is vital to know not only important beliefs themselves, but also the topics which the beliefs are about and the goals for which they are useful. Maintaining a list of knowledge sources that have produced widely held, accurate and useful knowledge is also important as an indication of sources people hold to be reliable.

University of Cape Town

5.4 Learning from EpiK

Learning from the content of the corporate memory model can happen in two ways, *intrusively* and *non-intrusively*. Non-intrusive learning occurs in our every day lives – one communicates an interest about a certain topic to another individual and that individual communicates relevant information back to us. This is called *non-intrusive* learning because one does not have access to other people's minds, one is completely reliant on communication from other individuals for obtaining knowledge from them.

However, having developed a knowledge model, it is important to discuss how it can help in the process of learning that goes "beyond reality" by giving individuals access to each other's belief systems. Giving access to others' minds provides users with the ability to learn from the content without being dependent on communicating an interest and expecting a response. This type of learning is called *intrusive* because one can "pry" into other people's minds to obtain knowledge.

The process of learning in EpiK would primarily comprise of specifying parameters around which to filter information. The only practical difference between intrusive and non-intrusive learning would be the data-set to which the criteria are applied; for intrusive learning the data set comprises all minds, while for non-intrusive learning the data set is comprised of only one's own personal mind and the consensus mind. Once the filtering process has resulted in a small enough result set, one is able to learn through navigation of related beliefs. Therefore, finding information in EpiK makes use of the three different searching techniques outlined in Section 4.4 (*Learning with corporate memory systems*), namely hierarchical search, attribute search and content search.

One of the major benefits of learning from this model comes in the ability to combine content-based queries and truth-based queries. For example, one could ask questions such as "Find all beliefs containing keyword *xyz* that are believed to a high degree by at least *n* individuals".

An important thing to remember is that the consensus mind offers users the ability to quickly and easily find trustworthy information about specific topics of interest, or information that is very helpful in attaining specific goals. Because the consensus mind is constantly changing according to the trust and usefulness that the organization as a whole assigns to beliefs, it would always remain a very good starting point for any learning exercise. Therefore, to learn about a new topic, one might want to start by learning from the consensus mind followed by a process of navigation to obtain more detailed information that might not be widely known.

University of Cape Town

5.5 Evaluating People and Knowledge from EpiK

It is important for a corporate memory system to provide knowledge managers useful information that describes the dynamics of knowledge and people in the organization. Below are examples of the types of management level reports one could obtain from EpiK about the knowledge contained in it, about the people in the organization and about persons outside of the organization.

Reports about Knowledge

- What are the most trusted beliefs?
- What are the most useful beliefs?
- What are the most widely known beliefs, topics, goals and knowledge sources?
- Through what routes does knowledge get disseminated in the organization?

Reports about People

- Which people or processes are the primary sources knowledge?
- Which people are the primary consumers of knowledge?
- Are there certain individuals that promote distribution of knowledge?
- Are there any communities that are developing inside the organization that are independent of imposed organizational structures?
- Which people outside of our organization have skills in fields beneficial to the organization?

Reports combining information about people and knowledge

- Which people or processes are the primary sources of widely known, trustworthy and useful and knowledge?
- How dependant is the organization on a small number of individuals for knowledge about specific topics?
- Is there knowledge in the organization that is known by a significant number of people that could assist in creation of new project ideas?

University of Cape Town

The reports described above have many benefits for the organization. Some can be used for evaluation of employees by how well they produce and distribute knowledge as well as how respected they are by their peers. Others can be used to optimize the routes that knowledge follows through a community as well as promoting the distribution of highly believed knowledge.

5.6 Comparison to other Corporate Memory Models

The corporate memory model presented in this chapter has a number of similarities to corporate memory models discussed in Chapter 4 (*Corporate Memory Models*).

Representation of different types of knowledge – EpiK does not limit the type of knowledge that can be stored inside it. In some ways this offers users great flexibility, but in other ways it is not geared at alleviating problems in knowledge representations of specific domains. Unlike the knowledge representations presented in Section 4.3.1 (*Representing Different Types of* Knowledge), the goal of EpiK is to represent knowledge as *generally* as possible and avoid issues involved in representing knowledge for specific domains.

Classification – corporate memory models traditionally make use of a single classification tree for use by many individuals. EpiK is different in that it allows each individual to maintain his/her own set of topics. By making use of communication channels, it is able to identify equivalent beliefs across many minds, and thereby propose topics that are similar across minds. In this way, global classification information is derived and stored in the consensus mind.

Meta-Knowledge – EpiK does not attempt to store all possible items of meta-knowledge about every belief. The criteria used to differentiate meta-knowledge from beliefs is that meta-knowledge is information that is completely certain and based on the epistemic concept of a priori knowledge (knowledge gained prior to experience). This eliminates meta-knowledge maintenance issues that have been discussed in Chapter 4 (*Corporate Memory Models).* Because of this definition, one is limited

University of Cape Town

to storing meta-knowledge that is derived from a source that is completely trustworthy and not susceptible to human bias.

Despite this definition, there are a number of items of information that are stored implicitly in EpiK that have been classified as meta-knowledge by other corporate memory systems discussed in Chapter 4 (*Corporate Memory Models*). Such information includes knowledge sender and knowledge receiver. If one takes the stance that information obtainable from a user can also be regarded as meta-knowledge (for the purposes of this comparison to other models), then EpiK also stores: domain of the knowledge, uses for the knowledge, knowledge source and relations to other items of knowledge previously gathered.

Multiple viewpoints – a number of systems that allow multiple viewpoints have been presented in Chapter 4 (*Corporate Memory Models*). EpiK is very similar to such systems in the fact that it allows users to maintain their own ideas. It seems however that conflict management software is geared towards understanding of the actual conflicts, whereas EpiK is used more to extract widely held, very accurate and very useful knowledge.

Furthermore, Epik treats the whole organization as a single group and allows for only one consensus mind that is the result of obtaining widely known beliefs across the whole organization. The conflict management systems presented earlier have more powerful functionality in this respect and allow for more granular management of consensus information by managing conflict among *subgroups* of an organization, and thereby forming *hierarchicies* of consensus information.

The model can be extended by allowing the system to generate multiple consensus minds, each for a selected group of individuals. These consensus minds can then be used for creating even higher level consensus minds between groups of individuals. This would give EpiK the ability to represent hierarchies of consensus information in similar ways to other conflict management software.

University of Cape Town

Connectionist Organization – As described in Chapter 3 (*Knowledge Representations in Artificial Intelligence*), connectionists understand the brain as a system of interconnected states, where learning rules and history of connections affect future connections made. EpiK represents a very similar concept to connectionist models; people can be seen as nodes and the arcs between them are that of communication channels. In this way, knowledge does not only lie in the brains of the people, but also in the strengths of the connections between them. It is however vital to understand that despite the concepts being very similar, the mechanisms used in either case are substantially different.

In the same way that a connectionist model obtains stimuli from its environment (input states) and thereby produces a result and possible changes its internal structure, so too do organizations get stimuli from their external environment that consist of customers, suppliers, owners and competitors [vKR95]. Such stimuli trigger new activities in the organization and therefore promote production of new knowledge and could possibly force the creation of new communication channels in the connectionist organization.

Corporate memory models presented in Chapter 4 (*Corporate Memory Models*) all make use of some levels of formality in the fact that they impose certain knowledge types, relationship types, and structural restrictions on the user. Epik purposefully sets out not to prescribe specific formalisms for beliefs and avoids imposing any "types" on either knowledge or relationships between knowledge.

5.7 Epistemic Basis of EpiK

Now that the knowledge model has been presented, it is important to discuss the epistemological basis of EpiK, that is, the features of epistemology it has attempted to incorporate.

University of Cape Town

EpiK uses the traditional account of knowledge ('True Justified Belief') as its basis for rating the reliability of people and processes as sources of true and justified knowledge. As processes of justification, it uses the theory of Reliabilism. Therefore, no specific perspective about justification (such as empiricism or rationalism) is assumed, but rather to allow for both such theories as well as any other theories whose basis lies in Reliabilism. As pointed out in Chapter 2 (*Epistemology*), there are a number of theories of knowledge that are independent of the theory of Reliabilism. The knowledge model presented here takes into account one such theory, Consequentialism.

Reliabilism is represented in EpiK by allowing users to rate the truth of knowledge with respect to topics this knowledge is describing. Consequentialism is represented by allowing users to rate knowledge with respect to how useful it is in attaining certain goals.

Despite the different views of reliabilism and consequentialism with respect to what constitutes knowledge, the final result is a rating of knowledge sources (both people and processes) with respect to the truth of the knowledge they produce, and the usefulness of a piece of knowledge in attaining certain goals. More specifically, the fact that people can be rated according to truth and usefulness, allows EpiK to successfully represent social epistemology.

The method of information exchange within EpiK is analogous to the process of knowledge exchange described in Section 2.4 (*Social Epistemology*). This processes deals with *knowledge discovery*, *transmission*, *reception* and *acceptance*. In this model, discovery of knowledge happens through processes that a person has deemed reliable based on previous truth and usefulness ratings. Transmission decisions can be answered by searching to other individuals that have produced knowledge about similar topics. One the intended recipient receives the information he/she adopts an attitude towards the information by rating its truth and usefulness.

University of Cape Town

By representing the usefulness of information, and thereby being able to derive the usefulness of a knowledge source, the pragmatic perspective of knowledge in a corporate memory system is emphasised. In the same light as pragmatism, scepticism about the existence of a world beyond our own existence goes against the very essence of building a corporate memory system. The nature of a corporate memory system is such that one is forced into assuming the existence of an external world and that a person is able to know something without the need to guarantee its truth.

One of the objectives of corporate memory systems is to enable learning about areas of interest. Therefore, EpiK completely ignores sceptical arguments and assumes the existence of things, both concrete and abstract. The existence of things in the world is represented by *topics*. The collection of topics can be seen as an ontology of the objects in the world about which beliefs are formed. Similarly, the collection of *goals* can be seen as an ontology of the objectives of the people in the organization.

Although EpiK does not impose a fixed ontology on every user, the consensus mind contains an ontology that would probably be very acceptable by most users of the knowledge model because it is a representation of widely held beliefs, topics, and goals in the organization.

The structure of knowledge (the way in which a person relates beliefs to one another) is not specified. Therefore, no knowledge structure such as foundationalism or coherentism is imposed. Instead, a person is given the opportunity to relate beliefs in any way they see fit. Such freedom is a true reflection of the different ways in which people store knowledge, but it is also a stumbling block to making use of the relationships between beliefs as part of the reliabilist structure, that is, it is difficult to take relations between beliefs into account when calculating the reliability of a piece of knowledge.

EpiK itself has an ontology implicitly contained inside it. This area has not yet been fully investigated, but an example of a graphical representation of such an ontology can be found in Figure 21.

University of Cape Town

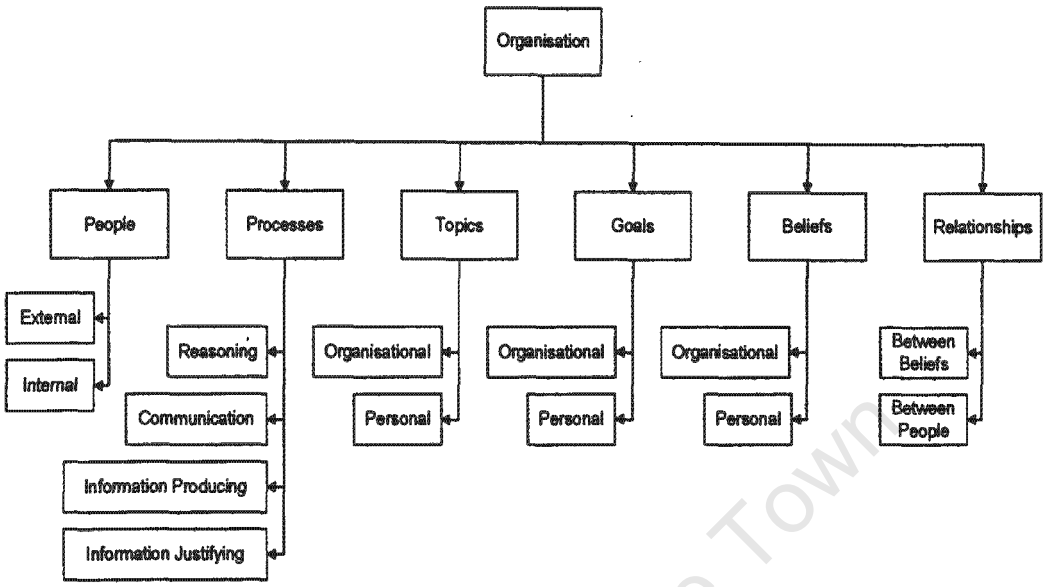


Figure 21: Ontology implicit in EpiK

The ontology in Figure 21 could be extended and modified in many ways. An example of a possible change to such a structure would be to take a worldly view by making "World" the root of the ontology and make the "Organization" a leaf. In this case, there would be "People" part of the world as well as part of the organization.

There are numerous ways in which to interpret the ontology implicit in EpiK, but one thing that can be noticed is that no matter how the ontology is viewed, there are a number of entities in it that overlap with entities in the Kant's ontologies (presented in Chapter 2, *Epistemology*). Such entities include Existence, Possibility, Causality and Community. EpiK represents *existence* through the representation of *topics*. *Possibility* is represented by *truth* and *usefulness* values. *Causality* is modelled by knowledge sources that are processes. *Community* is inherent in its structure through the clustering of information around certain topics and/or goals, as well as the clustering of people to form communicating groups.

University of Cape Town

The most important parallel that can be drawn between EpiK and epistemology is based on the existence of the consensus mind. If one takes the epistemological perspective of *reliabilism* and *consequentialism*, and therefore assumes that only very accurate or very useful information can attain the status of *knowledge*, then the goals, topics, knowledge sources and beliefs contained within the consensus mind describe completely the knowledge an organization possesses.

University of Cape Town

6 From Model to Corporate Memory

This chapter presents a detailed description of the corporate memory model described in the preceding chapter. If the reader is unfamiliar with the general structure of the EpiK, it is strongly suggested that Chapter 5 be read prior to this chapter. The aim of this chapter is to present a more detailed explanation of each of the components of EpiK and show that it is feasible to implement a corporate memory system based this model.

Section 6.1 discusses EpiK in more detail. Section 6.2 suggests user and system procedures to implement in a corporate memory system making use of EpiK. Section 6.3 presents prototype systems that have been developed using the model. Section 6.4 concludes by outlining the major components of EpiK and discussing some of the lessons learnt during development of the prototype systems.

University of Cape Town

6.1 Detailed model description

Due to the complexity of EpiK, it is better understood when viewed as being comprised of a number of different layers. These layers are:

- Individual Layer –the structures that maintain beliefs and relate beliefs to each other as well as to the topics, goals, and knowledge sources contained inside every mind.
- Organizational Layer –the mappings that relate topics, goals, beliefs and knowledge sources across multiple minds.
- Rating Layer – the structures that maintain the truth and usefulness values of beliefs with respect to topics and goals respectively.
- Meta-Knowledge Layer –the structures that store meta-knowledge about all entities and relationships in the system.
- Configuration Layer –the structures that contain the configuration parameters of a corporate memory system using this model.

This section describes each layer by means of an ER-Model and provides a general overview of the purpose of the entities and relationships. The purpose of the ER model is to help in the explanation of the knowledge model as well as to provide a conceptual model that serves as the basis for any corporate memory built to use the model.

6.1.1 Individual layer

The *individual layer* contains the structures that maintain beliefs, relations of beliefs to each other as well as to topics, goals, and knowledge sources for a single person involved in the creation and communication of knowledge.

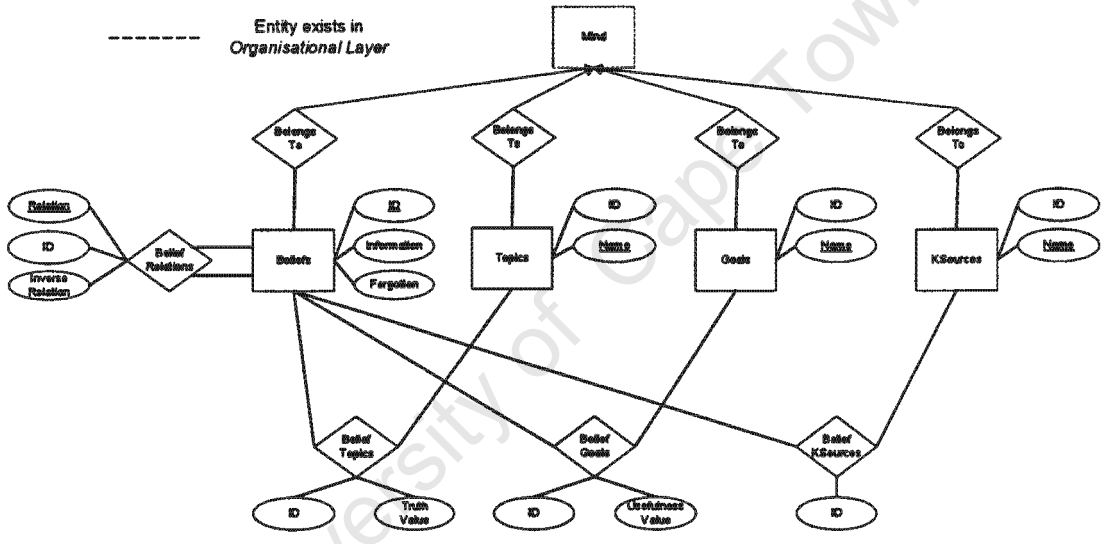


Figure 22: Individual Layer

The *Beliefs* entity contains the actual knowledge in the *Information* attribute. This knowledge is typically in the form of text, diagrams or documents. Beliefs are related to each other through *belief relations*. The *Forgotten* attribute is a flag that informs the system whether this belief has been deleted. It would not be correct to simply delete a belief because it may be part of a communication chain.

The *Topics* entity is a structure that contains the ontology for every person. That is, it is a structure that contains concepts both concrete and abstract about which beliefs are produced. The *Goals* entity describes all the goals of the individual, for example,

University of Cape Town

"improve customer service". The system would not enforce that every belief be related to one or more goals because in many cases a person would come across some knowledge that he/she might think could be useful at a later stage, but not right away.

Knowledge sources are contained in the *KSources* entity and could represent people or processes that produce knowledge. Epistemically speaking, processes could include introspection, different types of reasoning, or perception (to name just three).

Every belief entered into the corporate memory system would be related to a knowledge source through the *BeliefKSource* relation. It should be noted however, that in the situation that a belief is communicated to another person, the communicator is not necessarily the knowledge source. The knowledge source refers to the person or process that *authored* the knowledge, and not the communicator.

BeliefTopics relates beliefs to one or more topics. The truth-value represents the extent to which the user believes the knowledge to be an accurate reflection of a topic. *BeliefGoals* relates beliefs to one or more goals. The usefulness value represents the extent to which the user finds the knowledge useful in attaining the related goal.

The Topics, Goals and KSources entities could be implemented as sets or could have a hierarchical structure. This is, of course, an implementation issue.

University of Cape Town

6.1.2 Organizational layer

This layer represents the mappings that relate topics, goals, beliefs and knowledge sources across minds and therefore represents equivalences among many minds.

Every mind represents a person or a belief-producing agent with whom other minds can communicate. In general, a mind would represent a person who could be *in* the organization or *external* to the organization. When a person in the organization receives knowledge from an external individual, the system would create a mind for the external person, called a *virtual mind*.

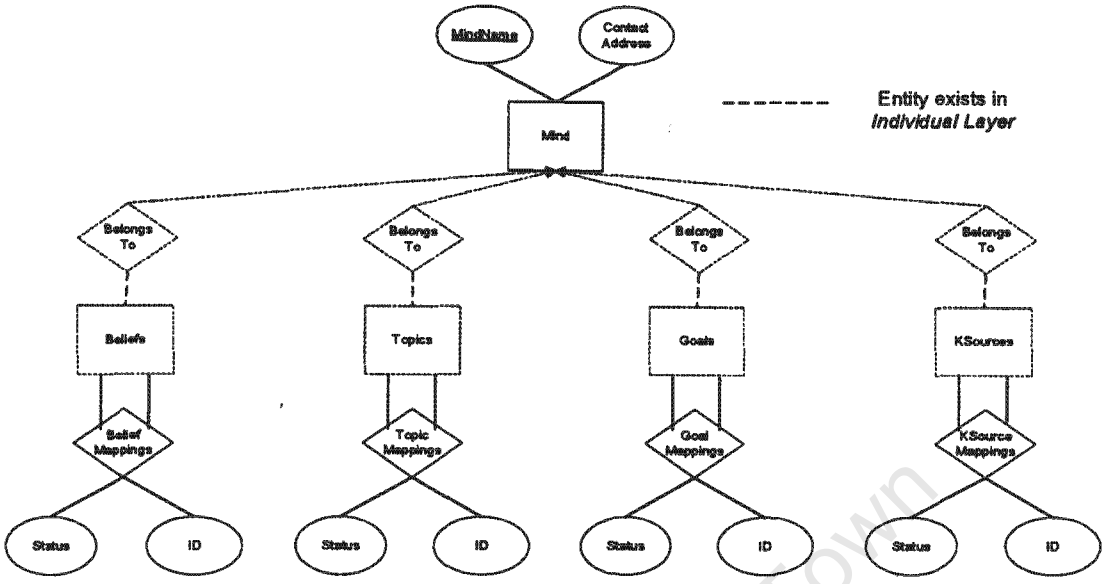


Figure 23: Organizational Layer

BeliefMappings, *TopicMappings*, *GoalMappings*, *KSourceMappings* link equivalent beliefs, topics, goals and knowledge sources found in different minds. *BeliefMappings* in particular represent communication paths of knowledge from one individual to another. Mappings can be automatically proposed, but will not be used until the knowledge manager has approved them. *Status* of mappings can either be *unknown* – when they have not yet been analysed, *accepted* by the knowledge manager, or *rejected*.

6.1.3 Rating Layer

The rating layer contains the structures that manipulate the truth and usefulness values of beliefs with respect to topics and beliefs respectively. Although the actual truth and usefulness values are stored in the BeliefGoals and BeliefTopics relations described in section 6.1.1 above, this layer makes use of these values to derive truth and usefulness values for knowledge sources when related to producing knowledge about certain topics or for attaining certain goals.

The entities in this layer also offer the user of the corporate memory model the ability to apply default truth and usefulness values to new knowledge based on criteria such as the knowledge source, the topic of the knowledge and the goal for which the knowledge is

University of Cape Town

useful. Not only are such default truth and usefulness ratings indicative of the way people perceive knowledge (based on the source, topics and goals), but also make EpiK more usable by not forcing the user to rate every piece of new knowledge that is acquired.

The rating layer could be seen as a set of rules that determine the method with which to automatically calculate truth and usefulness values for a belief with respect to a topic and/or goal. These types of default truth and usefulness ratings represent the initial perception a person has of a piece of knowledge based on who it is coming from, what it is about and what it is helping to achieve.

Although in reality humans apply many varying criteria and reasoning processes in determining the possible truth and usefulness of a piece of knowledge, allowing for a great amount of sophistication makes EpiK far less useable and prone to maintenance problems as well as detracting from the idea being put across.

University of Cape Town

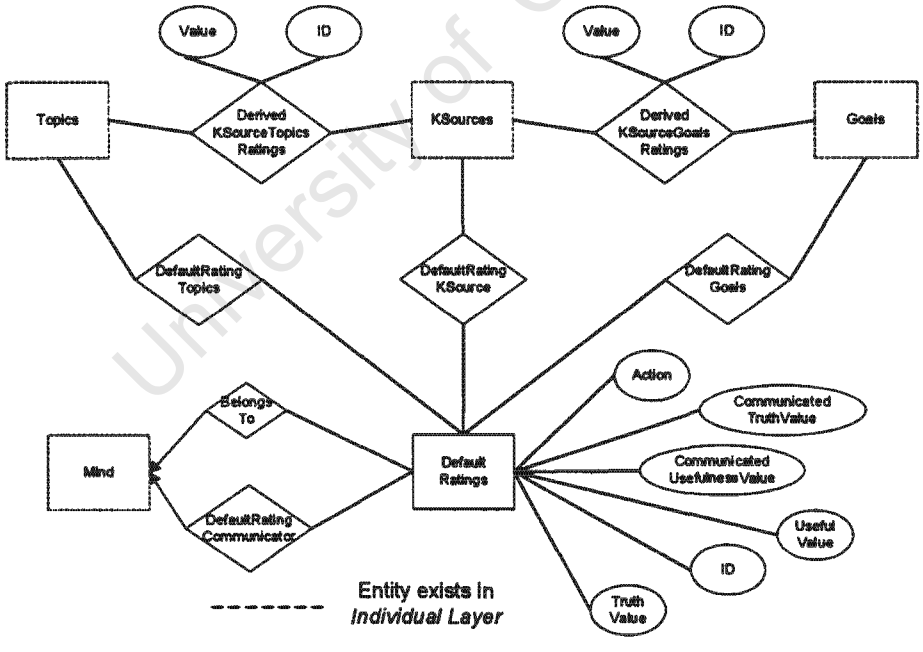


Figure 24: Rating Layer

As Figure 24 illustrates, each rule is based on four basic *conditions*: the knowledge source, the topics to which it relates, the goals for which it is useful and the communicator of the knowledge. Although each rule has the ability to be based on topics, goals, knowledge sources and other minds (based on the DefaultRatingTopic, DefaultRatingGoal, DefaultRatingKSource and DefaultRatingCommunicator relations), it is not essential that every rule be based on all four conditions. There could for example be rules based on only one of these conditions.

Based on these conditions, the user is able to take a number of different actions. These actions are operations that are applied to the numerical values contained in the other entities and relationships in EpiK such as the values found in the *value* fields of DerivedKSourceTopicsRatings and DerivedKSourceGoalsRatings relationships, or the truth and usefulness values that the communicator has sent along with the communication of a the belief. Possible actions include:

- *Apply value* – apply the truth and usefulness values found in the *TruthValue* and *UsefulValue* fields.
- *Apply derived values* – use values from DerivedKSourceTopicsRatings and DerivedKSourceGoalsRatings
- *Scale up derived values* - use values from DerivedKSourceTopicsRatings and DerivedKSourceGoalsRatings but scale them up by the values in the *TruthValue* and *UsefulValue* fields.
- *Scale down derived values* - use values from DerivedKSourceTopicsRatings and DerivedKSourceGoalsRatings but scale them down by the values in the *TruthValue* and *UsefulValue* fields.
- *Apply communicated values* - use values in the CommunicatedTruthValue and CommunicatedUsefulValue fields.
- *Scale up communicated values* - use values in the CommunicatedTruthValue and CommunicatedUsefulValue fields, but scale them up by the values in the *TruthValue* and *UsefulValue* fields.

University of Cape Town

- *Scale down communicated values* - use values in the CommunicatedTruthValue and CommunicatedUsefulValue fields, but scale them down by the values in the *TruthValue* and *UsefulValue* fields.

A far more complex set of actions is possible in an implementation, for example, actions that follow sophisticated mathematical models. Such issues are beyond the scope of this discussion.

The *DerivedKSourceTopicsRatings* and *DerivedKSourceGoalsRatings* relationships contain the *average truth-values* and *average usefulness values* of beliefs created by the related KSource about the relevant topic and goals. These values are derivable at runtime as a result of a query, but this information has been made explicit in EpiK because it is a vital part of the reliabilist theory this model represents. That is, every person starts to get a clear idea in his/her mind about how truthful or useful other people or processes are for creating knowledge. Such information is also helpful in evaluating the truth and usefulness of further knowledge attained from these knowledge sources.

6.1.4 Meta-Knowledge Layer

The Meta-Knowledge Layer contains the structures that store meta-knowledge about all other entities in the system. As mentioned in Section 5.2.5 (*Meta-Knowledge*), we avoid attempting to store all possible meta-knowledge in EpiK. Instead, a more epistemically correct definition of meta-knowledge is adopted; it is defined as "highly certain knowledge about beliefs that is not based on human input".

As a simple example, *Date Created* would be a piece of meta-knowledge stored when a new entry in any entity is created because the user does not need to input this value as it is obtainable from the operating system and is very probably correct.

The definition of meta-knowledge used does not limit the meta-knowledge to date and time values. In fact, if the system is able to obtain information from other highly reliable sources, then this information could also be used as meta-knowledge. For example, if one

University of Cape Town

has created a belief about stock market related issues, and the system is connected to a reliable source of stock market price information, then the current price of a certain stock item could be stored as meta-knowledge.

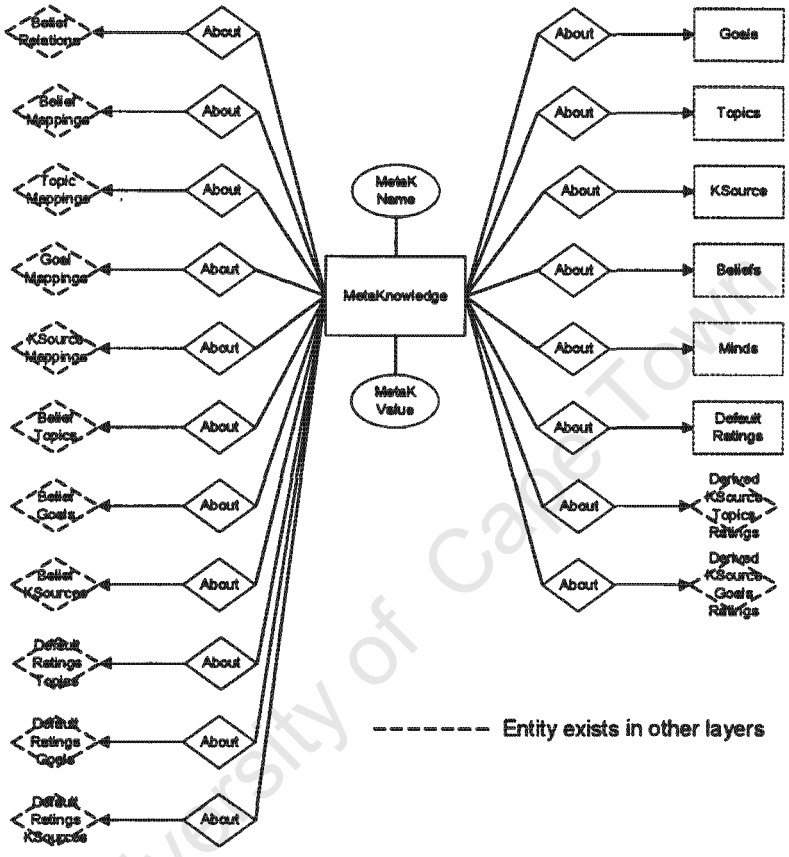


Figure 25: Meta-Knowledge Layer

Therefore, the meta-knowledge entity is comprised of only two fields, *MetaKnowledgeName* and *MetaKnowledgeValue* that contain the name and the value of the meta-knowledge. For example, if the system stores the date/time a belief is created, then the MetaKnowledgeName would be "Created Date-Time" and the MetaKnowledgeValue would contain the current date/time.

University of Cape Town

6.1.5 Configuration Layer

The configuration layer contains the structures that hold configuration parameters of EpiK to enable it to know which topics, knowledge sources, goals and beliefs to add to and remove from the Consensus Mind.

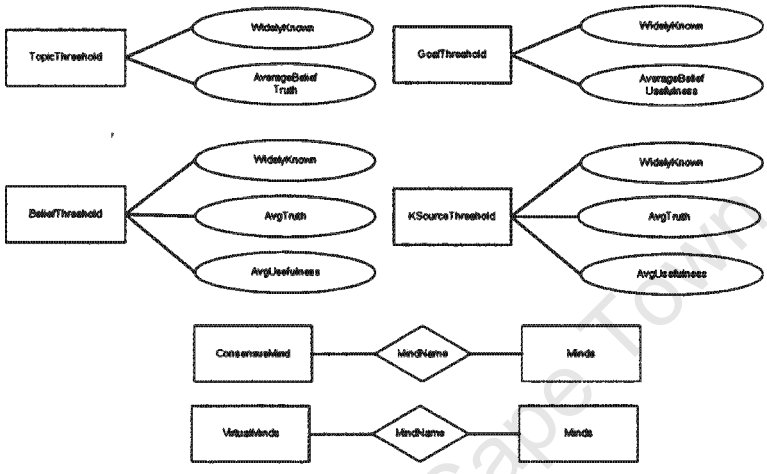


Figure 26: Configuration Layer

For a fully implemented system there would be numerous other entities in this layer, for example, information about presentation options, information about transport mechanisms, etc. Such information is not relevant to our discussions because this type of information is not relevant to the core of the knowledge model but is more relevant in a system design.

The *ConsensusMind* entity stores a reference to the shared organization mind that stores all widely known, and highly believed and useful topics, goals, beliefs and knowledge sources (Consensus Mind). The *VirtualMinds* entity stores references to the virtual minds. That is, minds that have been created as representations of people not in the organization, but with whom employees converse and share information.

The TopicThreshold, GoalThreshold, BeliefThreshold and KSourceThreshold entities maintain the conditions to which topics, goals, beliefs and knowledge sources must

University of Cape Town

adhere to be included in the consensus mind. These conditions include how widely known certain information is and how highly it should be rated with respect to truth or usefulness.

To calculate how widely an item of information is known, EpiK uses the *Mappings* relations described in Section 6.1.2. The mappings represent equivalence between information across minds, and therefore indicate the extent to which any information is known across the whole organization.

6.2 Using EpiK

The following processes describe (at a high level) the basic steps that the users would need to go through to use EpiK in a corporate memory system. These processes offer a guideline for system developers who wish to deploy a corporate memory system based on the knowledge model presented in this thesis. Although a full system has not yet been implemented, a number of screen-shots will be provided to allow the reader to visualize the type of system and user-interaction that envisaged for a corporate memory based on EpiK .

University of Cape Town

Outlines of the following processes are described:

1. Setup and Configuration
2. User processes
    a. Creating a mind
    b. Creating a virtual mind
    c. Creating a belief
    d. Communicating a belief
    e. Receiving a belief
    f. Rating a belief's truth and usefulness
    g. Finding a belief
    h. Deleting a belief
    i. Un-deleting a belief

3. Administrative processes
    a. Verifying mappings
    b. Maintaining the consensus mind

Every process is comprised of a number of steps. It is the responsibility of the user to complete some of these steps and the responsibility of the system to complete others. Either a *U* (to represent a user), or an *S* (to represent the system) will precede each of the steps described.

6.2.1 Setup and Configuration

This process will configure EpiK to enable it to make decisions about which beliefs, topics, goals and knowledge sources to bring into the consensus mind.

1. **U**: Create a mind for the organization (this is the Consensus Mind) by creating an entry in the *Mind* table with the name of the organization.
2. **S**: Create a reference in the ConsensusMind table that references the newly created mind.
3. **U**: Apply the necessary threshold values in the GoalThreshold, TopicThreshold, BeliefThreshold and KSourceThreshold tables.

6.2.2 User Processes

The processes outlined in this section describe how the user would create knowledge, communicate knowledge and learn.

Creating a Mind

A *User Mind* is created using the Consensus Mind as a template. This means that when a new user joins an organization, he/she does not start with a blank mind, but rather with the knowledge that his colleagues' feel is very truthful and useful at that point in time. Using the consensus mind as a template for new minds helps to get new users to use the system because of immediate availability of useful and truthful knowledge. The user is able to delete or edit the knowledge as necessary.

University of Cape Town

1. U: Create entry in the *Mind* table.
2. S: Duplicate the information found in the following tables from the *Consensus Mind* and relate them to the newly created mind: Goals, Topics, KSources, Beliefs, BeliefGoalRelations, BeliefTopicRelations, BeliefRelations, BeliefTopicRatings, BeliefGoalRatings, DerivedKSourceGoalRatings, DerivedKSourceTopicRatings

Creating a virtual mind

A *Virtual Mind* represents the mind of a person outside of the organization. It is necessary to create virtual minds when an employee is communicating with a person outside the organization. The difference between a normal user's mind and a *Virtual Mind* lies in the fact that virtual minds only contain beliefs and no other information.

When knowledge comes into the system from someone new who is outside the organization, or is communicated to a new person outside the organization, a virtual mind is created with the communicated belief inside it.

1. S: Add an entry into the Mind table
2. S: Add an entry into the Virtual Minds table with a reference to the newly created Mind

Creating a belief

This processes is used to add newly created knowledge into the corporate memory and relate it to relevant topics, goals and knowledge source.

1. U: Create the information and knowledge source (unless it already exists)
2. U: Relate the new belief to appropriate knowledge sources.
3. S: Based on the knowledge source, the system can *suggest* topics, goals and other beliefs that the new belief could be related to based on beliefs that have been previously related to this knowledge source.
4. U: Relate the belief to the topics, goals and other beliefs.

University of Cape Town

5. U: Rate the truth and usefulness of the belief. (See Process: *Rating a belief's truth and usefulness*]

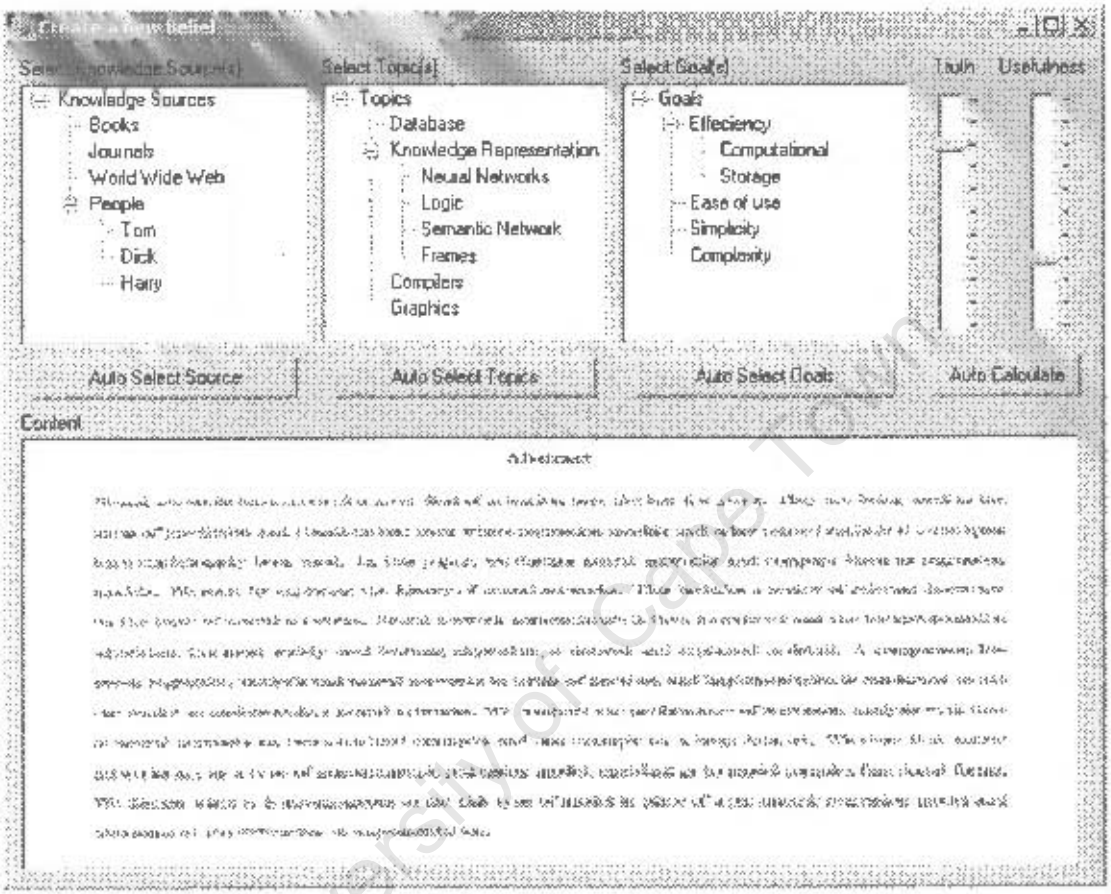


Figure 27: Screen-shot for a belief creation screen of a corporate memory system

Figure 27 illustrates a possible screen layout for the creation of a new belief. A user may choose to rate the usefulness and truth of a belief manually, or rate it according to the knowledge source. Furthermore, the system would have the ability to recommend topics and goals through analysis of the content of the belief.

Communicating a belief

This process is used to enable people to communicate and share knowledge. As part of the communication processes, it is not only the belief that is communicated, but also all

University of Cape Town

related topics, goals and knowledge sources and, if the communicator allows it, the truth and usefulness values are also communicated.

When a belief is communicated, an entry in the BeliefMappings table is created to keep track of communication paths of beliefs. Despite the fact that topics, goals and knowledge sources are also communicated along with the belief, no entries are made into the KSourceMappings, TopicMappings and GoalMappings tables. The reason for this is that it is the user's responsibility to relate the belief to the appropriate information; this is explained further in the *Receiving a Belief* process.

1. U: The user finds the belief that needs to be communicated using the *Finding a Belief* process.
2. U: Select the necessary beliefs from the result set and choose the recipients to whom this belief should be communicated.
3. S: Create a copy of the selected beliefs in the minds of the recipients. If one or more intended recipients are outside of the organization, then beliefs are copied into their virtual minds.
4. S: Recipients of beliefs will be notified of new incoming beliefs because communicated beliefs are not yet related to any topics or goals.
5. S: System creates entries in the BeliefMappings table.

Receiving a belief

Once a belief has been communicated from one person to another, the recipient needs to accept the belief, relate it to other beliefs, topics, goals and knowledge sources in his/her own mind as well as rate the truth and usefulness of this belief.

1. S: By making use of the TopicMappings, GoalMappings and KSourceMappings tables, the system can in many cases automatically relate received beliefs to topics, goals and knowledge sources in the mind.
2. S: In cases where the communicated belief is related to topics, goals and knowledge sources in the communicator's mind that are unknown to the

University of Cape Town

recipient's mind, the system offers the user the ability to discard these topics, goals and knowledge sources, create new ones, or relate them to topics, goals and knowledge sources that are similar in the recipient's point of view.

3. S: In the case where the user creates equivalences, entries are placed into the appropriate *mappings* tables and given a *status* of "unknown". The administrator of the system is left with the job of verifying the mappings (See Process: *Verifying mappings*)
4. S: Based on the knowledge source, the system can *suggest* topics, goals and other beliefs that the new belief could be related to based on beliefs that have been previously related to this knowledge source.
5. U: Relate the belief to the topics, goals and other beliefs.
6. U: Rate the truth and usefulness of the belief. (See Process: *Rating a belief's truth and usefulness*)

Rating a belief's truth and usefulness

When knowledge is created or received, the user is asked to rate the truthfulness and usefulness of the knowledge with respect to topics and goals respectively. As an additional step, the user can create default rating rules that will automatically rate knowledge according to conditions such as who the knowledge was received from, how or by whom the knowledge was created, what its about, or what it is being used for.

1. S: The system must first check if the belief in question complies with any of the rules specified in the DefaultRatings table. If such a rule exists, then the system must make use of the *action*, *truthvalue* and *usefulvalue* to apply truth and usefulness values in the BeliefTopics and BeliefGoals relations.
2. S: If no rule exists in the DefaultRatings table, then the system needs to suggest values for truth and usefulness based on the DerivedKSourceTopicsRating and DerivedKSourceGoalsRating tables.
3. U: User may use the suggested rating from the previous step or apply values manually.

University of Cape Town

4. U: Once values have been applied, the system should permit the user to create a rule in the DefaultRatings table in order to record this et of circumstances for the future.
5. S: The system updates the DerivedKSourceTopicsRating and DerivedKSourceGoalsRating tables to reflect the new truth and usefulness values.

Deleting a belief

When a user wants to forget a belief, he/she uses this process. Because beliefs are related to each other across minds, the system cannot actually remove the belief from the corporate memory because this could cause inconsistencies in the management reports.

1. U: The user finds the belief that needs to be deleted using the *Finding a Belief* process.
2. U: Select the necessary beliefs from the result set
3. S: In the Beliefs table, set the *forgotten* field to "true"

Un-deleting a belief

Users have the ability to undelete previously deleted beliefs because beliefs are never completely removed from the corporate memory.

1. U: The user finds the belief that needs to be undeleted using the *Finding a Belief* process.
2. U: Select the necessary beliefs from the result set
3. S: In the Beliefs table, set the *forgotten* field to "false"

Finding a belief

Due to the multitude of criteria that a user can use in trying to locate a belief, it would be extremely complex to require users to specify exact search criteria. A real-time, "active" search is proposed in which the user would always start off with a result set comprising of all beliefs in the corporate memory. The active set is then filtered as the user selects different conditions.

University of Cape Town

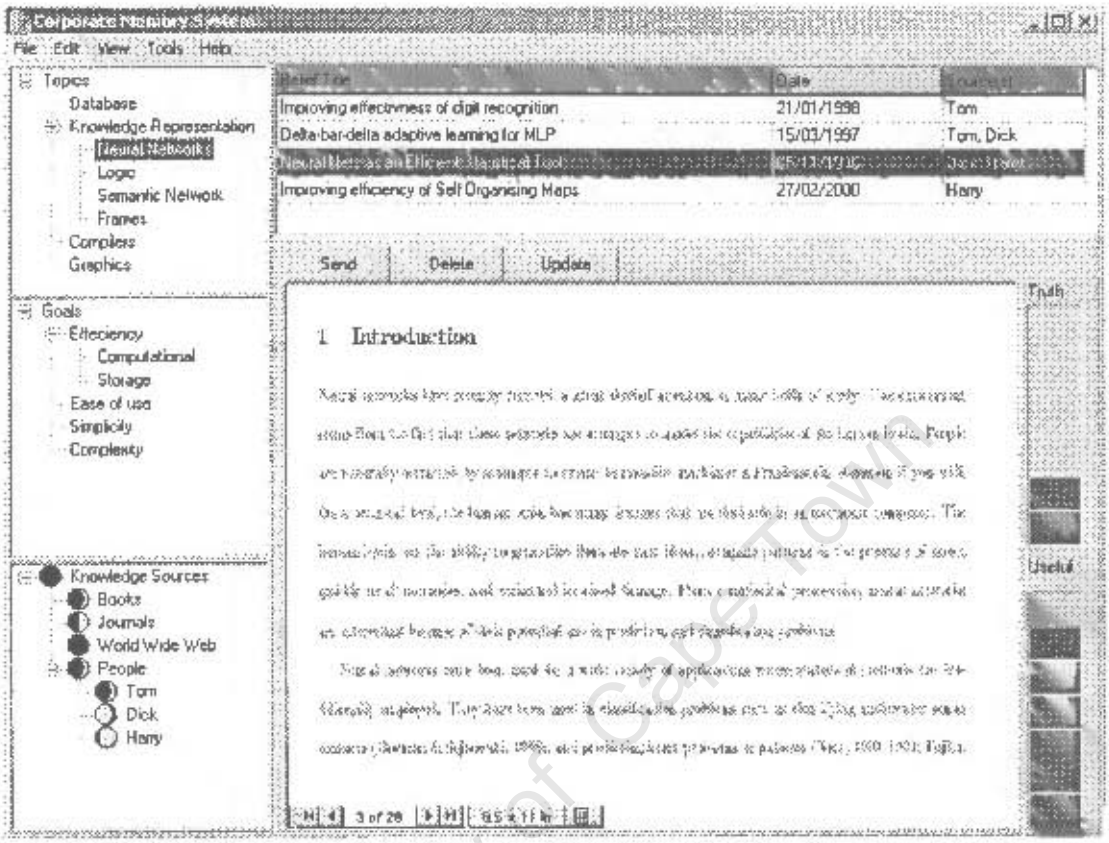


Figure 28: Screen-shot for a belief browsing screen of a corporate memory system.

As mentioned in the preceding chapter, and illustrated in Figure 28, there are many ways to query the model. Search criteria could include minds, topics, goals, knowledge sources, meta-knowledge, forgotten beliefs, truth-values, usefulness values, and communication channels details. Note how the screen-shot in Figure 28 emphasizes the truth and usefulness of a belief, as well as offers some type of metric as to the trustworthiness of knowledge sources.

6.2.3 Administrative procedures

The processes outlined in this section describe how the system administrator would maintain the consensus mind and ensure that goal, topic, and knowledge source mappings in the organizational layer (See Section 6.1.2, *Organizational layer*) are valid.

University of Cape Town

Verifying mappings

The knowledge manager needs to accept or reject all the mappings that are tagged as "unknown". Based on commonsense reasoning and understanding of the concepts being discussed, the status of mappings are changed to "accepted" or "rejected", whichever is appropriate.

Maintaining the consensus mind

The consensus mind can be maintained by automated processes that automatically identify widely known, true, and useful knowledge.

1. S: The system should make suggestions of new entries that are eligible to be included in the consensus mind based on averages of truth and usefulness values of beliefs, and making use of the mappings tables to calculate how widely known information is. These values are compared to the conditions found in the *threshold* tables.
2. S: For all the beliefs, topics, goals and knowledge sources that are accepted into the consensus mind, entries need to be added to the mappings tables.

In the same way as entries are added to the consensus mind, existing information in the consensus mind should be verified to be within the specified constraints, otherwise this information should be removed from the consensus mind.

6.3 Prototype Corporate Memory System

This chapter has presented details of how a corporate memory system should be built using EpiK. These details were used to develop a number of small implementations of the model along with a subset of the queries described in Section 5.5 (*Evaluating People and Knowledge from EpiK*). The aim of the experiment was to test the feasibility of implementing a corporate memory system based on the presented.

University of Cape Town

Twenty senior students were given a description of a subset of EpiK and asked to implement a system based on their own ideas and understanding of the concepts presented to them. A total of 5 different systems were developed by groups of 4 students each. Groups were given carte blanche as to their choice of language, deployment platform and interpretation of the problem at hand.

Four of the five groups decided to implement the system using the World Wide Web (WWW) as a platform for the user interface. One of the groups implemented a stand-alone windows application using a rapid application development tool. The groups making use of the WWW used either Java Server Pages or PHP3. The relational database used in all cases was Microsoft SQL Server.

6.3.1 Assignment Description

An extract of the description of EpiK given to students follows:

> *"The task is to build a corporate memory system that will manage the knowledge of individuals in an organization. The organizational memory must be comprised of many* ***minds*** *that represent individuals. Each mind is comprised of many beliefs. Beliefs are comprised of a* ***label*** *describing what is contained in the belief,* ***content*** *(the actual knowledge),* ***certainty value*** *that describes the extent to which this piece of knowledge is assumed to be true, a* ***certainty value description*** *that describes what factors were taken into account when rating the certainty of this belief, and* ***meta-knowledge****. Meta-knowledge is information that the system can obtain without asking the user for input, for example, the date and time.*
>
> *Within a mind, beliefs are related to each other via* ***relations****. Each relation can relate only two beliefs. Beliefs that are not contained in the same mind cannot be related to each other. Relations comprise a* ***label*** *that describes the relationship between the beliefs and* ***meta-knowledge*** *(as described above).*

University of Cape Town

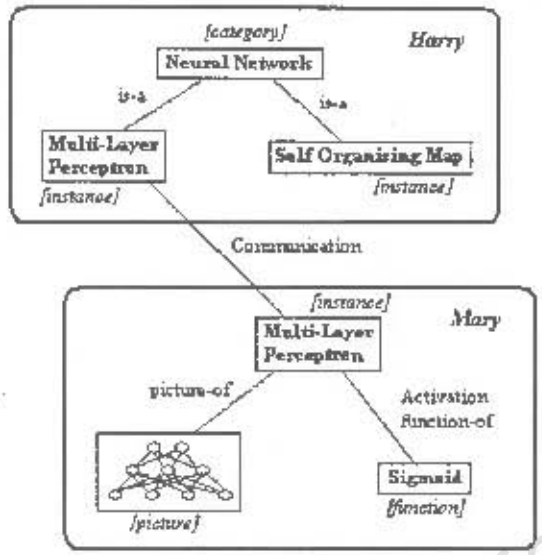


Figure 1: Example of two individuals' knowledge about Neural Networks

Individuals are given the opportunity to ***communicate*** *beliefs to each other, thereby forming communication lines between beliefs in different minds. An individual can communicate a belief to many people. Such a communicated beliefs can be stored in one of two ways, either as a duplicate of the belief that has been communicated, or as a* ***belief proxy*** *(for space efficiency). The key idea to remember is that the new person to whom the belief has been communicated can now relate this belief to other beliefs in his/her own mind. As with beliefs and relations, communication lines also have meta-knowledge related to them.*

Operations and Queries that the user should be able to execute include:

- *Adding a new belief into a mind*
- *Relating a belief to other beliefs in a mind*
- *Communicating a belief to other individuals*
- *Finding out which other individuals are in possession of a certain belief*
- *Finding out to what extent other individuals trust a certain belief*
- *Retrieving beliefs according to keywords in either relations labels, belief labels, belief content, certainty value, or meta-knowledge*

The knowledge managers in the organization need a different set of information from the system:

- *Who are the top knowledge creators?*

University of Cape Town

- *Who are the top consumers of knowledge? That is, people that get knowledge via communication but do not redistribute knowledge to other individuals*
- *Are there any communication groups forming? Are there certain individuals that communicate more with each other than with others?*
- *Which individuals are the most trusted, based on the certainty value that people have assigned to knowledge they have created?"*

It is clear from the extract above that for the purposes of this exercise EpiK had been simplified in the following ways:

- Students were not asked to represent goals, topics or knowledge sources.
- Beliefs are described as being related to only one topic through the existence of one *belief label* for every belief.
- Students were not required to implement the default rating rules; truth-values (or certainty-values) of beliefs with respect to a topic are values that a user must supply along with a reason for the supplied rating.
- A notion of a *belief proxy* was introduced as a mechanism for space efficiency.
- Knowledge manager queries were limited to retrieving information about the dynamics among the people.

6.3.2 Differing Interpretations of the Problem

The following observations were made during this exercise:

- All the groups developed two different views of the model; one for a user and one for a manager. All the groups implemented simple authentication mechanisms to identify the user and/or manager prior to system usage.
- One group identified the need for maintaining information about the originator of the knowledge. This was surprising seeing as the idea of a knowledge source was not mentioned to the students. Although this group only identified the origin of a belief as another person, it seems that their system could easily be extended to take into account other knowledge sources such as perception or reasoning.
- Four groups decided not to make use of belief proxies; however, only one of the groups elaborated on their reasoning for this decision. The reason given was that if an actual belief was modified and there exists a belief proxy that points to this

University of Cape Town

belief then the person in possession of the belief proxy can no longer claim to have understood that belief.

- One of the groups chose to represent meta-knowledge separately for each of the entities in the model, and not as one global set of information. The students claim that this makes the system more efficient as well as allowing for unique types of meta-knowledge to be stored for every entity.
- One of the groups identified the limitation of only being able to relate beliefs to one topic through the use of the *belief label*. To overcome this, the group defined a relation called "CATEGORY" and thereby allowed themselves to create beliefs that contain category names and relate beliefs to these categories. This group rightly claims that such an extension to the model results in substantial efficiency improvements in searching algorithms. This also allowed the group to retrieve information about which topics are "hot" in the company by analysing the topics most recently related to beliefs.
- One group correctly noticed that communication channels can be easily represented as belief relations with a pre-defined label, for example "COMMUNICATION".

6.3.3 Query Implementations

All the groups successfully implemented all the user queries specified in the assignment. However, the knowledge manager queries proved more difficult than expected for many of the groups. These difficulties were probably due to the need for complex, nested SQL SELECT statements.

It is interesting to note that one group not only implemented all the knowledge manager queries successfully, but also attempted to answer additional questions over and above the questions specified. They attempted to answer questions such as "*Who are the top communicators of knowledge?*". The group noticed that answering such a question is not simple because there is a need to take into account two variables; quantity of beliefs *sent* and quantity of beliefs *received*. The group attempted to combine these two variables into

University of Cape Town

one function that returned the communication strength of a person and were thereby able to answer questions of this nature.

6.3.4 Lessons Learnt

Some interesting lessons were learnt through the implementation of the prototype systems. Firstly, it seems easy to base a corporate memory system on EpiK, it also lends itself to development corporate memory systems with very different approaches. Secondly, details that were omitted to simplify the task were adopted as extraneous by some of the groups, thereby strengthening belief that the modelling primitives are both intuitive and useful. Thirdly, the complex graph structures of inter-related beliefs were very difficult to present to the user. This made navigation of beliefs difficult to understand, and the user often ended up being "lost" in the corporate memory, this problem was especially prominent with web browser based interfaces. Lastly, an implemented system making use of EpiK would need to maintain additional information to make it useful in practice. Such information includes: authentication information, information about beliefs that are unread, and belief presentation options.

6.4 Conclusion

EpiK has been described through the use of an ER-model divided into five layers. The *individual layer* represents the beliefs, goals, topics and knowledge sources of individuals. The *organizational layer* maps similar beliefs, topics, goals and knowledge sources across multiple minds. The *rating layer* maintains truth and usefulness values of beliefs. The *meta-knowledge layer* maintains meta-knowledge about all information contained inside the model, and the *configuration layer* stores the threshold values and conditions to which beliefs must conform to be accepted into the consensus mind.

Procedures that outline the usage of EpiK have been outlined along with procedures for maintaining the system and its components. Along with the ER-model, these procedures describe a design for a corporate memory system based on EpiK.

University of Cape Town

This rough system design has been simplified and implemented in different prototype systems in order to test the feasibility of developing a useable system based on EpiK. Despite the fact that the prototypes implemented only a subsection of the complete model, it has been shown that an epistemically based corporate memory model can be utilised effectively in a practical corporate memory system.

University of Cape Town

7 Evaluating and Extending EpiK

Numerous studies have been conducted in knowledge management literature that concern themselves with investigating human re-use of knowledge, sources of knowledge and methods of interaction between individuals.

[NEM00] investigates the frequency of usage of different media for collaboration among project participants. The survey collected data from nine virtual teams, distributed among nine companies. Seven out of the nine teams used email as a primary tool for collaboration among project members. The two teams that used other mechanisms of communication had email as the second most frequently used collaboration tool. An unrelated study done by [DP98] showed that people are more likely ask a person they know for information rather than searching for information themselves.

The studies done by [NEM00] and [DP98] makes one realize that communication, and particularly email, play a very important role in peoples' learning processes.

EpiK has many similarities to organizational email infrastructure. These similarities include:

- Individual minds – people maintain their own emails and attachments
- Individual Topics – people maintain their own classifications of information
- Knowledge Sources – every person has a unique address book of other people from whom they obtain information
- Communication channels – communication channels are represented by the existence of email from other people. In many cases it is possible to trace the route of emails through an organization by analysing the headers of the email message.

University of Cape Town

Fuelled by results of other studies as well as the similarities EpiK has with email, two studies were conducted in order to investigate other knowledge management issues as well as the possibility of integrating it with existing email infrastructure.

This chapter evaluates EpiK in two ways. Firstly, by presenting the results of the two studies, and secondly by comparing it's features to the corporate memory requirements outlined in Section 4.3 (*Corporate Memory Requirements*). Based on these evaluations, possible extensions are proposed.

Section 7.1 describes the two studies that were conducted. Section 7.2 compares the features of EpiK to the requirements outlined in Chapter 4 (*Corporate Memory Models*).

7.1 Studies Conducted

The first survey attempts to understand how people share information in the workplace, how they find information, and what role email plays in facilitating exchange of information.

The similarities EpiK has with email resulted in a second study being conducted that investigates the possibility of integrating it into existing email infrastructure. The only major feature of EpiK that does not exist in email is the ability to rate knowledge. Therefore, the second study was aimed at investigating if quantities of emails exchanged between people would be a good measure of trust and usefulness of information and thereby completely eliminating the need to rate truth and usefulness of information by basing these ratings on the amounts of emails exchanged.

It is important to note that these studies were conducted *after* the design of EpiK. This is in keeping with the objectives of this research: (a) not to be influenced by social design considerations for a knowledge model (b) to focus purely on epistemological concepts and (c) to see if epistemological concepts correlate to needs expressed in social studies.

University of Cape Town

This section reviews the experiments done and results obtained. Detailed results from both surveys can be found in Appendix A and Appendix B.

7.1.1 Knowledge Management Survey

This survey investigates the following:

- The extent of email usage in companies
- What type of information people find useful about other people
- The criteria used to rate the trustworthiness of information supplied by people
- Mechanisms used for distribution of information
- Mechanisms used for learning about new subjects
- The main sources of knowledge used
- The classification and retrieval habits of people in email environments

This survey is not aimed at being a definitive study of how knowledge is used in an organization, but rather as a validation for the relevance of EpiK. Because this research is focused more on the modeling aspect of organizational knowledge, knowledge management methodologies are not within the scope of this thesis, however, it is important to conduct a survey such as this as a "practicality" or "suitability" study.

The survey was completed by 200 individuals from different organizations and different fields of expertise including finance, information technology, marketing, professional (such as accountants and lawyers) and research and development environments. This section includes only relevant results and useful deductions that can be made from these results.

There are many more interesting results that were obtained from the online survey, the reader is encouraged to browse through Appendix A for information regarding: other uses of email applications, world-wide-web and interpersonal communication in knowledge acquisition, email packages used, personal habits in email application usage and effectiveness of current knowledge management tool.

University of Cape Town

Appendix A contains the complete questionnaire and a complete set of results.

Question #	Result	Implication
3, 12	Email is the primary mechanism for work related communication	A communication oriented corporate memory system should be integrated into an email tool.
4, 5, 7	Many people remember information about previous meetings with other people as well as other people's goals, attitudes and interests in their heads.	There is a need for a tool that would enable storage of information about people's goals, attitudes and interests
6	The most important factors determining whether a person would be asked for help are the trust built up from previous interactions, clarity of explanations and previous involvement in similar problems.	There is a need to represent trust and clarity of explanations about the people one interacts with in a knowledge model.
8, 9	Personal interaction is as important in obtaining knowledge as the world wide web.	Corporate memory systems need to model inter-personal communication. .
11	The majority of people say that they have only learnt approximately 50% of what they know from their colleagues.	There is a need to monitor communications external to the organization.
14	91.5% of people categorise knowledge that they receive from others.	Ability to categorise information is vitally important in a knowledge management system

Table 3: Summary results of knowledge management survey

One can conclude that an epistemically based knowledge model takes into account many of the needs expressed by respondents in the survey such as storage of trust relationships,

University of Cape Town

topics of interest and goals. An epistemically based model takes into account the importance of understanding inter-personal communication as a vital source of knowledge about specific subjects as well as about the dynamics of the workforce. Furthermore, integration into an email tool would be the most seamless way of introducing a corporate memory system into an organisation.

The study did however uncover areas where EpiK can be seen as lacking. In particular, it does not maintain information about knowledge *clarity*, nor does it maintain information about personal *attitudes* on organizational issues. Although EpiK could be easily changed to reflect these additional needs, these changes have not been made because it would no longer accurately represent epistemic theories of knowledge.

7.1.2 E-Mail Log Analysis

The aim of this experiment was to investigate the existence of a correlation between the quantity of email communication between people and the extent to which these communicating pairs trust each other and/or find each other useful. Finding such correlations would have great impact on EpiK; there would be no need to emphasise the feedback of users about the usefulness and truth of knowledge because these values could be deduced automatically based on the quantity of email communication between two people.

Permission was obtained from 22 people in the University of Cape Town Computer Science Department to make use of personal email logs over a period of 7 months. The 22 people comprised of 11 MSc students, 9 final year students, and 2 staff members (called *volunteers* for the rest of this discussion). The email log contained information regarding the sender's email address, recipient's email address and the date of communication. No details about the subject of the email nor the actual contents of emails were available.

The email log was used to tally emails sent and received between every person in the volunteer group and every other person with whom he/she communicates, not limited to

University of Cape Town

the volunteer group. In private interviews, the volunteers were asked to apply truth and usefulness values to people they have been communicating with as well as rate the approximate percentage of emails that is work related. Furthermore, volunteers were asked to identify increases or decreases in trust over time.

Due to the amount of communicating pairs produced, it was found to be unreasonable to request volunteers to complete truth and usefulness values for every person with whom they communicate. Therefore, there was a need to limit the dataset to people with whom the volunteers have *high communication strength* and *low communication strength*. This filtering process required a means of rating the strength of communication between two people.

High communication strength were attributed to people with high values for both incoming and outgoing messages, and low communication strength were associated with one-sided communication. With this in mind, an equation for communication strength was derived:

$$Communication\_strength_{AB} = \frac{(a+b)(a^2+b^2)}{2ab} \qquad (1)$$

where a represents the amount of emails sent from person A to person B, and b represents the amount of emails B sent to A. This equation was derived from the idea that the communication strength is proportional to the mean of the messages sent and received, and inversely proportional to the sum of the ratios between the sent and received email.

Correlations values of 0.93 were found to exist between the calculated communication strength value and the quantities of emails exchanged between people in both directions. This shows that the equation for communication strength was representative of actual communication quantities and therefore had the required characteristics.

Once the filtering process was complete the volunteers were left with a reasonable amount of people for whom they needed to apply truth and usefulness values.

University of Cape Town

Unfortunately, only 19 of the 22 volunteers could be prevailed upon to complete the required information: 11 MSc students, 7 final year students and 1 staff member.

A complete set of results can be found in Appendix B.

Results based on individual evaluation of trust and usefulness

The following results were obtained by analysing the values applied by volunteers to other individuals regarding truth and usefulness:

- Strong correlation between trust and usefulness was found.
- Strong correlation was found between the value used as an estimate of the percentage of emails that are work related to both trust and usefulness.

There was no correlation between communication strength on the one hand and trust and usefulness on the other hand. Therefore, one cannot make the conclusion that quantity of emails exchanged between people is representative of trust and usefulness perceptions. There was also no correlation between changes in communication strength over time and perceived changes in trust and usefulness.

Results based on group evaluation of trust and usefulness

The following results were obtained by analysing the values applied to groups of individuals regarding truth and usefulness. Here members of one group were asked to rate the usefulness and trustworthiness of another group. Groups included staff, MSc students, final year students, 3rd year students and 2nd year students.

- Weak correlations were found between communication strength, and usefulness and trust. This means that at a group level, quantity of inter-group communication could potentially be used as a measure of trust and usefulness of groups.
- Trust and usefulness values for groups were logical, having staff members as the most trusted and useful, MSc students as the next most trusted and useful group, followed by final year students, 3rd year students and 2nd year students in that order.

University of Cape Town

Comments about results obtained

During the interview process with the volunteers there were a number of interesting comments pointed out that could potentially be good explanations for the relatively poor results obtained in this experiment:

- It is difficult to rate the usefulness of a person in general; people are only useful for specific tasks.
- It is difficult to apply a rating describing the trust one has for another person in general. People primarily trust information produced about specific topics by specific individuals.
- Increases and decreases in communication over a short period of time do not represent increases or decreases in trust or usefulness. These fluctuations are primarily circumstantial, that is, they represent start and end dates of projects.

Despite the paucity of results, the findings are important and encouragingly supportive of ideas embodied in EpiK, as were the comments obtained as side-effects of this effort. With an extended dataset it is anticipated that more useful results for corporate memory developers would emerge and it is recommended that similar surveys be conducted at institutions where more complete communication information is available. More specifically, the existence of email content would enable more specific questions to be asked about the trust and usefulness of people with respect to specific subjects and goals. Furthermore, the existence of email logs over a number of years would have yielded more subtle, long term fluctuations in communication and thereby be more representative of growing trust relationships.

7.2 Evaluating Compliance to Corporate Memory Requirements

It is necessary to evaluate EpiK with respect to the relevant requirements outlined in Section 4.3 (*Corporate Memory Requirements*). Below is a table that lists the requirements as well as how EpiK meets those requirements.

University of Cape Town

#	Requirement	EpiK
1	Knowledge capture from multiple knowledge sources, at different levels of formalisation, using both declarative and procedural techniques	No limitation is placed on the internal structure of knowledge or on knowledge sources. Knowledge sources are explicitly represented.
2	Storage of user profiles that describe an individual's expertise, interests and goals	This information is derivable through analysis of the topics and goals the user relates to beliefs in his/her mind
3	Communication of knowledge according to the user profiles	There are no restriction on communications within EpiK. User profiles can be derived and used for effective distribution of information
4	Refinement of knowledge by exploiting user feedback that captures accuracy and usefulness of information	Users are able to rate knowledge according to its truth and usefulness
5	Storage of contextual and cultural information as well as knowledge annotations	These are not explicitly modelled, but there are no restrictions to the type of information being stored in EpiK
6	Storage of meta-knowledge	Meta-knowledge is determined according to the information that is automatically accessible to the system
7	Classification of knowledge	Knowledge is classified by the users according to topics as well as goals
8	Retrieval of knowledge through keyword search or navigation	EpiK can be used for learning through searching as well as navigation
9	Interaction with users through a variety of different sources	Standalone applications and web-browser interfaces have been developed in prototype systems
10	Representation of multiple viewpoints	Every user is able to maintain their own personal knowledge base within an integrated corporate memory

University of Cape Town

#	Requirement	EpiK
11	Integration into the workplace, including the ability to function in a geographically and temporally disparate organization	EpiK is very suited to a distributed environment. The model lends itself to integration into email infrastructure

Table 4: How EpiK meets the general requirements of corporate memory systems

Requirement 1 in Table 4 is concerned with the internal structure of an actual belief. EpiK does not prescribe knowledge representation at this level of granularity, although the general-purpose graph structure of the belief system could be used to represent almost any type of knowledge the user may require. It must be highlighted though that in this model, beliefs are intended to be self-contained objects such as documents.

Requirement 5 in Table 4 requires a level of formalisation to be applied to a corporate memory model. EpiK avoids predefining knowledge and relationship types such as 'Annotation' or 'Context', but could be extended to meet such requirements more specifically if appropriate knowledge and relationship types were to be defined.

7.3 Conclusion

The knowledge management survey showed that EpiK has correctly represented the needs of users by storing trust relationships, topics of interest, goals and inter-personal communication. However, the study found that there might be a further need to maintain opinions about the clarity of information as well as personal attitudes of people towards organizational issues as well as other people.

The study conducted to investigate the possibility of estimating truth and usefulness values from communication quantities, produced results that did not support a correlation between the quantity of exchanges between individuals and the trust relationship between those individuals. The participants in the survey struggled to rate the level to which they trusted other people because it was suggested that people only trust one another with respect to specific topics. Despite the disappointing results obtained when analysing communicating pairs, the results obtained when analysing communicating groups were

University of Cape Town

far more encouraging. It seems that inter-group communication is less ad-hoc and is limited to a few very specific topics. This is a possible explanation for the correlation found between email quantities and trust relationships between groups. It is possible that an extended dataset that could identify specific topics of communication would yield results that were far more representative of trust relationships between individuals.

As a final step in evaluation of EpiK, its features were compared to the corporate memory requirements outlined in Chapter 4 (*Corporate Memory Models*). The results of this comparison were very encouraging and showed that it complies with almost all the necessary requirements. The requirements with which it did not comply fully are requirements that are dependant on the existence of formalisms within the model. EpiK can be easily made to comply with these requirements by formalising knowledge and relationship types, however, to maintain generality and flexibility in the model these changes were not made.

University of Cape Town

8 Conclusion

This thesis presented work done in the fields of epistemology, artificial intelligence and corporate memory systems, and compared work in these fields. A new corporate memory model was presented along with two surveys conducted to validate the underlying assumptions of this model. This chapter reviews epistemology, artificial intelligence and corporate memory systems, summarizes the contribution of this work and makes recommendations for future work.

8.1 Summary

Thin spread of domain expertise, high levels of staff turnover and high costs of employee training are just some of the reasons that organizations are finding the need to manage the intellectual capital of their employees. In building systems that attempt to effectively represent many differing forms of information, knowledge-based system designers have been divided into two primary categories; that of artificial intelligence systems and that of corporate memory systems.

Artificial intelligence systems focus on execution of automated reasoning processes on small units of knowledge that have been structured so as to facilitate automated deduction, and thereby imitating experts in specific domains of expertise. These systems make use of knowledge representations such as first order logic, semantic networks, frames, production rules and neural networks. Each of these representations has its own advantages and disadvantages in terms of its effectiveness for specific problem solving tasks.

Corporate memory systems are used primarily for storage of knowledge in a manner that allows for effective information retrieval by humans. These systems are used to store large items of knowledge that are difficult to process computationally but are better understood by humans than by automated deductive processes. Typically, corporate

University of Cape Town

memory systems facilitate retrieval of information through keyword, attribute or navigational search methods.

Work done in artificial intelligence systems and corporate memory systems can be contrasted with work done in *epistemology*, the philosophical study of knowledge. Epistemology does not attempt to understand how experts solve problems, nor does it attempt to outline methods of information retrieval. Epistemology investigates only how people perceive the extent to which information is a true reflection of the world, and thereby suggesting theories about what people are justified in believing.

Social epistemology is the study of knowledge at a social level. It is the examination of communication processes used by people, as well as an examination of the knowledge obtained at a social level, that is, knowledge about the communication itself. In the same way that different processes are seen as reliable or unreliable in individualistic epistemology, so too are different people seen to have differing reliability values assigned to them. Such reliability values play an important role in initial evaluation of the quality of knowledge being received in a social environment.

Despite the differences that work done in corporate memory systems and artificial intelligence systems have with epistemology, one can identify a number of similarities between these fields of research. In epistemology, extreme rationalist viewpoints of knowledge have been found to have limitations when one considers that from a practical point of view, humans can never know things about the world without any doubt whatsoever. Similarly, knowledge representations that are able to represent uncertainty of information have proved to be very useful in solving real-world problems and have come to the forefront of knowledge representations used in artificial intelligence. Corporate memory systems have many parallels with concepts found in social epistemology. Systems have attempted to make use of mechanisms such as peer review processes and knowledge ranking during user feedback cycles in order to maintain a high level of knowledge quality in the corporate memory system.

University of Cape Town

8.2 Overview of EpiK

Chapter 5 (*An Epistemically Grounded Corporate Memory Model*) presented a corporate memory model called EpiK that is based on epistemic theories of reliabilism and consequentialism. It assumes that all information is contained in peoples' minds and not in a shared repository. Every person has his/her own set of beliefs, methods of rating the truth and usefulness of beliefs, view of the external world, set of goals, and set of processes and people that are sources of knowledge. To enable exchange of knowledge, EpiK has relations between beliefs in different minds that are representations of communications that occur between people.

By having the user rate the truth and usefulness of knowledge with respect to how accurately it represents the subject matter and how useful the knowledge is for attaining specific goals, EpiK is able to derive the reliability of different knowledge sources and can therefore help in the initial evaluation of the truth and usefulness of new knowledge.

An important construct represented in EpiK is that of the *consensus mind*. This consensus mind is a collection of widely known items of knowledge that considered very accurate, and very useful. To select the items of knowledge that are worthy of being part of the consensus mind, EpiK traces communication channels of beliefs to understand how widely a belief is known, and employs the usefulness and truth values to determine the reliability of that knowledge.

The *consensus mind* is useful for a number of reasons; firstly, to evaluate how useful different people and/or processes are in producing accurate and useful results. Secondly, to get an understanding of what the core competencies are within the organization. Thirdly, because it contains knowledge that has been perceived as very accurate or very useful by other individuals in the organization, it can be used as an entry point for users during a learning process. Lastly, and most importantly, if one takes the epistemological perspective of reliabilism and consequentialism, and assumes that only very accurate or very useful information can attain the status of knowledge, then the goals, topics,

University of Cape Town

knowledge sources and beliefs contained within the consensus mind describe the information an organization is justified in knowing.

8.3 Evaluation of EpiK

EpiK was evaluated in a number of ways. Firstly, a knowledge management survey was conducted that investigated how people find information and how they share information in the workplace. Secondly, an attempt was made to investigate how well it would integrate into the email infrastructure of an organization. Thirdly, a number of small prototype systems were developed to prove the feasibility of building a working system based on EpiK. Lastly, EpiK's features were shown to comply with corporate memory system requirements outlined in relevant literature.

The knowledge management survey showed that EpiK possesses many of the features needed by respondents. Such features include representation of people's interests and goals, as well as being able to represent the knowledge of persons external to an organization. Along with the ability to enable learning through traditional searching mechanisms, EpiK also enables learning by allowing one to find other people with similar interests. Features that it does not represent are indicators of knowledge clarity and personal attitudes.

The second study was done to investigate the possibility of integrating EpiK into existing email infrastructure by trying to estimate the truth and usefulness of people through analysis of communication strengths between individuals and between groups. Finding such correlations would mean that it would not be essential to force users to rate knowledge according to truth and usefulness because these values could be derived automatically. There was no correlation between quantity of communication and trust when analysing communicating pairs of individuals. However, correlations were found when analysing the communication strength and trust relationships between groups of people. It is possible that a more complete data set that contains actual topics of communication between people (and not simply quantity of communication) might yield better results in analysis of trust relationships between individuals.

University of Cape Town

The development of a number of prototype systems helped to understand some of the extensions that were needed to EpiK at an implementation level. Such extensions included an authentication system and a powerful user interface. However, the most important thing that the development of these prototypes showed was the fact that a real-world corporate memory system could easily be built using EpiK.

Over and above the evaluations conducted, it was shown how EpiK could be used to evaluate individuals involved in exchange of knowledge, obtain an understanding of the social dynamics within an organization.

The strength of EpiK lies in its elegant simplicity. It is a realisation of the fundamental theories about the philosophy of knowledge. Whereas most work in the field of corporate memory systems is focussed on areas of usability and functionality, the focus of EpiK is simply to represent the concept of epistemic justification by modelling the extent by which a person is justified in believing information.

Through the innovative combination of personal levels of justification, EpiK is able to answer some of the most crucial questions organisations can ask about their intellectual capital. With a quick glance at the queries EpiK can answer (Section 5.5, *Evaluating People and Knowledge from EpiK*), one immediately notes the type of high-level knowledge management information EpiK can supply.

EpiK has shown that representing corporate knowledge using well-established philosophical theories results in a very person-oriented understanding of the knowledge within an organisation. With this view of corporate knowledge, EpiK is able to supply one with the most crucial information about any organisation, that is, what a company can truly claim to *know*.

University of Cape Town

8.4 Future Work

There are many possibilities for extending EpiK at both an implementation level as well as at a conceptual level. Possible extensions to the model are the inclusion of a measure of knowledge clarity, personal attitudes, a more sophisticated meta-knowledge structure, more sophisticated default rating rules, more granular consensus management, and the implementation of text classification algorithms. The theoretical grounding of the model could also be extended through the repetition of the email log analysis study, and through the investigation of technologies that can be used for automation.

Knowledge Clarity and Personal Attitudes

Section 7.1.1 (*Knowledge Management Survey*) concluded that *information clarity* and *personal attitudes* are important aspects in evaluating the usefulness of knowledge and a decisive parameter used by many people for deciding whether or not to approach a certain individual for information. It seems that such measures would be very useful to many users of corporate memories. They could be applied by adding fields to the *Beliefs* entity and creating an *Attitudes* entity for every mind and relating it to goals, topics and knowledge sources. The reason that clarity and attitudes were not included in EpiK along with truth and usefulness is because they have no epistemic grounding. The model would also become complex, compared to the proposal here which offers power through simplicity.

Meta-Knowledge Structures

If the system is implemented to store all possible meta-knowledge it is capable of storing every time information in any of the entities is altered, then the meta-knowledge table would get too large for any useful manipulation. In a practical system implementation additional configuration information describing when to store meta-knowledge could be maintained. This would enable the system to store limited meta-knowledge and thereby making the system less cumbersome.

Rating layer extensions

The rating layer uses a very simple set of conditions and actions. The conditions could be implemented using very complex decision trees (rule-bases) and the actions could be

University of Cape Town

implemented to model sophisticated mathematical equations. The rule structure that has been presented in Section 6.1.3 only allows for rating of knowledge according to its usefulness and truth. A very important aspect to rating the truth of knowledge that has been left out in EpiK is *coherence*. Therefore it should be extended to allow for rating of knowledge based on the coherence an item of knowledge has with the rest of a person's belief system. For example, to measure the coherence a belief has to the rest of a belief system, one could ask the question "count how many relations labeled 'agree-with' that relate to the belief in question exist in my belief system". Therefore, the rating layer of EpiK should be able to rate knowledge according to relations that have been attached to the knowledge in question.

Hierarchies of Consensus Minds

EpiK only has the ability to represent consensus information for the organization as a whole. It may prove useful to represent more granular consensus information for subgroups of selected individuals in an organization, as well as consensus between such subgroups. The model can be extended by allowing the system to generate multiple consensus minds, each for a selected group of individuals. These consensus minds can then be used for creating more higher level consensus minds between groups of individuals. This would give EpiK the ability to represent hierarchies of consensus information.

Extending the theory

In order to simplify the usability of the model, it would be a worthwhile exercise to repeat the email log analysis study (section 7.1.2) on a larger and more complete set of data to investigate the possibility of automating the rating of trust and usefulness of people and processes according to communication quantities. Future work should also look at mechanisms for exploiting the model effectively through the use of algorithms and data structures that allow for efficient computation of the consensus mind through the use of the mappings, developing appropriate user interfaces for learners and managers, and agent technology. Agents can partially automate the processes of knowledge classification, derivation of people's interests and competencies, and learning from the corporate knowledge.

University of Cape Town

University of Cape Town

9 References

[ABH+98] Abecker A., Bernardi A., Hinkelmann K., Kuhn O., Sintek M., (1998). *Towards a Technology for Organizational Memories,* IEEE Intelligent Systems and their Applications, 13(3), 40-48

[ALS89] Alston W.P., (1989), *A 'Doxastic Practice' Approach to Epistemology.* Empirical Knowledge: Readings in Contemporary Epistemology, Second Edition, Rowman & Littlefield Publishers Inc., pg 269-301

[ANN78] Annis D.B., (1978), *A Contextualist Theory of Epistemic Justification.* Empirical Knowledge: Readings in Contemporary Epistemology, Second Edition, Rowman & Littlefield Publishers Inc., pg 205-217

[AUD93a] Audi R., (1993), *The Foundationalism—Coherentism Controversy.* Empirical Knowledge: Readings in Contemporary Epistemology, Second Edition, Rowman & Littlefield Publishers Inc., pg 143 – 188

[AUD93b] Audi R., (1993), *The Structure of Justification,* Cambridge University Press

[AUD98] Audi R., (1998), *Epistemology, A Contemporary Introduction to the Theory of Knowledge.* Routledge London and New York

[BG80] Bobrow, Goldstein, (1980) *Representing Design Alternatives,* Proceedings AISB Conference, Amsterdam

[BK86] Bobrow D.G., Katz R.H., (1986), *Context Structure/Versioning: A Survey, On Knowledge Based Management Systems,* Springer-Verlag

University of Cape Town

[BLD00] Barrett W., Lau L.M.S., Dew P.M., (2000), *Facilitating Knowledge Transfer.* An R&D Environment: A Case Study in Internet-Based Organizational Memory and Knowledge Management, 2000, Edited by David G. Schwartz, Monica Divitini, Terje Brasethvik, Idea Group Publishing

[BMD99] Buckingham Shum S., Motta E., Domingue J., (1999), *Representing Scholarly Claims in Internet Digital Libraries: A Knowledge Modeling Approach.* Proceedings of ECDL '99: Third European Conference on Research and Advanced Technology for Digital Libraries, Paris France, September 1999.

[BON85] BonJour, (1985), *The Structure of Empirical Knowledge*, Cambridge, Mass.: Harvard University Press

[BRE33] Brentano F., (1933), *Kategorienlehre*, translated by R.M. Chsholm & N. Guterman as *The theory of Categories*, Martinus Nijhoff publishers, The Hague, 1980

[BUC96] Buckingham Shum S., (1996), *Negotiating the Construction of Organizational Memory Using Hypermedia Augment Spaces.* Workshop on Knowledge Media for Improving Organizational Expertise, 1st International Conference on Practical Aspects of Knowledge Management.

[BUC97a] Buckingham Shum S., (1997), *Representing Hard-to-Formalized, Contextualized, Multidisciplinary, Organizational Knowledge.* AAAI '97 Spring Symposium on Artificial Intelligence in Knowledge Management, Stanford University CA. AAAI Press

[BUC97b] Buchingham Shum S., (1997), *Balancing Formality with Informality: User Centred Requirements for Knowledge Management Technology.* AAAI '97 Symposium on Artificial Intelligence in Knowledge Management 1997, Stanford University, CA, AAAI Press.

University of Cape Town

[BUX97] Buxton H., (1997) *Visual Interpretation and Understanding.* Artificial Intelligence, University of Edinburgh, http://www.dai.ed.ac.uk/CVonline/LOCAL_COPIES/ECVNET/BUXTON//interpretation.html

[BW77] Bobrow D.G., Winograd T., (1977), *An Overview of KRL, a Knowledge Representation Language.* Readings in Knowledge Representation, 1985, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque. Originally in *Cognitive Science 1(1), 1977, 3-46.*

[CKI88] Curtis, B., H. Krasner, & N. Iscoe, (1988), *A Field of the Software Design Process for Large Systems.* Communications of the ACM, 31 (11)

[CON89] Conklin, J., (1989) *Design rationale and maintainability*, Proceedings of the Twenty-Second Annual IEEE International Conference on System Sciences, Vol 2 Hawaii

[CON96] Conklin, E., (1996) *Designing organizational memory: Preserving intellectual assets in a knowledge economy.* Electronic publication by Corporate Memory Systems, Inc. [Online]. Available: http://www.gdss.com/wp/DOM.htm

[CSSvH96] Clancey W., Sachs P., Sierhuis M., van Hoof R., (1996), *Brahms: Simulating Practice for Work Systems Design.* In Compton P., Mizoguchi R., Motoda H., Proceedings of PKAW'96

[DBS77] Davis R., Buchanon B., Shortlife E., (1977), *Production Rules as Representation for a Knowledge Based Consultation Program.* Readings in Knowledge Representation, 1985, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque. Originally in *Artificial Intelligence 8(1), 1977, 15-45.*

University of Cape Town

[DCGR98] Dieng, R., Corby, O., Giboin, A., and Ribiere, M. (1998). *Methods and tools for corporate knowledge management.* Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'98). Banff, Alberta, Canada, 1998. Edited by Gaines, B.R. & Musen, M

[dHBM+94] de Hoog R., Benus B., Metselaar C., Vogler M., Menezes W., (1994). *Organization model: Model definition document.* Technical report, University of Amsterdam and Cap Programmator. Deliverable DM6.2c of ESPRIT Project P5248 (KADS-II)

[dKDSS77] de Kleer J., Doyle J., Steele G.L., Sussman G.J., (1977), *AMORD: Explicit Control of Reasoning, Readings* in Knowledge Representation, 1985, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque. Originally in *Proc. Symposium on Artificial Intelligence and Programming Languages, SIGPLAN Notices 12(8).*

[DP98] Davenport, T.H., Prusak, L. (1998), *Working Knowledge: How organizations manage what they know*, Harvard Business School Press, Boston.

[DSS93] Davis R., Shrobe H., Szolovits P., (1993), *What is Knowledge Management.* AI Magazine, 14(1): 17-33, http://medge.lcs.mit.edu/ftp/psz/k-rep.html

[EAS91] Easterbrook S., (1991), *Handling Conflict Between Domain Descriptions with Computer Supported Negotiation.* Knowledge Acquisition: An International Journal, Vol 3, pp 255-289

[EAS93] Easterbrook S., (1993), *Domain Modelling With Hierarchies of Alternative Viewpoints*, IEE International Symposium on Requirements Engineering

University of Cape Town

[EFKN94] Easterbrook S., Finkelstein A., Kramer J., Nuseibeh B., (1994), *Co-ordinating Distributed Viewpoints: the Anatomy of a Consistency Check.* Concurrent Engineering: Research and Applications

[EUZ96] Euzenat J., (1996), *Corporate Memory through co-operative creation of knowledge bases and hyper-documents*, KAW'96, Banff Alberta, Canada, Proceedings online at http://ksi.cpsc.ucalgary.ca/KAW/KAW96/KAW96Proc.html

[FB99] Faure A., Bisson G., (1999), *Modelling the Experience Feedback Loop to improve Knowledge Base re-use in industrial environment*, KAW'99, Banff Alberta, Canada, Proceedings online at http://sern.ucalgary.ca/ksi/KAW/KAW99/papers.html

[FES98] Fensel, D., Erdmann, M., Struder, R. (1998), *Ontobroker: The Very High Idea*, Proceedings of the 11th International Flairs, Conference (FLAIRS-98), Sanibal Island, Florida, May 1998

[FGKN89] Finkelstein, A.C.W., M. Goedicke, J. Kramer, & C. Niskier (1989), *Viewpoint oriented software development: Methods and Viewpoints in Requirements Engineering.* Proceedings, Second Meteor Workshop on Methods for Formal Specification, Springer-Verlag.

[GET63] Gettier, E., (1963), *Is Justified True Belief Knowledge?* Analyzis 23: 121-123

[GOL79] Goldman A., (1979), *What is Justified Belief?* Justification and Knowledge, edited by George Pappas. Dordrecht: Reidel

[GOL91] Goldman A., (1991), *Epistemic Folkways and Scientific Epistemology.* Liaisons: Philosophy Meets the Cognitive and Social Sciences, Cambridge, Mass: MIT Press

[GOL99] Goldman A., (1999), *Knowledge in a Social World*, Clarendon Press. Oxford

University of Cape Town

[HAY79] Hayes, P.J., (1979), *The Logic of Frames*. Readings in Knowledge Representation, 1985, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque. Originally in *Frame Conceptions and Text Understanding, 46-61, edited by D.Metzing, 197.9*

[HEI75] Heidegger M., (1975), *Die grundprobleme der Phanomenologie*, translated by A. Hofstadter as *The Basic Problems of Phenomenology*, Indiana University Press, Bloomington.

[HOT00] Hotdispatch.com, (2000), *The online marketplace for technical expertise*, http://www.hotdispatch.com

[HUS13] Husserl E., (1913), *Ideen auf einer reinen Phanomenologie und phanomenologischen Philosophie*, translated by F. Kersten as *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*, Kluwer, Dordrecht, 1982

[IBM00] IBM Corporation, (2000), *IBM's text analyzis tools*, http://www.software.ibm.com/data/iminer/fortext/tatools.html

[KA97] Kuhn, O., Abecker, A. (1997). *Corporate memories for knowledge management in industrial practice: Prospects and challenges*. Journal of Universal Computer Science, 3(8), 929-954

[KAC86] Katz R.H., Anwarrudin M., Chang E., (1986), *Organising a Design Database Across Time*. On Knowledge Base Management Systems: Integrating Artificial Intelligence and Database Technology, Springer-Verlag

[KAN87] Immanuel Kant, 1787, *Kritik der reinen Vernunft*, translated by N. Kemp Smith as *Critique of Pure Reason*, St. Martin Press, New York.

University of Cape Town

[KHJ00] Kantrowitz M., Horstkotte E., and Joslyn C., (2000), *FAQ: Fuzzy Logic and Artificial intelligence*,
http://www.cs.cmu.edu/Web/Groups/AI/html/faqs/ai/fuzzy/part1/faq.html

[KR95] von Krogh G., Roos J., (1995), *Organizational Epistemology*, New York: St. Martin Press

[LB84] Levesque H.J., Brachman R.J., (1984), *A Fundamental Tradeoff in Knowledge Representation and Reasoning* (Revised Version). Readings in Knowledge Representation, 1985, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque. Originally in *Proc. CSCSI-84, London, Ontario, 1984, 141-152.*

[LEH86] Lehrer K., (1986), *The Coherence Theory of Knowledge*. Empirical Knowledge: Readings in Contemporary Epistemology, Second Edition, Rowman & Littlefield Publishers Inc., pg 121-142

[LEH90] Lehrer K., (1990), *Theory of Knowledge*, Boulder, Colorado: Westview Press

[LG90] Douglas B. Lenat, R.V. Guha, (1990) *Building Large Knowledge-Based Systems*, Addison-Wesley, Readings, MA

[MAKM99] Menzies T., Althoff K.D., Kalfoglou Y., Motta E.. (1999), *Issues with Meta-Knowledge*, Software Engineering Knowledge Engineering, 1(1).

[MBD99] Motta E., Buckingham-Shum S., Domingue J., (1999), *Case Studies in Ontology-Driven Document Environment*, 12th Banff Workshop on Knowledge Acquisition, Modelling, and Management (KAW '99)

University of Cape Town

[McC77] McCarthy J., (1977), *Epistemological Problems of Artificial Intelligence.* Readings in Knowledge Representation, 1985, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque. Originally in *Proc. IJCAI-77, Cambridge, MA, 1977, 1038-1044*

[ME98] Martin, P., Eklund, P. (1998), *WWW Indexation and Document Navigation Using Concept Strictures*, Proceedings of the 2nd International Conference on Intelligent Processing Systems, IEEE Press, pp. 217-221

[MEN99] Menzies T., (1999), *Knowledge Maintenance Heresies: Meta-Knowledge Complicates KM*, Software Engineering, Knowledge Engineering, SEKE '99

[MH99] Maurer F., Holz H., (1999), *Process-Oriented Knowledge Management for Learning software organizations*, KAW'99, Banff Alberta, Canada, Proceedings online at http://sern.ucalgary.ca/ksi/KAW/KAW99/papers.html

[MIN74] Minsky M., (1974), *A Framework For Representing Knowledge*, Memo 306, Artificial Intelligence Lab MIT.

[MOO82] Moore R.C., (1982), *The Role of Logic in Knowledge Representation and Common-sense Reasoning.* Readings in Knowledge Representation, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque, 1985. Originally in *Proc. AAAI-82, Pittsburgh, PA, 1982, 428-433.*

[MOS96] Moser P.K., (1996), *Empirical Knowledge: Readings in Contemporary Epistemology*, Second Edition, Rowman & Littlefield Publishers Inc.

[MW99] Masterton S., Watt S., (1999), *Oracles, Bards and Village Gossips, Or, Social Roles and Meta-knowledge Management*, Simon Masterton, Stuart Watt (KMI-TR-82), August 1999

University of Cape Town

[MYL86] Mylopoulos J., (1986), *On Knowledge Base Management Systems*. On Knowledge Base Management Systems: Integrating Artificial Intelligence and Database Technology, Springer-Verlag

[NEM00] Nemiro J. E., (2000), *The Glue That Binds Creative Virtual Teams*. Knowledge Management and Virtual Organizations, Idea Group Publishing, edited by Yogesh Malhotra, 2000

[ORA00] Oracle Corporation, (2000), *ConText Cartridge* http://www.oracle.com/st/o8collateral/html/xctx5bwp.html

[PLA93] Plantinga, (1993), *Warrant: The Current Debate*, Oxford: Oxford University Press

[REI98] Reiner U., (1998), *Knowledge Integration for Building Organizational Memories*, KAW'98, Banff Alberta, Canada, Schedule online at http://ksi.cpsc.ucalgary.ca/KAW/KAW98/KAW98Sched.html

[RK91] Rich E., Knight K., (1991), *Artificial Intelligence*, 2nd edition, McGraw-Hill

[ROB89] Robbins, S.P. (1989), *Organizational Behaviour: concepts, controversies, and Applications*, (fourth edition) Prentice Hall, NJ.

[ROB90] Robinson, W. N. (1990), *Negotiation Behaviour During Multiple Agent Specifications: A Need for Automated Conflict Resolution*. Proceedings of the International Conference on Software Engineering.

[RUM89] Rumelhart, (1989), *The Architecture of Mind: A Connectionist Approach. Foundations of Cognitive Science*, Cambridge, MA: Bradford/MIT Press

University of Cape Town

[SAY97] Sayre K.M., 1997, *Belief and Knowledge: Mapping the Cognitive Landscape*, Rowman & Littlefield Publishers Inc.

[SR96] Scaife M., Rogers T., (1996), *External Cognition: How do Graphical Representations work?* International Journal of Human-Computer Studies, 45, 185-213.

[SC00] Snis U., Carstensen P.H., (2000), *On Knowledge Management: A Field Study.* Internet-Based Organizational Memory and Knowledge Management, 2000, Edited by David G. Schwartz, Monica Divitini, Terje Brasethvik, Idea Group Publishing

[SCHI91] Schubert L.K, (1991), *Semantic Nets Are In The Eye Of The Beholder*, Principles of Semantic Networks: Exploration in the Representation of Knowledge, Edited by John F. Sowa, Morgan Kaufmann Publishers

[SDB00] Schwartz D.G., Divitini M., Brasethvik T., (2000), *On Knowledge Management in the Internet Age.* Internet-Based Organizational Memory and Knowledge Management, Edited by David G. Schwartz, Monica Divitini, Terje Brasethvik, Idea Group Publishing, pg 1-23

[SEM00] Semino, (2000), *Semino text mining tool* http://www.qsr.com.au/QSRHome.htm

[SFB+87] Stefik M, Foster G, Bobrow D.G., Kahn K., Lanning S., Suchman L., (1987), *Beyond the Chalkboard: Computer Support for Collaboration and Problem Solving in Meetings*, Communications of the ACM 30 (1)

[SG88] Shaw M.L.G, Gaines, B.R. (1988) *A Methodology for Recognising Consensus, Correspondence, Conflict and Contrast.* Knowledge Acquisition Systems, Proceedings, Third Knowledge Acquisition For Knowledge-Based Systems workshop, Banff, November

University of Cape Town

[SIM96] Simon G., (1996), Knowledge Acquisition and Modelling for corporate memory: lessons learnt from experience, KAW 96, Banff Alberta, Canada, Proceedings online at http://ksi.cpsc.ucalgary.ca/KAW/KAW96/KAW96Proc.html

[SMI82] Smith B.C., (1982), *Prologue to "Reflection and Semantics in Procedural Language"*, Readings in Knowledge Representation, 1985, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque. Originally in *PhD. Thesis and Tech. Report MIT/LCS/TR-272, M.I.T. Cambridge, MA, 1982.*

[SOS91] Sosa E., (1991), *Reliabilism and Intellectual Virtue.* Empirical Knowledge: Readings in Contemporary Epistemology, Second Edition, Rowman & Littlefield Publishers Inc., pg 189-204

[SOW01] Sowa J., (2001), *Conceptual Graphs.* http://www.bestweb.net/~sowa/cg/index.htm, Accessed June 2001

[SPA98] Sparrow J., (1998), *Knowledge in Organizations: Access to thinking at work,* Sage Publications

[STE96] Steup, 1996, *An Introduction to Contemporary Epistemology*, Prentice Hall

[SZO86] Szolovits P., (1986), *Knowledge-Based Systems: A Survey.* On Knowledge Base Management Systems: Integrating Artificial intelligence and Database Technologies, Springer-Verlag, http://medge.lcs.mit.edu/ftp/psz/Szolovits85b.mss

[TAHZ00] Tschaitschian B., Abecker A., Hackstein J., Zakraoui J., (2000), *Internet-Enabled Corporate Knowledge Sharing and Utilisation* in *Internet-Based Organizational Memory and Knowledge Management*, Edited by David G. Schwartz, Monica Divitini, Terje Brasethvik, Idea Group Publishing

University of Cape Town

[TAS97] Tschaitschian, B., Abecker, A., Schmalhofer, F. (1997), *Information tuning with KARAT: Capitalizing on existing documents*. In E. Plaza and R. Benjamins (Eds.), Proceedings of the 10th European Workshop on Knowledge Acquisition, Modeling and Management (EKAW'97), Sant Feliu de Guixols, Catalonia, Spain, October 15-18, 269-284

[vdSS97] van der Spek, B.R. and Spijkervet, A.L. (1997). *Knowledge Management: Dealing intelligently with knowledge*, Kenniscentrum CIBIT in Internet-Based Organizational Memory and Knowledge Management, Edited by David G. Schwartz, Monica Divitini, Terje Brasethvik, Idea Group Publishing

[vHvdSK96] van Heijst, G., van der Spek, R., Kruizinga, E. (1996). *Organizing corporate memories*. 10th Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'96). Banff, Alberta, Canada, 1996. Edited by Gaines, B.R., Musen, M. (Eds.), Proceedings online http://ksi.cpsc.ucalgary.ca/KAW/KAW96/KAW96Proc.html

[vHIP88] von Hippel, E. (1988), *The Sources of Innovation*, MIT Press

[vKR95] von Krogh G., Roos J., (1995), *Organizational Epistemology*, St. Martin Press: New York

[VNJ00] Voss A., Nakata K., Juhnke M., (2000), *Concept Indexes: Sharing Knowledge from Documents*. Internet-Based Organizational Memory and Knowledge Management, Edited by David G. Schwartz, Monica Divitini, Terje Brasethvik, Idea Group Publishing

[WBW86] Wiederhold G., Blum R.L., Walker M., (1986), *An Integration of Knowledge and Data Representation*, On Knowledge Based Management Systems, Springer-Verlag

University of Cape Town

[WEB86] Webber B.L., (1986), *Questions, Answers and Responses: Interacting with Knowledge-Base Systems*, On Knowledge Based Management Systems, Springer-Verlag

[WEB00] WebMind text classification products, (2000), http://www.webmind.com

[WF87] Winograd T., Flores F. (1987). *On Understanding Computers and Cognition: A New Foundation for Design.* A Response to the Reviews. Artificial Intelligence, 31:250-261

[WHI29] Whitehead A.N., (1929), *Process and Reality: An essay in Cosmology*, corrected edition, D.R. Griffin & D.W. Sherburne, eds., Free Press, New York 1978

[WIN75] Winograd T., (1975), *Frame Representations and the Declarative / Procedural Controversy.* Readings in Knowledge Representation, 1985, Morgan Kaufmann Publishers, edited by Ronald J. Brachman, Hector J. Levesque. Originally in *Representation and Understanding: Studies in Cognitive Science, 185-219, edited by D.G. Bobrow and A.M.Collins, 1975*

[WOO91] Woods W.A., (1991), *Issues in Knowledge Representation*, Principles of Semantic Networks: Exploration in the Representation of Knowledge, Edited by John F. Sowa, Morgan Kaufmann Publishers

[ZM95] Zdrahal Z., Motta E., (1995), *The Trouble with What: Issues in Method-Independent Task Specification.* Proceedings of the 9th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop edited by B.R.Gaines and M.Musen, pp. 30-1 - 30-17.

[ZM96] Zdrahal Z., Motta E., (1996), *Improving Competence by Integrating Case-Based Reasoning and Heuristic Search.* Proceedings of the10th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop, edited by Gaines B.R., Musen M., November 1996, pp. 13-1 - 13-12

University of Cape Town

Appendix A

The following is a duplicate of the questionnaire used in the knowledge management survey including results obtained.

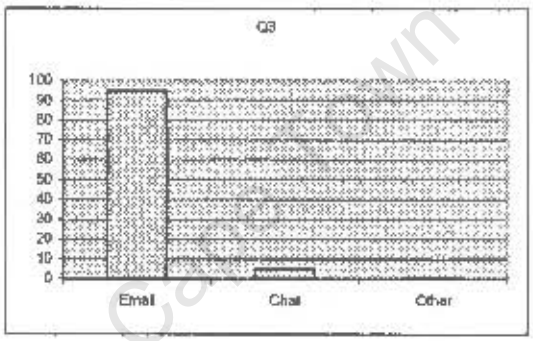
1. What is your email address? (optional)

2. What is email address of the person that referred you to this questionnaire? (optional)

General Info

3. Which medium of electronic communication do you *most* frequently use in the *workplace*?

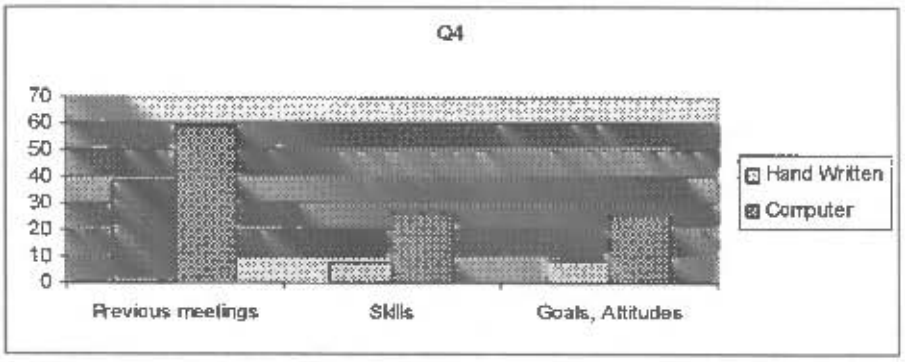
- ☐ Email
- ☐ Chat
- ☐ Video and/or Audio conferencing
- ☐ Other: .....



Information about people

4. Do you keep records of any of the following information about people you interact with at work?

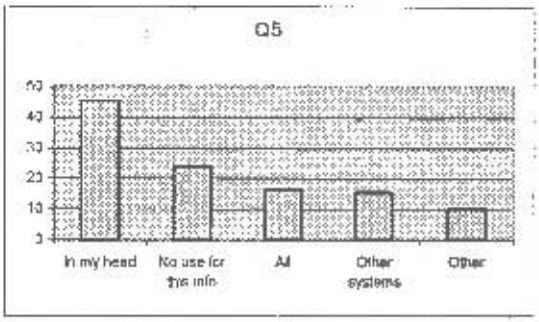
Hand written	On computer	
☐	☐	Notes about previous meetings and/or conversations you had with this person
☐	☐	Work related skills that a person has
☐	☐	Goals and attitudes of this individual regarding work related issues
	☐	I don't keep any records



University of Cape Town

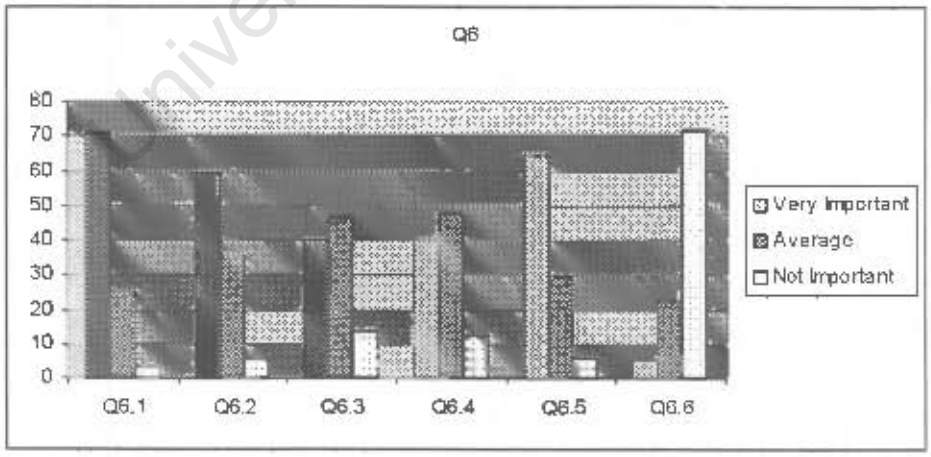
5. Why did you not select some of the items in Question 4 above?

- I made a selection on every line!
- I have no need to remember such information
- I remember this information in my own head
- I have access to such information through systems that are maintained for the company as a whole
- Other



6. Are there people from whom you seek advice about work related issues? If so, rate the importance of the reasons why you turn to these people for advice:

	Not Important	Average Importance	Very Important
i) These person(s) offer reliable information based on previous requests	☐	☐	☐
ii) These person(s) are clear in their explanations of complex topics	☐	☐	☐
iii) These person(s) widely accepted as leaders in their fields	☐	☐	☐
iv) I think these person(s) are leaders in their fields	☐	☐	☐
v) I know the person was previously involved in solving a similar problem	☐	☐	☐
vi) I have nobody else to turn to	☐	☐	☐

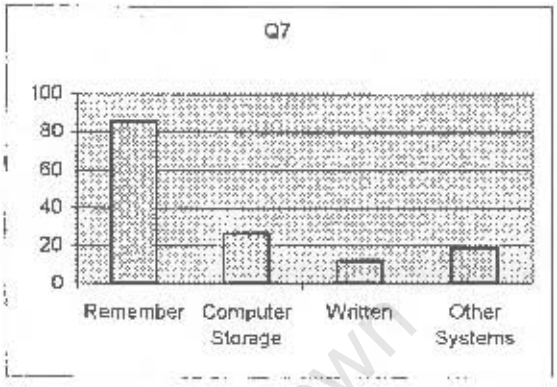


University of Cape Town

Information Distribution

7. Being in possession of an item of information that you think might be valuable to others, what conditions would you use to decide to whom you should send this information?

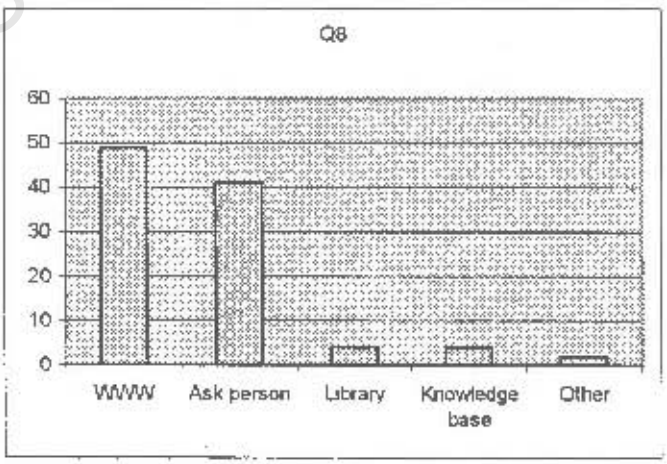
- I use my own judgement based on past experiences that I *remember*
- I use my own judgement based on past experiences I have *stored on my computer*
- I use my own judgement based on past experiences I *write down*
- I have access to such information through systems that are maintained for the company as a whole
- I never send any information to others
- I never receive any information worth sending to others



Learning

8. Given the task of learning about a new work related subject, what would be your *first* step in attaining knowledge about this field?

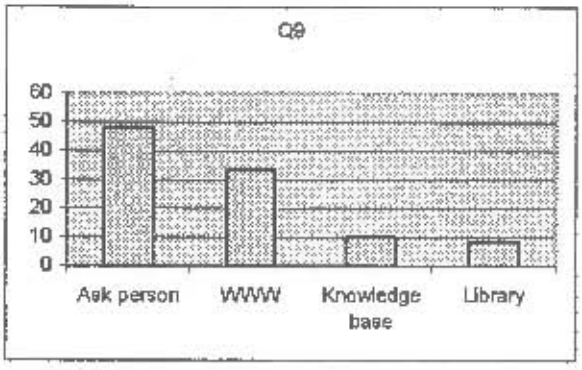
- Ask a person who has knowledge of this field to explain the primary concepts to you and recommend further readings
- Try to find information using the World Wide Web (WWW)?
- Try to find information using a library?
- Try to find information using your companies document repository or knowledge management system?
- Other: _____
- I have never found the need to learn anything new



University of Cape Town

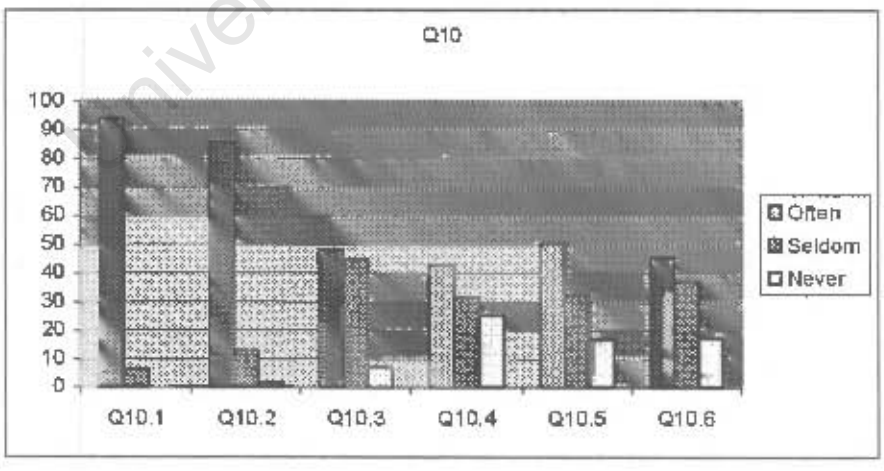
9. Assuming your first step (Question 8 above) was *unsuccessful*, what would be your *second* step in attaining knowledge about this field?

- ☐ Ask a *different person* who has knowledge of this field to explain the primary concepts to you and recommend further readings
- ☐ Try to find information using the World Wide Web (WWW)?
- ☐ Try to find information using a library?
- ☐ Try to find information using your companies document repository or knowledge management system?
- ☐ Other ……
- ☐ I have never found the need to learn anything new



10. How often do you make use of the following sources of information in the workplace?

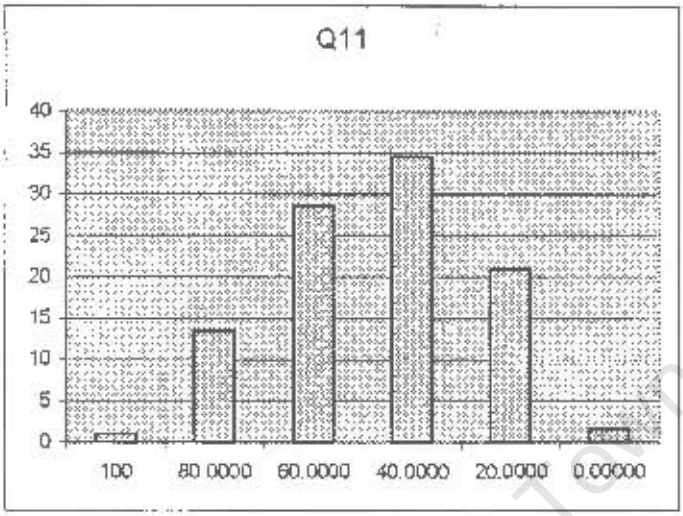
	Never	Seldom	Often
i) Communication with people	☐	☐	☐
ii) World Wide Web	☐	☐	☐
iii) Books & Journals	☐	☐	☐
iv) Document repository	☐	☐	☐
v) Research Results	☐	☐	☐
vi) Magazine articles	☐	☐	☐



University of Cape Town

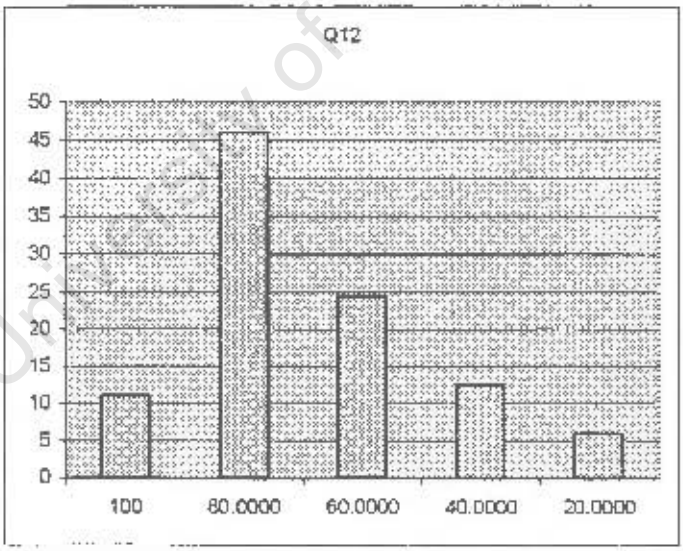
11. How much of what you know about your job do you estimate you learn or have learnt from colleagues?

☐ Nothing ☐ 20% ☐ 40% ☐ 60% ☐ 80% ☐ 100%



12. How much of your work related communication is done electronically?

☐ Nothing ☐ 20% ☐ 40% ☐ 60% ☐ 80% ☐ 100%

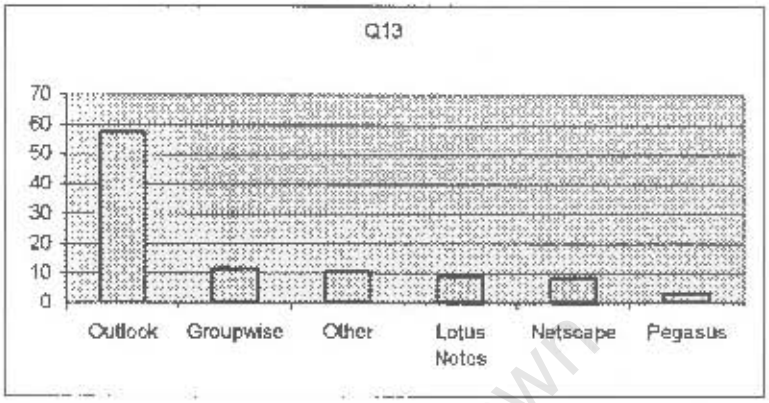


University of Cape Town

Email related questions

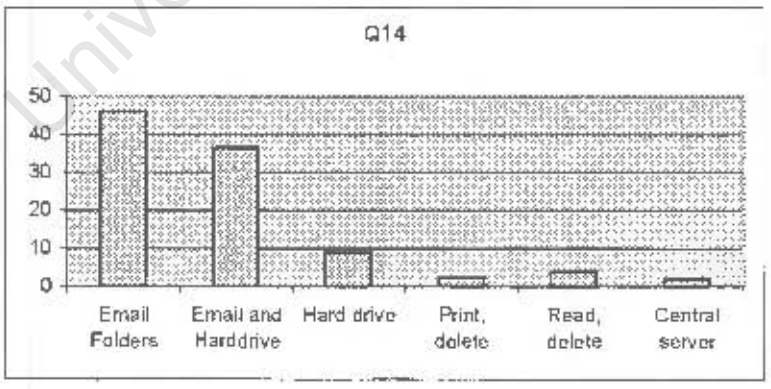
13. What *email* software do you use in the workplace?

- ☐ Microsoft Outlook
- ☐ Groupwise
- ☐ Lotus Notes
- ☐ Netscape
- ☐ Pine
- ☐ Pegasus Mail
- ☐ Other
- ☐ I don't know



14. What do you do with any *work related* material you receive via email?

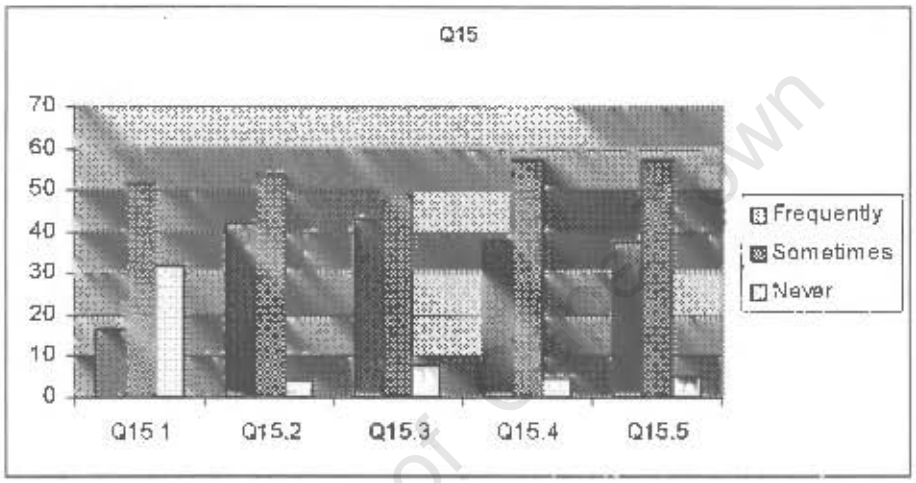
- ☐ I file it into folders in my email software
- ☐ I file relevant information into folders on my hard-drive
- ☐ I file some information into email folders and some to my hard-drive
- ☐ Read it and then delete it
- ☐ I read it and delete it because my company stores work related documents in a central archive
- ☐ Print it and delete it
- ☐ I don't know
- ☐ I don't receive work related email



University of Cape Town

15. Do you ever refer back to previously *received* email? If so, please rate the frequency of reference for every reason below:

		Never	Sometimes	Frequently
i)	To see who has been sending me email about certain topics or issues	☐	☐	☐
ii)	To see what somebody had to say about certain topics or issues	☐	☐	☐
iii)	To find an answer to a question or problem	☐	☐	☐
iv)	To forward the email on to other people that would have an interest in the information contained in the email	☐	☐	☐
v)	To retrieve a document (possibly stored as an attachment) that might otherwise be lost	☐	☐	☐



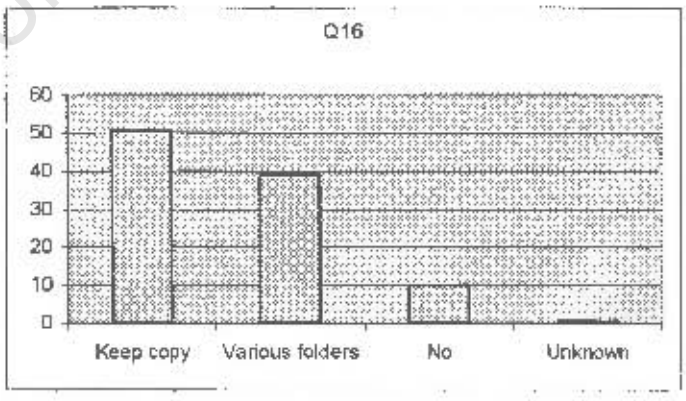
16 Do you organise (file away) email that you *send*?

☐ No

☐ No, but I keep a copy in a folder in my email application (for example, in the *Sent* folder)

☐ Yes, in various folders

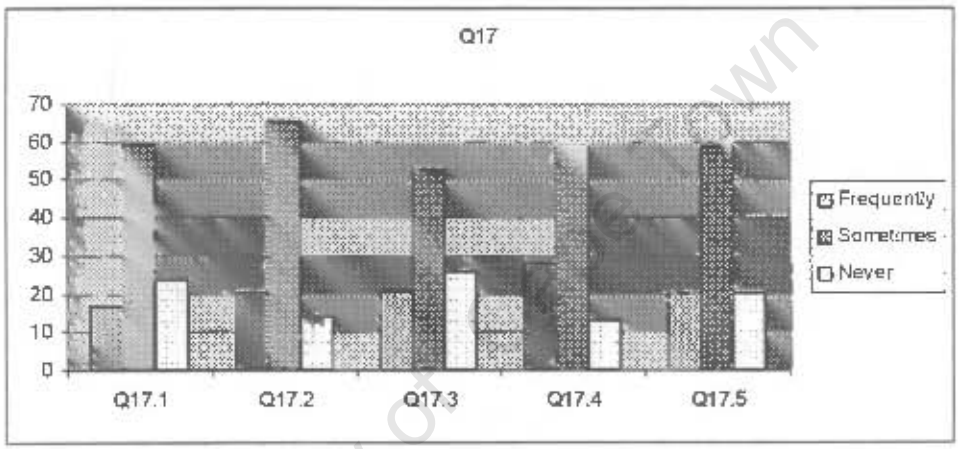
☐ I don't know



University of Canberra

17. Do you ever refer back to previously *sent* email? If so, please rate the frequency of reference for every reason below:

		Never	Sometimes	Frequently
i)	To see with whom I have been communicating about certain topics or issues	☐	☐	☐
ii)	To see what I had to say about certain topics or issues	☐	☐	☐
iii)	To find an answer to a question or problem	☐	☐	☐
iv)	To forward the email on to other people that would have an interest in the information contained in the email	☐	☐	☐
v)	To retrieve an document (possibly stored as an attachment) that might otherwise be lost	☐	☐	☐



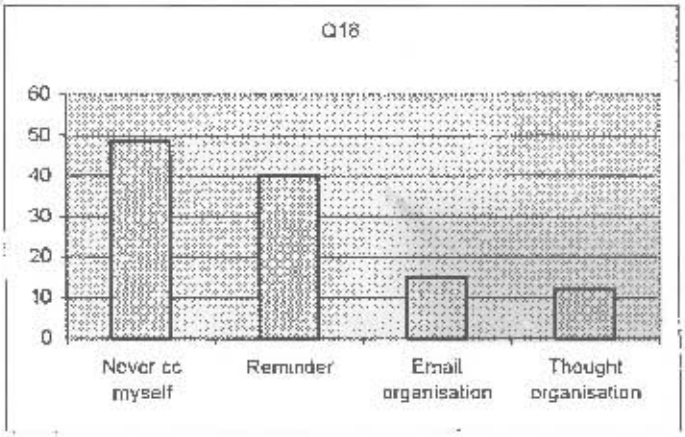
18. Do you ever use email to send messages to yourself?

☐ Yes, to remind me to do things

☐ Yes, as a means of organizing my thoughts

☐ Yes, as a means of organizing email that I send out

☐ No, I never send email to myself

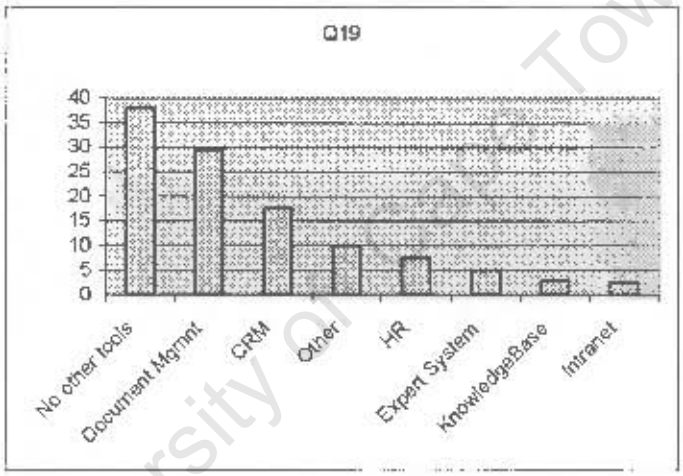


University of Brown

Other information tools used in the workplace

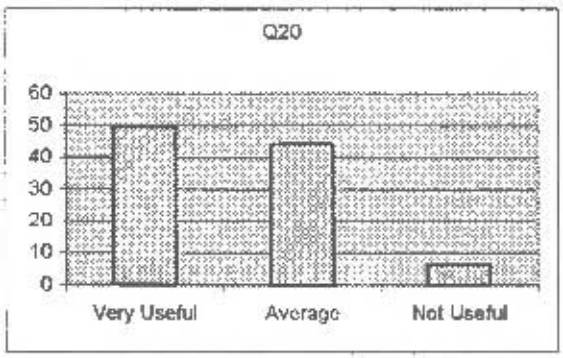
19. Do you use other information storage and retrieval facilities at work? If so, what kind?

- ☐ Customer Relationship Management (CRM) software
- ☐ Document management system (document archiving facility)
- ☐ Expert Systems
- ☐ Human resources (HR) systems
- ☐ Other _____
- ☐ Other _____
- ☐ Other _____
- ☐ I do not use any other tools



20. How would you rate the usefulness of these other tools (in 19 above) for finding information that helps you to solve a problem, answer a question or learn about a new subject?

- ☐ Not useful
- ☐ Average
- ☐ Very useful
- ☐ I do not use any other tools



University of Cape Town

Appendix B

Following are correlation matrices derived from analysis of email log data. Relevant correlations have been highlighted. Marked correlations are significant at $p < .05000$.

Description of the variables:

ATOB – quantity of email sent from A to B

BTOA – quantity of emails sent from B to A

COMMSTR – the communication strength calculated by equation (1) (Section 7.1.2)

TRUST – the trust value applied by every volunteer to other people

USEFUL – the usefulness value applied by every volunteer to other people

PWORK – the percentage of emails estimated by the volunteers to be work related

GRADIENT – the gradient at which communication has increase or decreased over time

Correlation between all communicating pairs

N = 497

	ATOB	BTOA	COMMSTR	PWORK	TRUST	USEFUL
ATOB	1					
BTOA	0.817278372	1				
COMMSTR	0.940328315	0.927275632	1			
PWORK	0.321534706	0.32450066	0.306022651	1		
TRUST	0.129370927	0.182313644	0.169091594	0.109168881	1	
USEFUL	0.000673462	0.015271116	0.012193211	0.130723126	0.038598332	1

Correlation for communication between individuals, limited to the volunteer group

N=23

	TRUST	ATOB	USEFUL	COMMSTR	PWORK
TRUST	1				
ATOB	0.354698101	1			
USEFUL	0.777496624	0.190315983	1		
COMMSTR	0.399593764	0.672785931	0.311839789	1	
PWORK	0.830695437	0.311577969	0.719183703	0.385682654	1

University of Cape Town

Correlations between fluctuations in communications over time and trust and usefulness of people

N= 131

	GRADIENT	TRUST	USEFUL
GRADIENT	1		
TRUST	-0.10075977	1	
USEFUL	0.152102562	0.288481561	1

Correlation between communications and ratings at a group level

N=212

	COMMSTR	TRUST	USEFUL
COMMSTR	1		
TRUST	0.24247524	1	
USEFUL	0.30047462	0.69024847	1

Average of trust and usefulness assigned to every group

Group Name	Average Trust	Average Usefulness
Staff	92.9166667	72.0416667
MSc	85.1428571	66.5714286
4th year students	76.8181818	61.5909091
3rd year students	45.7692308	20.8421053
2nd year students	33	15.6315789

University of Cape Town

University of Cape Town