METOKIS
Methodology and Tools Infrastructure for the Creation of Knowledge Units

DELIVERABLE

Information Society Technologies

<Knowledge Service Methodology>
<METOKIS Deliverable D08>

- edited for online distribution, 13.9.04

Task Number - Name	**8 – Knowledge Service Methodology**
Issuing Date	**July 2004**
Distribution	***edited for on line distribution, 13.9.04***
Document Web Link	**http://blog.monkeymagic.net**
File name	**Piers Young KSM v1.0.doc**
Author/Editor	**P. Young**

Table of Contents

1 Executive Summary

- ***Knowledge worker productivity is a critical issue*** .
  Conservative estimates indicate that just under a tenth of knowledge workers' time is wasted thanks to failure to find documents, having to recreate documents in different formats, and similar tasks. The cost to the organisation of this lost time is thousands of Euros per knowledge worker per annum.

- ***Improving this pr oductivity means optimising the way content is moderated,..***
  whether it is through "IT agents" (such as search engines, web services, or publishing tools), "human agents" (such as work colleagues, experts or journalists) or a blend of both.

- ***The Knowledge Service Methodology provides a framework for specific "solution blueprints" to the problem of optimising this moderating layer*** . With the range of different possible IT-human agent combinations, a prescriptive "one size fits all" *method* is clearly not feasible. Instead, a methodology is needed that outlines the stages needed to be met in any economically feasible solution. Off each stage this methodology, hang libraries of methods some of which may be generic, and some tailored to specific organisational needs.

- ***Optimizing this moderating layer is at heart a context problem.***
  The Knowledge Content Carrier Architecture (KCCA) aims to improve the moderating layer by improving computer to computer communication is a key facet in optimising the moderating layer. But it is impossible to fully optimise the moderating layer without knowing the context within which the actor is acting, and so what content is relevant to him or her, and what tasks are helpful and which choices need to be made.

- ***Context is multi -faceted concept.***
  'Context' is a classic Wittgenstein cluster or family concept with many facets. Many of the key facets have been the subject of separate research over the last 20 years, from the fields of computer science, the social sciences, and the humanities.

  The Knowledge Service Methodology seeks to provide a synthesis that does justice to the individual facets but which also provides a practical basis for factoring a greater contextual awareness into specific IS applications that might use the KCCA.

- ***Complexity Theory provides a model for synthesis of the theoretical facets***
  Though the different theoretical approaches to context come from a broad range of systems, research tends to be underpinned by one of two alternative points of view: positivist or interpretivist.

  Both viewpoints can be successfully merged under a complex systems model. Under this interpretation, the subjective, fluid, fragmented individual perspectives of actors (and human agents) are continually interacting and negotiating. From these local interactions, there emerges an order, such as the vocabularies used by a Community of Practice. Ontologies, taxonomies and models are efforts to model and condense that order. Importantly, those models feedback and affect the local interactions of individual perspectives.

- ***Practical facets such as the developments in Social Software conform to the model***
  Social Software such as weblogs and wikis are beginning to make use of similar insights to the ones afforded by the complexity model. They allow loose, unstructured communication (thus satisfying the interpretivist theoreticians) but they also allow for emergent taxonomies (thus satisfying the objectivists). Intriguingly, the

notion of personal, shareable "task taxonomies" look to be able to follow the same model.

- ***Context is Consensus***
  Using this complex systems model, and some of the insights it affords us by extension, we can determine a means of embedding context into the moderating layer in a form that is tractable to IT agents and Knowledge Content Objects.

  Given the synthesis, the key frame through which we can analyse the actors' needs (and so moderate for them) is the project or decision. Projects (or decisions) are a combination of consensus and action. Once the consensus (however quickly) is achieved over what the decision is about, the actions follow on.

  To optimise the moderating layer, then, context needs to be embedded in such a way that the KCCA, the Semantic Web and IT agents can make use of it. Context comes from consensus, and the most effective and only economically viable way of embedding it is through modelling the emergent consensus among actors involved in a project. The implications for metadata are that the only economically viable and the only tractable metadata is that which is automatically generated (rather than human authored). Moreover, to make this as useful to the actor as possible such the metadata should, when it can, be "scraped" from consensus rather than heuristically determined.

- ***The Knowledge Service Methodology***
  From this a first version of the Knowledge Service Methodology has been outlined. There are six steps

  Step 1: Map system - "Landscape"
  Identify project consensus points. Identify actors. Identify communication channels ("formal" and "emergent"). Perform content & tools audit.

  Step 2: Seed system – "Sow"
  Given the map of the knowledge system, identify those individuals who have the greatest effect on the group's consensus.

  Step 3: Encourage emergence – "Grow"
  Encourage emergence through conversation-based social software such as blogs and IM, and through document-based social software such as wikis and discussion boards. Support human intervention in this process, such as face-to-face networking.

  Step 4: Remove obstacles to emergence – "Prune"
  Isolate unwanted behaviours, of both human and IT agents in the system and remove.

  Step 5: Develop taxonomies – "Harvest"
  Use emerging taxonomies for content, user and task to construct models on which IT agents can act.

  Step 6: Bridge consensus & action – "Plough"
  Feedback models and constraints into the KCCA. Return to Step 2.

- ***Knowledge Service Blueprints***
  The next phase in the METOKIS Project is developing Knowledge Service Blueprints from this Methodology to support the three environments of Clinical Trails, eLearning production and Executive Support.

2 Introduction

This paper aims to define a methodology to develop knowledge service blueprints that help translate the domain tasks (i.e. what has to be done) into appropriately detailed service requirements (i.e. how those tasks are to be carried out). In particular, this Knowledge Service Methodology will help identify the dividing line between those tasks carried out by the knowledge workers under investigation that can be provided in a cost-effective and pragmatic way via technological solutions and those tasks that continue to require substantial human intervention.

The paper begins with an overview of the problem under consideration (Section 2). The major alternative methodological approaches to this problem are then described and evaluated (Sections 3 and 4), with the result that both point towards the most appropriate knowledge service methodological solution being based on a complex systems model (Section 5). This top-down theoretical discussion is then supported by a bottom-up pragmatic discussion of those technological tools that are already being used to support such complex systems in the world today (Section 6). This then leads to conclude on the key requirements of a knowledge service methodology (Section 7), and thus a knowledge service methodology is proposed (Section 8), which can then be tested in different contexts.

2.1 Basic Knowledge Systems

Knowledge management has many different definitions, but at heart it is the study and implementation of IT-supported knowledge systems (referred to as knowledge systems hereafter) that support knowledge workers. The METOKIS project aims to make both systems and workers more efficient and more effective.

Very broadly, a knowledge system can be viewed as a combination of Actors[1], Content, and a Moderating Layer, consisting of IT and human agents. This combination is shown in Figure 1.

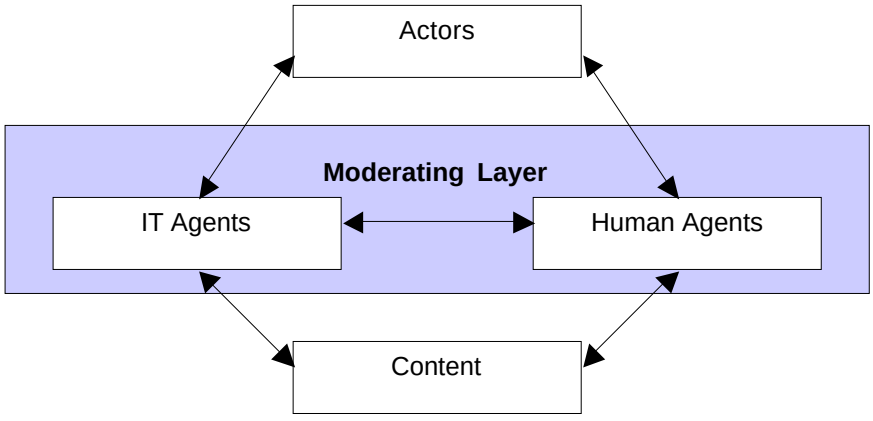

Figure 1: Basic Knowledge System

[1] The term “actor” is used in a number of different endeavours, such as actor network theory or UML, and these different uses bring different nuances to the term. In this instance it is used simply to demarcate the individual who needs access to content to help with a task – that is, to help them “act”.

2.2 Hard problem because of embedded recursion in this model

While this serves to highlight the role of the moderating layer in knowledge systems, the diagram above is over-simplified. "Human agents" are actors in their own right. As a consequence there is an implicit recursion in this simple model. Human agents access content through further (though perhaps differently balanced) moderating layers, and these layers include further human agents (who are actors) actors and further layers.

2.3 Need to optimise moderating layer

The proliferation of content, good and bad, over the last twenty years has created a pressing need to optimise this moderating layer. The need is brought into sharp relief by looking at current knowledge worker productivity.

Research on this productivity, such as IDC's "The High Cost of Not Finding Information" (Feldman 2001), highlight the drawbacks of failing to optimise this layer. IDC found that knowledge workers spend between 15% of their time looking for content, and half of their searches are unsuccessful. So, given that knowledge workers spend six hours a week (15% of their time), that half this time is fruitless, and that the average knowledge worker salary is $80,000, IDC further concluded that a yearly minimum of $6,000 was "wasted" per knowledge worker.

It is a minimum because the figure only includes "pull" – no account is taken of the time wasted through knowledge workers paying attention to content being "pushed" at them unnecessarily (Sieloff 1999). Add to these problems such as needless recreation of content and poor tools, and the waste cost begins to look very significant.

2.4 At heart a context problem

Optimizing this moderating layer of IT and human agents is essentially a context problem. All this costly, time-sapping pushing and pulling of content cannot fully be optimised without knowing the context within which the actor is acting, and so what content is relevant to him or her. But how, if at all, do we helpfully embed context into the system? Whose context do we embed? With what blend of IT and human agents? At what cost to the actor or the actor's organization?

Many, perhaps most, current efforts at understanding context (and how to deal with it) fall into one of two broad camps. The first is a positivist, rationalist approach and is the predominant approach from computer science. Underpinning this view is the belief that there is a fundamental order to the way things work, and that this order can be captured and then used to improve the moderating layer. Life, for this camp, may be complex, but it is essentially "tidy". This approach has enabled us to develop some sophisticated means to embed context into knowledge systems – and some powerful IT Agents as a result.

The second camp takes an interpretivist, constructivist view approach to context, and this is the predominant approach from the humanities and the social sciences. For these interpretivists, people, both individually and collectively, are inherently complex. While a fundamental order may indeed exist, these complexities, when combined with the view that context is human-information interdependent, prohibit any quick, objective understanding of that order. Life, as we can know it, is "messy". As such, the tools the positivists try to develop are likely to be too time-consuming to be viable, or inherently biased, or worse, context-free.

If we are successfully to begin to optimise the moderating layer, then we need to broach a compromise between these two approaches. Given make-up of a moderating layer, a view of

context that is workable for both IT and human agents. Thankfully, lessons from network science and complexity theory provide a basis for such a compromise, and developments in the world of “social tools” indicate that such a compromise is not simply theoretical, but practicable too.

2.5 The Knowledge Service Methodology.

The Knowledge Service Methodology has its basis in this compromised view of context afforded by complexity theory. It aims to provide a series of steps by following which an organization can begin to optimise their moderating layers and their knowledge worker productivity as a result.

Now, given that today’s business and technology environment is complex and quick to change, organizations develop and support all types of knowledge systems. In other words, they use moderating layers built from a broad combination of agents, both human and IT.

As a consequence, a prescriptive “one size fits all” *method* is clearly not a sensible answer to the problem of knowledge worker productivity. Instead, a methodology is needed that outlines the stages needed to be met in any economically feasible solution. Off each stage this methodology, hang libraries of methods. Some of these methods will be generic blueprints, others specialised and specific to the particular constraints, characters and needs of a company or working environment.

This Knowledge Service Methodology, together with the arguments for it, is outlined below.

3 The Positivist View of Context ("Life is Tidy")

3.1 Abstraction

As with mathematics, computer science gains much of its power from abstractions. Functions, predicates, forms of logic and other abstractions enable repeatability, predictability and efficiency of code.

Take for example a basic function (Figure 2): there is an input, something happens, and an output. This "black box" approach means output is determined by input, and input alone.

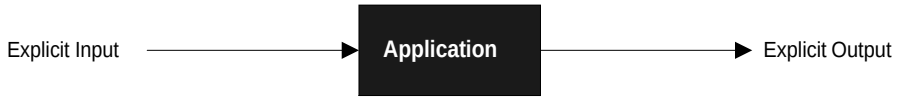

Figure 2: Black-box approach

Importantly, abstraction allows us to simplify our environments. We do not need to be automotive engineers to drive a car from London to Oxford. Equally, we do not need to understand everything going on in a company to be able to effectively manage it.

Unfortunately, though, the cost of these (admittedly significant) benefits is context-independence. While "garbage in, garbage out" makes sense when talking of computer functions and mathematical abstractions, it is less clear that it applies to the human processing of information. The users' situations – their project, their priorities, their place, and their preferences – have not been dealt with very well. Tellingly, users' mental models of a system's capabilities are very often incomplete, sometimes even intentionally so. Norman (1983) indicates that they adopt these sub-optimal models protect them against the possibility or consequences of errors.

3.2 Context-Aware Computing & User Modelling

Now, to say that computer science is unaware of the problem would be grossly unfair. IBM's Lieberman and Selker (2000), for example, have made a strong case for the need for "context-aware computing". As shown in Figure 3, they look to extend the black box model by making context an "implicit input and output to the application. That is, the application can decide what to do, based not only on the explicitly presented input, but also on the context, and its result can affect not only the explicit output, but also the context."

The major approach to make IT agents more context-aware is user modelling. Here, the idea is that the user has a repertoire of tasks and attributes and by accurately modelling both these and the system in which the user is situated the software can then make more context-aware decisions. With its pedigree coming from Expert Systems of the 1970s and 1980s, it is by no means a new approach, but it is becoming more and more sophisticated and more and more powerful. Over the last couple of decades, approaches such as user cognitive modelling (e.g. Dueck, 2001; Wilson et al., 2002) have looked to integrate findings from psychology which indicate that, though different people approach the same tasks in different ways, there are a bounded number of general approaches.

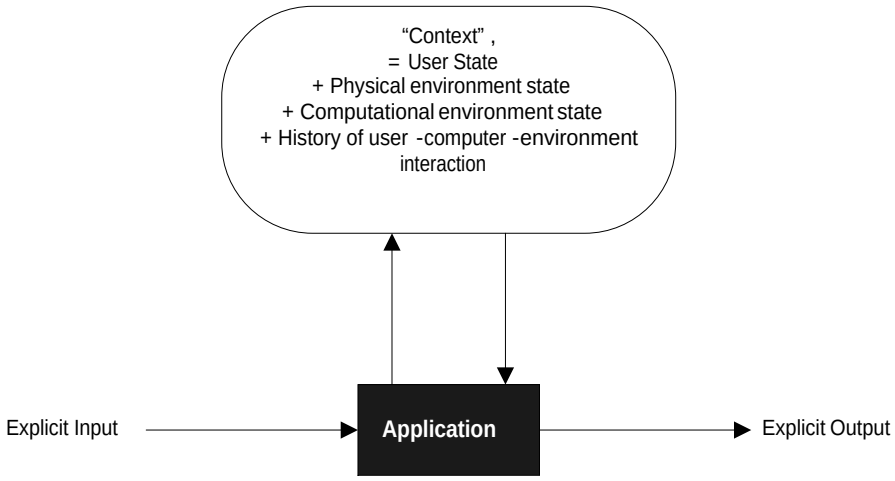

Figure 3: Black Box & Context

3.3 The Semantic Web

A recent boon to modellers is the Semantic Web. This, in theory (and hopefully in practice) allows concepts to be semantically mapped. This mapping is of critical importance to ratcheting up IT Agents ability to deal with concepts and context. As Berners-Lee describes it,

> "the Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation …
>
> The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users. Such an agent coming to the clinic's Web page will know not just that the page has keywords such as "treatment, medicine, physical, therapy" (as might be encoded today) but also that Dr. Hartman *works* at this *clinic* on *Mondays*, *Wednesdays* and *Fridays* and that the script takes a *date range* in *yyyy-mm-dd format* and returns *appointment times*. And it will "know" all this without needing artificial intelligence on the scale of 2001's Hal or Star Wars's C-3PO." (Berners-Lee et al, 2001)

There are four cornerstones that need to be in place if the Semantic Web is to be built. These are: data identification and retrieval; the development of vocabularies; model constraints; and assertion and proofs.

3.4 Metadata

Underpinning this all is metadata. Two types of metadata can be distinguished: human-authored and automatic.

Human-authored metadata is either created "in-process" or "out-process". That is to say, it is either authored at the same time as the data, or afterwards. With both, the metadata can be created by either the originator of the data, or a different individual.

The second way metadata is derived is through automatic inference – through the recognition of patterns and trends in data, and through semantic assumptions concerning the correlations

between data and metadata. This metadata can be created in two ways: heuristically or transparently. With transparently created metadata, humans perform a semantic analysis to create the metadata, and mechanically extracted information is then associated with it. By contrast, in heuristic metadata creation an algorithm performs analysis on the data set and so aims to artificially reproduce intelligent behaviour – human intervention is not needed.

The success stories of metadata are compelling. Google's PageRank algorithm, for instance, infers page importance from the topology of web hyperlinks. It is a hugely successful moderating layer agent. Google's initial public offering raised $1.67 billion in August 2004 and its market value was set at $23 billion, which put it on a par with General Motors. Any value that Google has, though, is only made possible by the fact that semantics of hyperlinks are well determined, thus understandable by machines. Google's value is dependent on an ontology of hyperlinks.

3.5 Ontologies

3.5.1 What is an ontology?

The AI community has various definitions of ontologies, many of which not helpfully contradict one another. What we can say is that:

a) an ontology is a means of ordering, structuring and classifying the concepts, objects and relations within a subject domain, and
b) there are a number of different types of ontology, differentiated horizontally (by their levels of precision) and vertically (by their levels of abstraction).

3.5.1.1 Levels of Precision

The aim of an ontology is to clarify what can and what cannot be inferred from content. This helps distinguish the areas where human and software agents can act on content. An ontology does this by ordering, structuring and classifying the concepts, objects and relations within a subject domain.

Ontologies consist of:

- The general topic you're interested in (the domain)
- A formal description of the concepts used in the domain (classes or concepts)
- Properties of those concepts, features and attributes (roles, slots or properties)
- Restrictions on those properties (facets or role restrictions)

There are various levels of precision in ontologies (Figure 4). At their vaguest, ontologies can be catalogues or glossaries; at their most precise they are fully-fledged axiomatised theories. As these types become more precise, restrictions can be put on concepts and relations defined between them.

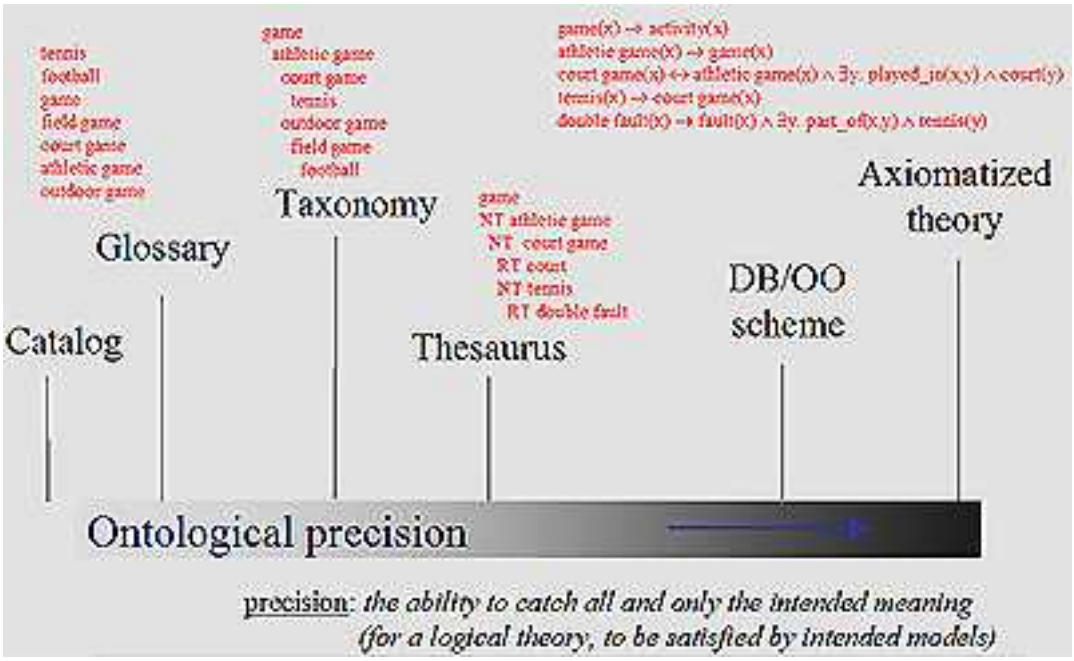

Figure 4: Ontological Precision

Some of the relationships in the domain will be implicit (e.g. a human cannot be a type of fish because a human is a mammal and animals are fish *xor* mammals *xor*…). Other, more functional relationships can be added through, for example, Middleware. In this case, the middleware might work in such a way that dictates if x is a type of a (through checking the ontology), then do x1, else x2.

3.5.1.2 Levels of Abstraction

Ontologies also differ in their levels of abstraction. Put another way there are different ontological layers. This allows the inheritance of concepts and properties.

Each ontology in the diagram below above might be equally precise, but they have the following differences.

- The foundational ontological layer is as abstract as it gets. It is fully domain independent, i.e. it deals with concepts that apply to all domains, such as space and time. As such it is the bedrock of all other ontologies.

- The domain ontological layer is the least abstract. This deals with concepts within a specific domain (molecular biology in the example above)

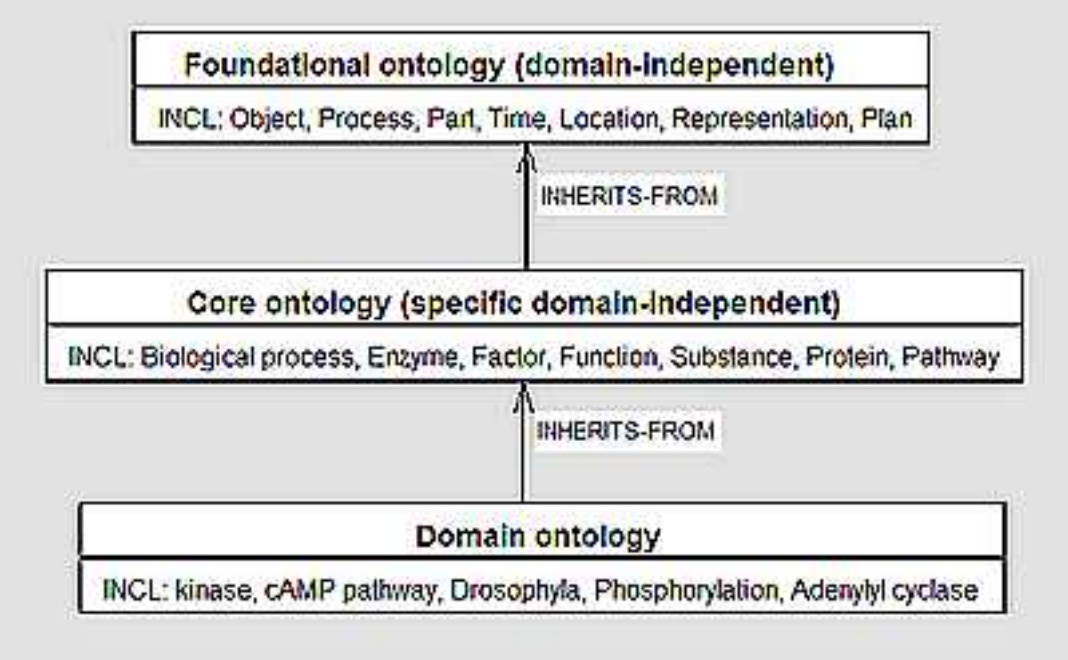

Figure 5: Levels of ontological abstraction

- The core ontological layer is "specific domain independent". While foundational and domain specific layers provide respectively the most and the least abstract representations of a domain, there are "shades of grey". There are times when it is useful or necessary to abstract away from a specific domain ontology and generalise, while not generalizing to the level of concepts such as space and time. So, for instance, in the example above, the core ontology covers aspects of biology. Many of these aspects can be inherited from more specific domain ontologies. In this case the specific domain is molecular biology, but the core "biology ontology" would equally apply to neurobiology, botany and the like.

3.5.2 How does an ontology get used?

By themselves, ontologies do not *do* very much. They just offer a structured logically consistent picture of a domain. There are (at least) five reasons why you might want to develop them.

a) To share common understanding of the structure of information among IT agents.
If a number of web sites deal with a similar subject (e.g. European Funding), they might use different terms for the same concepts (e.g. money, funding, donation, fool's gold). This makes it hard for software agents (and uninformed human agents) to compare those sites. If the sites share and publish the same underlying ontology of terms used, then computer agents can extract and aggregate information from each. The agents can use this aggregated information either to answer more effectively user queries or as input data to other applications.

b) To enable reuse of domain knowledge and so save time, money and headaches.
So for example, if one ontology represents engines, cars and automotive engineering for Formula 1 racing then this could be reused in developing a more generic "car" ontology, which could then be further used in a more generic transport ontology, and so on.

c) To make domain assumptions explicit.
This is useful for two reasons. First, if domain knowledge changes, then this explicitness means you have an easily accessible one-stop point of change (as opposed to searching for

assumptions that are embedded in code). Second, it useful for new users learning about a domain

d) To separate domain knowledge from the operational knowledge
(where domain knowledge is generalised knowledge - such as how to support executives or clinicians – and operational knowledge is tool- or ingredient-specific.)

e) To analyse domain knowledge
This is mainly valuable for logicians when trying to sort out when you can reuse an ontology.

As mentioned at the beginning, an ontology does not actually do anything on its own. Ontologies are reference structures, typically one document, that need to be combined with middleware or a web service. Normally, ontologies are used by being exported to one of various XML-based representations. Examples of these representations are: XML (DTD); XML (Schema); RDFS (RDF Vocabulary Description Language); DAML+OIL/OWL; and ISO Topic Maps

3.5.3 Why not just use XML/RDF etc?

For IT agents to conduct automated reasoning of the type hoped for in the Semantic Web, they need to have access to structured collections of information and sets of inference rules.
XML lets everyone create their own tags—hidden labels such as those that annotate Web pages or sections of text on a page. Scripts, or programs, can make use of these tags in sophisticated ways, but the script writer has to know what the page writer uses each tag for. In short, XML allows users to add arbitrary structure to their documents but says nothing about what the structures mean.

Meaning *is* expressed by RDF, which encodes it in sets of triples, each triple being rather like the subject, verb and object of an elementary sentence. These triples can be written using XML tags. In RDF, a document makes assertions that particular things (people, Web pages or whatever) have properties (such as "is a sister of," "is the author of") with certain values (another person, another Web page).

Unfortunately, that is not the end of the story. Two databases may use different identifiers for what is in fact the same concept, such as *zip code* and postcode. A program that wants to compare or combine information across the two databases has to know that these two terms are being used to mean the same thing. Ideally, the program must have a way to discover such common meanings for whatever databases it encounters.

And that is where ontologies come in.

3.6 Economics of ontologies

There are concerns about the economic feasibility of endeavours such as the Semantic Web. While metadata is a crucial step towards IT Agents becoming more context-aware and more able to “reason” about the content, they need to be supported by logical frameworks such as ontologies. Ontologies, though, are hard work. They take time, they need expertise and they cost money.

Stefano Mazzocchi (2002), the founder of the Apache Cocoon project has raised a number of issues with the development of the Semantic Web. Rather than fundamental concerns with the positivist approach, these are concerns with the economic viability of the Semantic Web, including the Babel Problem, the ROI Problem, the Screen-Scrape Problem, and the Marginal Costs Problem.

3.6.1 The Babel Problem

The Babel Problem is that XML's flexibility means that new mark-up languages can be written on spec to fit each and every need. Given that many existing markups are too complex to learn quickly, more and more markup languages are being written. Nor is it clear that this boom will level off of its own accord. It is by no means obvious that this trend will come to a saturation point (especially with the advent of SOAP-based web services[2]). Customization in and of itself is no bad thing, but automatic translation between markups is not always algorithmically possible. This means that there is a likelihood that experts will have to make the mapping for you (and they have a cost associated with their work), and that as that likelihood will increase with the explosion in markup languages, the costs of translation between markup languages will rise too. In many ways, the problem is a computer-to-computer variant of the "terministic screen" problem outlined above. Berners-Lee et al., with a nod to the interpretive tradition, state that, to deal with the range of individual preferences for and differences in dealing with information.

> "Semantic Web researchers … accept that paradoxes and unanswerable questions are a price that must be paid to achieve versatility. We make the language for the rules as expressive as needed to allow the Web to reason as widely as desired" (Berners-Lee et al., 2001)

This is all well and good. Interpretivists would argue that it is crucial for any Semantic Web undertaking to acknowledge these complexities. Nevertheless, the Semantic Web Researchers look to have ignored the wooden beam in their own eye. Just as much as human complexities, it is the range of software preferences for dealing with information that has the capability to undermine the Semantic Web by making it unfeasible.

3.6.2 The Return on Investment (ROI) Problem

The ROI Problem is that, currently, writing 'semantised' information is expensive, certainly more so than writing 'non-semantised' information. The benefits to semantic markup need to outweigh the costs of producing it or the dream of the Semantic Web is unlikely to become a reality.

3.6.3 The Marginal Cost Problem

Linked to ROI problem is the issue of marginal cost. If the cost of adding the relevant semantic metadata to text have a linear relationship with the size of the text, there is further cause for doubting the Semantic Web will scale economically.

3.6.4 The Screen-Scrape Problem

Lastly, Mazzocchi identifies the Screen-Scrape Problem. HTML is still the most popular standard for publishing on the web, which is semantically "weak". If the only means to semantically enable that content is to "screen scrape", then again costs will exceed the benefits.

3.6.5 Solutions to Economic Infeasibility and Metadata Constraints

Mazzocchi has suggested a number of ways in which to address these problems. First and foremost is XML-based web publishing. With time, this removes the Screen-scrape problem, and it has the potential to show some immediate benefits to the user, thus partially addressing the ROI and marginal cost problems. Second, and if used in conjunction with XML-based web publishing solutions, easy to use, semantic-aware content editors may remove these

[2] SOAP (Simple Object Access Protocol) is an effort to allow different applications running on different servers to communicate using the Internet. It uses HTTP as its transport and XML as its payload, i.e. the breadth of different configurations are dealt with through different XML translations.

problems altogether. Unfortunately, simple, cost-effective semantically aware content editors have not yet been built. Third, and as already implemented by the Apache Cocoon project, it is possible to ask via the HTTP/1.1 protocol for variants of a resource. A standardization of “semantic variants” together with XML-based publishing systems could remove the Screen-scrape problem, and would certainly address it more quickly than XML-based web publishing alone.

These problems, and in the current environment the Babel problem, look likely to make the Semantic Web economically infeasible. At best, they impose significant constraints on the type of metadata that we can realistically expect to create, edit and use. As already noted, there are two types of metadata: human-authored and automatic. Systems that feel intelligent, though, need human-authored metadata and this, as can be seen from the above concerns, is prohibitively expensive. The Babel problem indicates that the more metadata humans author, the more metadata humans have to author, and the costs spiral accordingly. This leaves us with automatically inferred metadata. This type of metadata comes in two flavours: heuristic (all computers) and transparent (human structured but mechanically filled). Accordingly, for the solutions to feel human, it makes most sense to aim first for transparent or “scrapeable” metadata, with heuristic metadata as a last resort. Indeed it is this sort of metadata on which Google’s PageRank algorithm and Amazon’s “Customer’s who also bought” feature relies. They are indeed compelling success stories of the value of metadata, but not any metadata: automatically created, economically viable metadata.

4 The Interpretivist View of Context ("Life is Messy")

Members of the interpretivist thinkers come from a range of different disciplines. And that is natural enough given the positivist thinkers' definition of context. Context, for them, is "everything that affects the computation except the explicit input and output." Put another way, context is everything that is implicit in the environment that affects the computation. Their studies have considered the impact of context on human perceptual and cognitive processes in various sociological and psychological research. Weick (1995) on sense making, Wolfgang Welsch (1997) on aesthetics, Haugaard (1997) on power and knowledge, Sillince (1992) on new rhetoric have all stressed the strong interdependence between context and human information processing. It is precisely this interdependence that any attempt to extend the basic IS black box model must strive to capture.

It is difficult to do justice to such a broad range of thinking, but there do appear to be a number of common complaints with the positivist approach. These are that:

- humans are complex (not simple),
- every facet of the user-task-system model is open to bias (not objective)
- the environment is dynamic (not static)
- the conduit metaphor of information transfer is inappropriate.

As will be seen, there is substantial overlap between these complaints, but it is worth briefly looking at each of these in turn.

4.1 Human complexity

A key assumption of the positivist thinkers is that at some level people can be generalised and sensible conclusions can be drawn from those generalizations. They have common features, those in specific roles do specific tasks, by definition, and those, and that if two users share the same "description" what is relevant to one is relevant to another. But the interpretivist thinkers continually highlight the danger of the approach. Humans are more complex than that. A common metaphor is that of the iceberg. While it may be possible to model to our visible tips, we each have much larger, shifting invisible area below water level. And it is this "underwater" level that has the potential to sink even the most Titanic of tools.

Specifically, the complaints are that: identities are fragmented and fluid, and while people may have the same number of glasses through which they view the world, those glasses have a myriad number of tints.

The notion that we have single identity has been under attack since the beginning of the last century. The thinking is that, rather than being one unified entity, we are what George Herbert Mead (1968) called "a parliament of selves". Moreover, that parliament is continually shifting and adjusting to the world outside. Cooley suggested the metaphor of a mirror; we continually adjust our "selves" depending on how we see them reflected back in our external lives. (Cooley, 1992) This fragmentation and this fluidity are a key part of the "underwater" part of the user which modelling fails to capture. As Weick puts it,

> "Personal identities are shifting and multiple. And when those shifting identities are embodied members of the top-management team (a frequent target for accuracy studies), outside observers who try to predict the behaviour of this team using the model of object perception are in trouble … Questions in organizations that look like they involve global accuracy and object perception

> tend to be translated into questions of intentions and personalities." (Weick, 1995)

The second complaint is that users view the world through a number different "screens". Kenneth Burke (1945), one of the grandfathers of New Rhetoric, drew attention to the notion of "terministic screens". These screens are a set of symbols that become a screen through which we make sense of the world. In Burke's philosophy, social interaction and communication can be understood in terms of a *pentad*, which includes act, scene, agent, agency, and purpose. On a similar track, Kelly developed his Role Repertory Grid (Kelly, 1995; Kelly, 1970). Both of these approaches have had practical applications. Critically, many of the results indicated that users themselves might not be aware of many of the different terministic screens involved in separate communities.

As Sillince (1992) has reported in his work on classification

> "1. If two groups of people construct thesauri in a particular subject area, the overlap of index terms will only be 60%.
>
> 2. Two indexers using the same thesaurus on the same document use common index terms in only 30% of cases.
>
> 3. The output from two experienced database searchers has only 40% overlap.
>
> 4. Experts' judgements of relevance concur in only 60% of cases."

This has serious implications for the positivist thinkers.

The argument from the interpretivist thinkers is this. User modelling needs to be aware of the complexity of the individual. Often it is not, which undermines it. User modelling can be done – studies have shown that - but it is not a one-stop shop. Models need to be done in a real-time fashion. Users are not static, nor are their understandings of their environment, and although they may not change quickly, models need to be able to change with them. Lastly, the modeller or modellers are users, with their own "terministic screens" and "personal constructs". To assume these screens are objective and "one size fits all" is to make a mistake. Even then, if observer bias is removed, the complexities of identity necessitate modelling each one individually.

4.2 Model bias

Returning to a positivistic user model, there are three key elements: user, task, and system. Inherent in the modelling process is the notion of objectivity. However, linked to the notion that humans are complex is the notion of model bias and interpretivists might argue that this is perhaps a deeper issue than Kelly's observer bias.

There seems, in effect, to be a triangle of model bias. As shown in Figure 4, the user can bias the system, and the system the user; the user can bias the task, and the task the system, and the task can bias the system and the system the task.

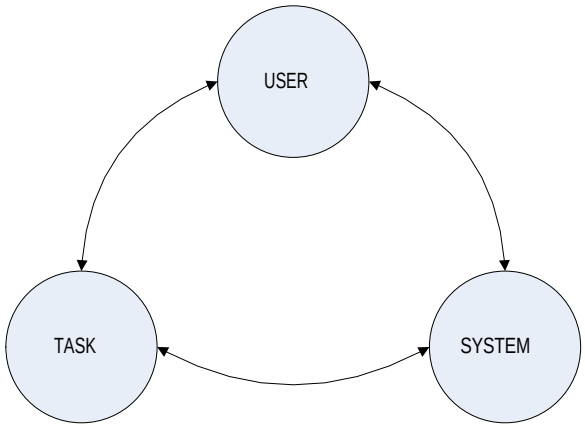

Figure 6: Triangle of Model Bias

First, there is the actor-system facet. This, like the other facets, is prone to two-way bias. It has already been seen that, because of the complexity of the individual, each user is likely to see the same system through different lenses. Modelling, even with observer bias removed through tools such as the Role Rep Grid, is prone to bias unless a representative sample can be found. If there is to be any context embedded in user models then that context is at best system-specific.

However, while individual lenses may bias the system model, the system also biases the user model. In effect, interpretations can have no grounding outside the rhetorical exchanges taking place within institutional and cultural politics. Hermeneuticians such as Heidegger, Habermas and Ricoeur all argue for the constraining nature of the system on the user. Studies of socialization indicate similar system-user bias. Shutz's (1964) analysis of the stranger, for example, suggests that newcomers need to learn both how to interpret and express themselves in the natives' vernacular.

Embedded in this socialization theme are talk and conversation and the notion of "shared meanings". Although Weick (1995) argues that this may be too general a catch-all (there are for example distributed meanings, overlapping views of ambiguous events, non-disclosure intimacy, equivalent meanings), the underlying worry is the same. Pushing for an objectivist model is fundamentally difficult because meanings, contexts and knowledge are negotiated, not given.

Second, there is the user-task facet. Users, and their positions in a system bias their views of task priority, task importance and task relevance. Lanir, Fischoff and Johnson (1988) have argued that military command and control structures connect people at the top to people at the bottom. These connecting structures are necessary because the mindsets of those at different levels of the organization vary significantly. Broadly, those in the higher ranks focus on strategic thinking, while those at the bottom are concerned with tactical and local manoeuvres. Clearly, projects at the top and bottom differ dramatically, but more important for any study of context, so do readings of the same events.

Interestingly, Follet (1924) has suggested that this type of bias means that, rather than trying to model in advance what needs to be done and by whom, and then persuade users to fall in to line with the plan, instead one should "confront the activity environment" and the complexities within it.

> "Thus we need not make anticipatory judgement; there may be opposition, there may be resistance, but this definition leaves it possible for us to wait until we find them … This confronting would make apparent many incompatibilities of interests, but does not judge the case beforehand as to what shall be done about it. Confront does not mean combat. In other words, it leaves the possibility of integrating as the method of meeting the difference."

Third there is the system-task facet. Even if there is enough stability in an environment for it to be modelled, then the model at least has to take into account the notion the system biases the perception of tasks. Different organizations, doing essentially the same tasks, have very different styles of achieving them. New Rhetoric thinkers such as Olbrechts-Tytecta and Perelman (1969) have argued that different organisations have very different modes of reasoning about and arguing for different approaches. Again, the classifications given by the system determine what tasks are permissible and what tasks aren't. A recent example of this was the Bush Administration's declaration of "war on terror". (Newman 2004) This reconceptualisation and the change of metaphor to war enables, and up to a point demands a different view of what is permissible and what is not. Newman points to some of the steps taken since September 11 that directly challenge long-standing protections guaranteed by the U.S. Constitution and federal legislation

> "Law enforcement officials now have broader power to use wire tapping and Internet surveillance techniques to gather information about citizens.
>
> Noncitizens can now be detained on minor visa violations.
>
> The FBI can demand access to sensitive business records without demonstration that a crime has been committed.
>
> Law enforcement officials can monitor communications between lawyers and their clients.
>
> The U.S. attorney general can order a secret search of a U.S. citizen's home and, based on the information discovered, secretly declare him or her an "enemy combatant." Enemy combatants can be investigated, jailed indefinitely, interrogated without legal representation, tried, punished, or deported by a military tribunal.
>
> The Pentagon now has authority over the U.S. Northern Command, a domestic unit responsible not only for homeland defense but also for providing military assistance to civil law enforcement authorities."

Different system models enable very different tasks.

4.3 Fluidity of environment

The interpretivist thinkers also take issue with the static nature of models. Key to modelling is the notion that the world is static enough to be modelled, but the interpretivist argues that the opposite is the case. Just as the individuals in it are complex and ever changing, so too is their environment. This is grist to the hermeneutic mill.

Dilthey argued that the ever fluid, ever moving characteristics of the environment mean that we can only ever make provisional assumptions (Makkreel, 1992). The fluidity of the world around us means that there are no absolute starting points – and consequently no static certainties - on which we can build because we always find ourselves in the middle of changing, complex situations. The best we can do is to try to iterate towards understanding by making, then revising, provisional assumption.

On the same track, there is Heidegger's concept of thrownness. Heidegger (1978) argues that we have our being through taking up a place in a world that is already 'set up' in advance. We inherit, for example, the great metaphysical divisions between subjects and objects, selves and others, bodies and minds, and cannot do otherwise than to grasp what we are in terms of these prior ways of thinking. That which seems most 'ready-to-hand' in our thought (the pronominal 'I', for example) is not our property, or rather it is not proper-to we as individuals. It belongs elsewhere. We are then 'thrown' into a world marked by the activities and language of those who come before (e.g. Stenner 1998; Munro, 1997). As such, there is

a historical dimension to the way we find ourselves in discourse, as was recognised by much of the initial work in the area (e.g. Shotter, 1975; Harré, 1993; Gergen, 1994).

How then are we to make sense of this fluid, inherited world? Weick (1995) has argued that projects, by allowing us to bind this flow of events in a time frame, provide a critical starting point. Without delving into this too deeply,

> "Our interactions with other people and with the inanimate world we inhabit puts us into a situation of thrownness, for which the metaphor of the meeting is much more apt than the metaphor of the detached scientist." (Winograd & Flores, 1995)

Modelling assumes that one can be a detached scientist, but if this is the case, then it is the case only in part.

> "If people are in the middle, what are they in the middle of? One answer … is 'projects'. And if people are in the middle of projects, then what they see in the world are those aspects that bear on their projects. In other words, even though people are immersed in flows, they are seldom indifferent to what passes them by." (Weick 1995)

The only sensible framework for modelling the environment is one constrained by time. The most obvious such frame, especially in a business context, is the project.

4.4 Conduit Metaphor

This conduit metaphor underlies a range of different approaches to knowledge and knowledge transfer. Essentially, it assumes that communication is a source putting ideas into words and sending the words to the receiver, who then receives the ideas. The underlying metaphor is of putting objects into a container and sending them through some sort of conduit to the receiver who receives the containers and takes the objects out. It is, understandably, a popular model when talking about computer-mediated communication and knowledge management.

> "Most of the traditional Knowledge Management Systems rely on the assumption that knowledge can be assimilated to objects that can be identified, separated from their initial context, and handled in information systems." (Nabeth et al., 2002)

Indeed, the "knowledge content objects" in METOKIS are close in spirit to this metaphor.

But the conduit metaphor sparks off a wide range of complaints from interpretivists. How do the 'objects' get into the 'containers'? In other words, how do we succeed in putting meanings 'into' words and how does somebody else succeed in taking the meanings 'out of' words? Transmission models do not deal with meaning. They were never meant to, admittedly, but there application to communication between humans, whether via IT agents or not, is misleading.

Again, communication differs very greatly according to the context in which it takes place, and the audience (or lack of audience) involved. Communication is also political. Our intentions fundamentally affect the exchange of information. Lastly, the medium and channel for communication affects the message. Conduit metaphors pay no attention to the effect of the channel used. If I want to communicate with you, do I get someone else to pass on my message, 'phone you, send you a letter, send you a memo or seek you out to talk to you directly? The choice makes a difference. And there are also differences in the features of different media that make them more or less appropriate for saying what we want to say.

Some people write songs, some people write poems, some people say it with flowers – none of these seem particularly appropriate when, as above, declaring a war on terror.

The issue here is that much of the context we need as actors in a knowledge system comes from presentation. If the system we use for communication is unable to deal with that presentation layer, then optimising the moderating layer will be extremely difficult.

4.5 Cartesian Anxiety & Ontological Oscillation

Now, the reaction of many in the positivist thinkers to the sorts of arguments laid out above has been labelled “Cartesian Anxiety”. This

> “is best put as a dilemma: either we have a fixed and stable foundation for knowledge, a point where knowledge starts, is grounded, and rests or we cannot escape some sort of darkness, chaos, and confusion. Either there is an absolute ground or foundation or everything falls apart.” (Varela, Thompson & Rosch, 1991)

If we are to make useful tools, they argue, there needs to be a fixed and stable foundation for knowledge.

A second argument is that the interpretivist thinkers are guilty of ontological oscillation (Burrell & Morgan, 1979). That is to say, it is all very well to stress a highly subjectivist stance, even one that denies existence of social structures, but all too often attempts to operationalise these ideas try to smuggle in a more rationalist ontology through the back door.

Perhaps more practically, the concern is that if one cannot label content in sufficiently “sensible” ways, and if one cannot label them in ways that IT agents can make use of, then the goal of this paper – to understand how best to optimise the moderating layer in a knowledge system – is doomed. If each actor really does see the world in different ways, then positivists would suggest that there is no way to know *in general* how best to optimise the layer.

5 A Non-Cartesian Model for Context

5.1 Need to blend both approaches

There is clearly a need to consider both the positivist and the interpretivist thinkers. Neither camp on its own provides an approach that could convincingly optimise the moderating layer.

The argument from positivists is that if we cannot at some level tidy up the way we deal with content, by labelling it sensibly, by giving it the right metadata, then IT agents will necessarily communicate in a sub-optimal fashion. Moreover, since the majority of content is becoming digital, there is an ever greater need to forge a solution that uses digital tools, and that entails abstractions at some level.

Interpretivists, although they differ in the names they give to their arguments, share a common perspective. This is that the top "ontological layer" (the black box) represents the visible tip of a much larger, more complex 'iceberg'. The true impact of context can only be properly understood by reference to the larger, submerged structure underpinning the ontological layer. Current positivistic approaches to this either fail to take account of the depth of the iceberg, or are in danger of becoming economically infeasible. And yet, interpretivists have few alternatives.

This tug of war is driven by interpretivists' reactions to the limits, as they see them, of the Cartesian, positivist perspective of earlier work in their fields. Interestingly, by drawing on contemporaneous developments that reflect a similar reaction within systems thinking, it is possible to map all the different studies, from both sides, onto a single complex systems model.

5.2 Complexity Theory and Emergence

Complexity theory is concerned with systems comprising networks of interacting agents. These networks can be grey networks of neurons, ecosystems, or economies. It is a unifying theory that claims that much of the apparent complex behaviour can be explained by a few simple rules. Research has indicated that these simple rules can explain multinational corporations, mass extinction, rainforests and human consciousness. As one might expect with any theory this grand, there is a lot of enthusiasm.

Armed with the mathematics of Descartes, Newton and Leibniz, science has successfully modelled many of the workings of the universe. The world was seen to run like a complex clock, characterised by repetition and predictability. Putting men on the moon depends on that predictability. Once the trajectory of the spacecraft is known, the path and destination can be predicted using the laws of motion. The assumption is that the world is linear, and much of what we do, whether rocket scientists or not, depends on the order and predictability that comes from this assumption. These assumptions marry well with the positivist approach outlined above.

Most of nature is non-linear, though, and not easily predicted. Complexity theory focuses on these non-linear systems. The plight of the weatherman is the traditional example. Many components interacting in complex ways lead to notorious unpredictability. Ecosystems, developing embryos, economies and the human brain all defy mathematical analysis or simulation. Classical physics has regarded these non-linear systems as complex systems that, when powerful enough analytical tools were available, would be understood through complex descriptions. The central discovery of chaos theory is that, in this case, classical physics is wrong. While non-linear systems may indeed appear complex on the surface, they may be governed by a relatively simple set of subprocesses. Broadly speaking, complexity theory has gone one step further than chaos theory in that it studies not just the unpredictable results attained from the interaction of these simple subprocesses, but also the way in which

these results can then feed back to and affect the underlying subprocesses. Chris Langton describes this through a diagram like the one below. (Lewin 1993)

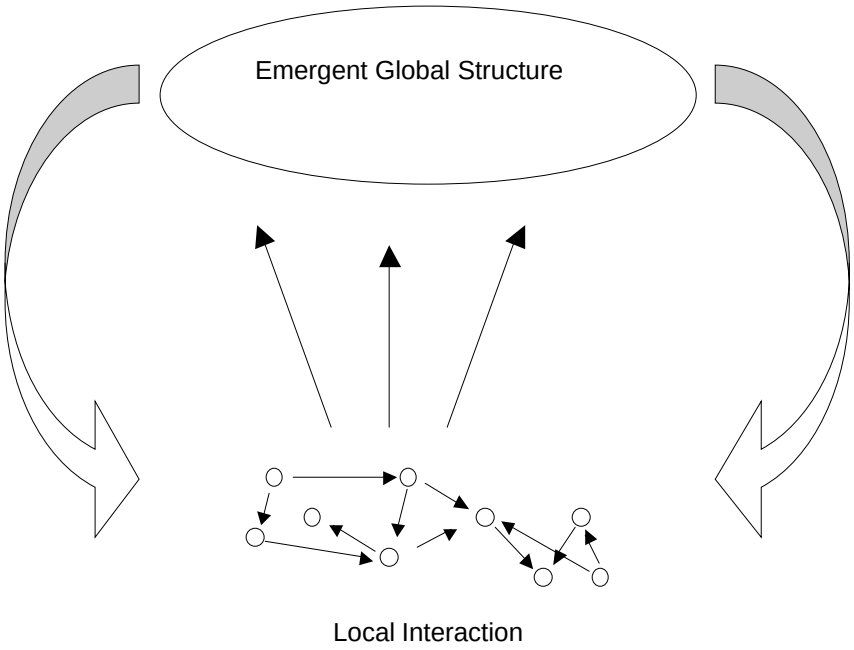

Figure 7: Complex System Model

From the interaction of the individual components, some kind of global system property emerges, which could not have been predicted from what is known of the component parts. And the global property, or rather its emergent behaviour feeds back to influence the behaviour of the individuals that produced it. So for an ecosystem, the interaction of species within it might confer a degree of stability on the system to say the ravages of droughts, floods and other such traumas. Here stability would be the emergent property. For industrial societies, the aggregate behaviour of companies, financial markets and consumers is the capitalist economy. The behaviour that Adam Smith saw in these economies, that they acted as if guided by an invisible hand is a property of all non-linear dynamic systems. It is what Stuart Kauffman has termed “order for free” (1996).

5.3 A Model For Compromise

Many of the concerns of both the positivist and the interpretive camps can be addressed with such a model. Effectively, a highly complex and diverse surface layer creates a fluid and dynamic emergent deep structure. The more stable elements of this structure then support a further level comprising an ontological snapshot of this emergent structure.

In the language of the interpretivist thinkers, personal perspectives, whether they are terministic screens, personal constructs, or fragmented and fluid identities are the engines driving the subjectivist approach. These can be mapped on to the “small pieces loosely joined” that interact locally at the bottom of the complex model. From these and from events affecting them emerges a common approach or language or set of behaviours. This negotiated set of meanings then feeds back to and affects the interactions of at the local level.

The compromise between the two camps comes from the more stable emergent structures as a basis for the ontologies, user models and the tools needed to build more context-aware IT Agents. These are “harvested” from that emergent order. If there is no negotiated meaning, or the negotiated meaning is too transient (when the system is in too rapid a state of flux) then such endeavours are fruitless. But if there is an emergent order, then this harvesting can be

used to help improve the general interaction at the bottom of the diagram. And one of the benefits reaped by improving that interaction might be improving the moderation of content.

Figure 6 shows this mapped onto the complexity model.

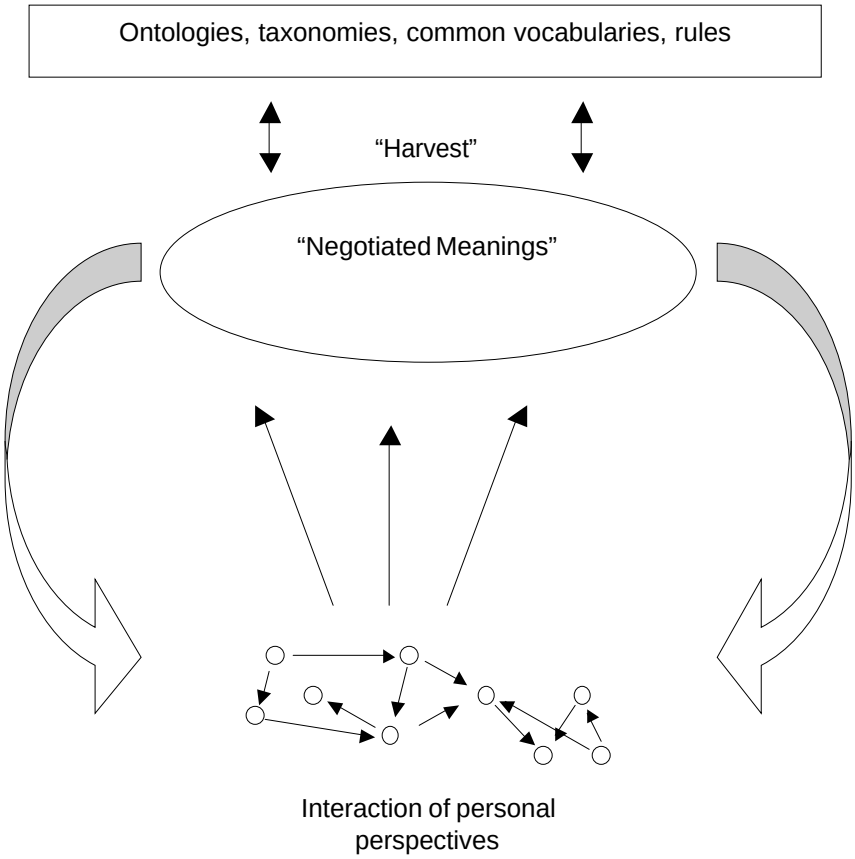

Figure 8: Model Underlying the Knowledge Service Methodology

5.4 Concerns with the Model

5.4.1 The plural of anecdote is not empirical data

There remains the still broader question as to how much credence should be given to systems theory as a base model for optimising the modelling layer. There are two reasons for querying its value. One is that the history of similar theories would seem to indicate a degree of scepticism is called for. The other, which follows on from the first, is that the enthusiasm for the theory needs to be grounded in fact.

Sociometry, while offering a new mathematical representation of social systems, is now viewed by all but a small clique within the intellectual community as a method (i.e. network analysis) rather than a paradigm, thus dashing the messianic claims by its founder Moreno. Similar trajectories can be traced for von Neumann and Morgenstern's introduction of Game Theory, Thom's Catastrophe Theory, and Zadeh's Fuzzy Sets Theory.

A pattern (Back, 1997) can be seen in the fate of these new techniques following their introduction:

the application of this procedure to unresolved questions in social science;

a sceptical reception in the established science with claims of faddism and similar reactions;

the formation of a following for this mathematics;

the claim for a new science resolving many previous problems;

the diffusion into popular – non-scientific culture;

reaction and disillusionment;

the adaptation of the technique to the established model.

The recent surge of interest in self-organizing systems perspective is precariously poised at stage 5, threatening to lurch on to stage 6 on this trajectory. The history of, and reaction to, earlier mathematical theories described above suggest that disillusionment sets in when the public tires of the metaphor and the research community fails to see formalised intellectual advances.

The non-scientific culture of stage five includes management and leadership theory. The excitement that systems theory has brought is palpable. M.J.Wheatley (1992) remarks that "ever since my imagination was captured by the phrase 'strange attractor', I have wondered if we could identify such a force in organizations … my current belief is that we *do* have such attractors at work in organizations and one of the most potent … is meaning." In a similar vein, Stacey (1996) argues that organizations are complex adaptive systems, and shows that he knows his Prigogine by further commenting that they have dissipative structures. Perhaps as telling as any of this is the title of Warneke's work "The Fractal Company" (1993).

Lilienfield (1978), though, has been especially unimpressed by the findings of such systems thinkers.

> "Systems thinkers exhibit a fascination for definitions, conceptualizations, and programmatic statements of a vaguely benevolent, vaguely moralizing nature … They collect analogies between the phenomena of one field and those of another … the description of which seems to offer them an esthetic [*sic*] delight that is its own justification … no evidence that systems theory has been used to achieve the solution of any substantive problem in any field whatsoever has appeared."

The argument against such systems thinking, then, is this: the plural of anecdote is not empirical data.

However, it is not helpful to construe the management theorists' work as literally meant. Rather, these findings of systems theory are a useful and productive metaphor. As with any metaphor, its value lies in its ability to allow us to attack a problem in a different way. And it is through using the complex systems metaphor that the Knowledge Service Methodology hopes to attack the problem of optimising the moderating layer, and improving the context-awareness of the system, in a different, more fruitful way.

5.4.2 Not all networks are the same

Now, although networks are pervasive, they do not all operate with the same constraints. Social networks have three major differences to networks of routers, roads, and other physical entities. And it is worth quickly looking at these as qualifiers to the network model before seeing where social network analysis can help us optimise the moderating layer.

The first characteristic that differentiates social networks from others is that what flows through the network changes as it goes. Whether we call it information or knowledge or content, as it flows from node to node (person to person, or group to group) it is subject to change. As already indicated, this is because people add their own spin, their own contexts,

and their own interpretations, good or bad. The classic example of this is the game of Chinese Whispers. One of the more famous Chinese Whispers is supposed to have happened during the First World War. The original message from the trenches to British HQ was 'Send reinforcements, we're going to advance'. In transit, it became 'Send three and fourpence, we're going to a dance.' Or, alternatively,

> "the reason why the Kaiser's last big push of 1918 almost succeeded was due to the fact that a message to British HQ became garbled in transit. The original message from Col. Henry Wilmimgton-Bottomley read "send reinforcement, Germans advancing on west flank" It was received at British HQ as "send three and fourpence, Gerda's dancing on wet plank. As Gerda was a popular German stripper of the time the three and fourpence was quickly raised after a whip round in the trenches but once the troops found out she was not dancing they became so annoyed they proceeded to kick shit out of the German troops." (Brains Trust, 2004)

Neither the jingoism nor the truth of the account is the issue here. More importantly, both the misunderstanding and the fact that there are two versions of the same misunderstanding indicate that the information passed round social networks can and does change as it passes each node. More generally, Professor Kim Kirsner (2000), a psychologist researching the cognitive processes that are at work when we interpret events, suggests

> "any event - accident, crime or political incident - will generate a huge variety of versions of what actually occurred from those who witnessed or participated in it. And as time goes on, even further distortion occurs in accounts of the event, either deliberately or by accident"

A second key differentiator is that social networks are "political". Just as nodes can affect the information that flows through them, that information can affect the nodes it passes through. Much of this can be inferred from the above discussion of personal constructs, fluid identities and context. Backing this up, evolutionary psychologists (McAndrew & Milenkovic, 2002) have found that

> "people actively seek information about others that will be most useful in social competition. We seek exploitable, damaging information about high-status people and non-allies; we actively disperse status-enhancing information about our allies; and we keep a very watchful eye on our friends."

Though the researchers are cautious about their findings, it would follow that the information that passes through nodes in social networks can affect the status of those nodes, and so their "identities". Less prosaically, knowledge is power.

And the third differentiating characteristic of social networks is that it is the nodes themselves that change the structure and performance of the network. Different job specifications, office plans, work shifts all affect who communicates with who. Changing these, which is a decision made by a "node" in the network such as for example the head of Human Resources, impacts on the network map.

All of which suggests that social networks, and the contexts developed by them, are unpredictable. Care needs to be taken assuming that simply increasing the "flow" or "connectivity" of a network allows for better, context.

5.4.3 Communication Network Theory

Although Network Analyses may offer us some useful tools, a word of caution is in order. There is as yet very little research indicating how either to support networks or on how different types of content affect and are affected by network topologies.

In their excellent review of communication network theories, Monge and Contractor (2000) identify two holes. The first is that communication networks theories tend to address

> "the creation of networks [rather] than their maintenance or dissolution. This imbalance reflects a serious shortcoming in current theoretical perspectives and empirical research".

That is to say, while models of why communication networks emerge may help us to grow them, they do not help us understand how to support them on a day to day basis. While SNA can give us insights into where to start "growing" a context network, it does not help us understand how to manage them. This is quite possibly a fruitful area for further research.

The second hole recognised is that

> "Missing from the network literature is any systematic theoretical or empirical work aimed at examining the relationship between the structure of networks and the content of messages, symbols, and interpretations that produce and reproduce them. Consequently, we know very little about the manner in which different network configurations (e.g., centralised networks, dense networks) are likely to facilitate the creation of certain types of messages (e.g., supportive, critical). Conversely, little is known about how the production and reproduction of certain types of messages or symbols are likely to influence the structural emergence of communication networks."

5.5 Empirical yardstick: Mapping communication within organizations

Nevertheless, empirical studies of communication *can* be successfully mapped on to it. Boland and Tenkasi (1995) have viewed knowledge intensive firms as a composition of multiple communities with highly specialised, separate knowledge domains. They examine in detail how these communities communicate internally and externally, given their essentially private perspectives (terministic screens) and their associated language models (group ontologies). Essentially, though, they identify two main processes: reflexivity (whereby individuals and groups generate visible representations of the private views and understandings embodied in their perspective); and interpretation (whereby individuals seek through dialogue to understand, if not accept, other perspectives). Two further processes can be inferred from their work, namely persuasion (whereby a group seeks to persuade others to adopt some or all of their perspective without compromise) and negotiation (whereby a group seeks to persuade others to adopt some or all of their perspective through compromise).

So to rephrase, the four processes involved in communication

- Reflexive
- Dialogue
- Negotiation
- Persuasion

These map on to the model as follows. Reflexive communication happens at the local interaction level; negotiation is the process that drives the emergence of an order, persuasion is the argument from that order to constrain and feedback to the reflexive communication; and dialogue is the picture of the system.

5.6 Extensions & Insights afforded by the model

Adopting such a model allows us, by extension, to adopt some of the findings of those involved in the study of networks and non-linear systems. These in turn provide valuable insights, which can be fed into any methodology looking to support such a model. The most significant findings (for the purposes of this paper) are outlined below, and they are referred to later in the description of the Knowledge Service Methodology itself. These are that:
the system is unmanageable; the system is fractal, a design perspective helps us simplify the situation; as do some generic properties of networks; that there are practical, political issues concerned with trust, openness and power in the network; and last, but not least, that context is consensus.

5.6.1 Unmanageability of the system

Perhaps the most obvious upshot of the complex system model is that these systems are fundamentally unmanageable. The way people allocate contexts to situations is fluid, dynamic and unpredictable – it is bottom up. Although top down initiatives to manage the system can work, they do so at the expense of constraining behaviours. Dave Snowden of the Cynefin Centre has an example that captures this wonderfully well. He suggests (2004) we

> "Imagine organizing a birthday party for a group of young children.
>
> Would you agree a set of learning objectives with their parents in advance of the party aligned with the mission statement for education in the society to which you belong? Would you create a project plan for the party with clear milestones associated with empirical measures of achievement? Would you start the party with a motivational video so that the children did not waste time in play not aligned with the learning objectives? Would you use PowerPoint to demonstrate to the children that their future pocket money is linked to achievement of the empirical measures at each milestone? Would you conduct an after action review at the end of the party, update you best practice database and mandate future process?
>
> No, instead like most parents you would create boundaries to prevent certain types of behaviour, you would use attractors (party games, a football, a videotape) to encourage the formation of beneficial largely self-organization patterns (Identities); you would disrupt negative patterns early, to prevent the party becoming chaotic, or requiring the draconian imposition of authority."

The argument here is not so broad. The behaviour we are concerned with capturing and nurturing is how they label content: what emergent patterns do people use to navigate content, and how can a moderating layer both support and take advantage of that. Nevertheless, the unmanageability of such a model suggests that the biggest, and most resilient benefits would come through an approach that marries with the bottom up dynamic, rather than one that imposes a model top down.

5.6.2 Complex Systems are Fractal

In keeping with its complexity pedigree, the model is fractal. That is to say, the same structure applies at successive levels. As a consequence, the same model can be used to expand nodes at one level into systems at the next, for example individuals in a group structure into isomorphic system models of individuals, and vice versa.

The "small pieces loosely joined" at the bottom of Figure 6 might be terministic screens with an emergent "individual" set of classifications; they might be individuals with an emergent group set of classifications; and they might be groups with an emergent organizational set of classifications. The key processes in the model, such as fluidity, emergent order, feedback and self-organization still hold.

Without wishing overly to echo the "pan-unifying theory of everything" excitement, this feature of the system does allow us to straddle an increasingly deep crevasse in Knowledge Management. It stems from the question "Are communities aggregations of individuals or do individuals only have identity as part of a collective?" Traditional KM systems have been "collective" oriented. They have sought to provide one stop shop solutions for groups, and have paid increasing attention to mechanisms such as storytelling (with an assumed audience) that provide a means to negotiation meaning, sharing knowledge and making sense within a group. Identities come second. As such Collective Knowledge Management (CKM) is top-down. More recently, advocates for a more personal knowledge management (PKM) have found their voices. These see communities as aggregations of individuals, and focus on tools and methods for those individuals to manage their knowledge/information to suit their own ends[3]. As such, PKM is focused on the bottom up.

The model from complexity dictates that view both are needed. It also dictates that the stress between both camps is found, not simply in the group-individual perspective, but in the parliament of selves-self perspective.

5.6.3 Network Properties

The study of complex systems' Siamese twin is the study of networks. The science of networks has begun to demonstrate that very different real world networks all share some common rules, whether they be Milgram's "six degrees of separation", the Internet, flight paths or gene expressions. The science is complicated, and there is a very real risk of misinterpreting its insights, especially if guided by lay readings of the research. Treated with due care and attention, though, these common rules would seem to offer an "in" for supporting the model for context emergence.

5.6.3.1 Degree Distributions

The first "in" is that there are broadly three types of degree distribution found in networks. (The degree is the number of direct links with other nodes one node has; the degree distribution is the range of different node degrees contained in the network). Networks can have a range of different topologies, and vary wildly in size, but their degrees all seem to be distributed after a fashion that is: dense, normal, and power-law (scale-free).

Dense networks are those where each node is connected to every other node in the network. Typically these networks are found in small groups or clusters such as offices, juries, sports teams. The degree distribution of these networks is roughly linear: each node is linked to each node, everybody "knows" each other. Indeed, Grannoveter's work on the strength of weak ties can be seen to be looking for an understanding of how, given this interconnectedness, individuals navigate information and content held by other densely interconnected networks.

Normal (bell-curve) distributions are commonly found in random networks. Road networks can be thought of as a random network (with cities, towns and villages thought of as nodes, and the connecting roads as links), as can drinks party networks (with people as nodes, and being previously acquainted as the links). The degree distributions show that most nodes have an average number of links, and progressively fewer have more or less. So in the drinks party example, most people will know or have met the same amount of people, some will know very few, and some, such as the host, will know most if not all.

There are a couple of features of random networks that are worth drawing attention to. First, a random network is resilient but prone to random failure. If we take the road network example, it is tough to entirely destroy or disrupt it because of the distributed nature of its nodes; on the other hand it is prone to random failure, as every node is equally as important as another. Thus, in a random network, we cannot guarantee any kind of outcome – and we

[3] Interestingly, Mark Turner has argued that storytelling is a tool that lone individuals, as well as groups, use to decide on different courses of action. It is an argument that is gaining ground with cognitive scientists.

cannot begin to construct any form of effective, adaptive system. A second characteristic of random networks is the possibility of very long path lengths between any one node and any other one. There is not always a motorway from A to B. This means that, in terms of communication flow, the network does not have an efficient pathway. This feature becomes especially important when we start discussing communication networks – messages (and contextual interpretations) distort, for various reasons, as they pass each node. Gossip is perhaps the clearest example of this and it is discussed later on in this section.

The third common degree distribution that is found in "scale-free" networks, which are networks in which a power-law degree distribution applies. Here, most nodes have only a few links, held together by a few highly connected hubs. One example of this is the Internet, another is the typical airline hub-and-spoke system, and many biological processes (e.g., gene expression, neural networks) appear to follow the model.

Again there are a few noteworthy features of such networks.

First, this distribution delivers a very robust network. There is a redundancy of paths across which information or action can flow, and any random "attack" on the network is unlikely to succeed as a result. By the same token, though, failure of the hubs results in dramatic failure of the networks, thanks to the dependency of most nodes on these link stations. In contrast to the road network example, if hubs such as Heathrow, Charles de Gaulle, and Frankfurt are taken out the airline network, the efficiency of the network as a whole is greatly reduced.

Second, in contrast to networks with bell-curve degree distributions, scale free networks, thanks to the presence of hubs, guarantee short paths lengths from one node to another. A corollary of this is that it is possible to keep adding to the network without reducing efficiency, which has dramatic implications for enterprise growth.

Perhaps most important, these networks are emergent. They do not describe "command and control" top-down system – the hubs can appear anywhere. This ties in with the management concerns outlined above. But they also exhibit enough stability (in the hubs) for us to be able to isolate points of influence. Albert-László Barabási (2003) indicates two laws governing the emergence of these scale-free networks: growth, which means that new nodes are continuously added, and preferential attachment, which states that new nodes prefer to attach to the more connected nodes – a "rich-get-richer" kind-of phenomenon. Taken together, these two laws govern the evolution of scale-free networks, and result in the emergence of a few highly connected hubs. Barabási also introduces the notion of fitness, the ability to make friends relative to everyone else in the neighbourhood, a quantitative measure of a node's ability to stay ahead of the competition. Google's success is an indication of the "fit-get-rich" behaviour of networks. Microsoft is an example of "winner-take-all", meaning that the fittest node grabs all the links (a state described as the Bose-Einstein condensate), a situation where the network develops a star topology.

5.6.3.2 Dunbar Numbers

The second "in" comes from a focus on the nodes in communication networks – humans. Studies on network sizes and their effects on behaviour indicate that Mother Nature puts some fundamental constraints on the people acting in such networks.

Dunbar numbers are named after the theory posited by the anthropologist Robin Dunbar (1993). This is that

> "there is a cognitive limit to the number of individuals with whom any one person can maintain stable relationships, that this limit is a direct function of relative neocortex size, and that this in turn limits group size ... the limit imposed by neocortical processing capacity is simply on the number of individuals with whom a stable inter-personal relationship can be maintained."

Dunbar convincingly supports his hypothesis through studies by a number of field anthropologists. These studies measure the group size of a variety of different primates; Dunbar then correlates those group sizes to the brain sizes of the primates to produce a mathematical formula for how the two correspond.

With this formula, which is based on 36 primates, he predicts that 147.8 is the "mean group size" for humans. Compellingly, this matches census data on various village and tribe sizes in many cultures. That said, without “grooming”, the continual nurturing of the group and relations in it, the mean group size is likely to fall to around 60.

> “The group size predicted for modern humans by equation (1) would require as much as 42% of the total time budget to be devoted to social grooming ... My suggestion, then, is that language evolved as a "cheap" form of social grooming, so enabling the ancestral humans to maintain the cohesion of the unusually large groups demanded by the particular conditions they faced at the time.” (Dunbar 1993)

These findings have been quickly seized by organizations trying to foster Communities of Practice. As Gray has noted (2000), they indicate some simple rules with which to guide the fostering of human networks.

- Group size effects the dynamics of social networks - a community ethos is more likely to arise in human groups smaller than 150
- Network formation depends on social interaction - effective networks arise from regular personal contact that creates a shared sense of community
- Networks can be costly to maintain - time and resources are required to maintain the social ties that support a network
- Hierarchy becomes important as group size grows - more complex societies require authoritarian structures to clarify and enforce social relationships”

The findings overviewed above, both the network degree distribution and Dunbar’s numbers, are extremely tantalizing. They indicate an underlying, simple set of rules at work.

Adapting Ross Mayfield’s (2003) neat synthesis to include the social grooming caveat, we get the following simplified view of network properties.

Network	SIZE	Distribution
Political Network	~1000s	Power-law (scale-free)
Social Network (groomed)	~150	Bell-curve (random)
Social Network (ungroomed)	~60	Bell-curve (random)
Creative Network	~12	Dense (equal)

Table 1: A Simplified View of Network Properties

Such a view highlights a difficult issue. A social network, the network which fits the constraints of our neo-cortex, is not a scale-free network. This is important. It suggests that the benefits deriving from power-law degree distributions – resilience, emergent hubs, and shorter paths (which indicates fewer terministic screens for content to be moderated through) – do not necessarily apply to communication networks in an organizational context. If we are looking for ways to find emergent orders so as to better harvest semantic metadata, then the easiest level is the political network. However, these network properties indicate that, tantalizing as they are, more work needs to be done at the social level, and perhaps the creative level, if the context is to be applicable to an individual. Still, one step towards getting

the benefits of the scale-free network would be to keep a knowledge system as "open" as possible.

5.6.3.3 Social Network Analysis

Social Network Analysis is a useful tool to visualise the emergent communication patterns within a group. Social Network Analysis (SNA), and its subset Organisational Network Analysis (ONA), are used to identify, map and measure the flows between what might be called "knowledge processing entities". The maps consist of nodes (people or groups) and links (communication channels used) and effectively provide a mathematical analysis of human relationships.

The traditional "Hello World!" example is an analysis of an organization. In an organization, there are two models of communication structures. The first is the "ordered" model captured by organagrams, organizational charts, reporting structures and the like. This assumes an order, top down, of how members of that organization communicate. It is dangerous, though, to assume that this faithfully represents how members of that organization actually communicate.

Models of how individuals actually communicate –what might be called "evolved models" – are the business of SNA. These show bottlenecks in communication, who is *actually* central to the connectivity of the network, and peripheral, isolated nodes. There is often a striking difference between these two views of communication.

SNA looks to be able to help us optimise the moderating layer in three ways.

First, should we so wish or indeed have the authority to, it enables us to change the network of actor-moderators in a knowledge system so as to improve the "flow" of the network. This need have nothing to do with IT agents. Distance between knowledge workers, and the physical space in which knowledge workers work can adversely affect the "connectivity" of the network (Monge & Contractor (2000)). Rob Cross and Andrew Parker (2004) note that

> "distance between potential collaborators and other barriers such as hall layout or office design … has a dramatic effect on who collaborates with whom. For example, at a technology organization our network analysis showed that tall cubicle walls were blocking the communication between two teams. When the walls were lowered, the teams discovered greater commonality in their work and pushed for additional space for shared brainstorming."

A second boon to optimising the moderating layer that SNA looks to provide is that it allows us to visualise the emergent "dialogue" in the system – the second level of communication highlighted in Boland and Tenkasi's analysis. SNA can be used to analyse content access of members of a group. Valdis Krebs (2004a) for instance, has analysed the purchase patterns of the top hundred political books on Amazon. The analysis was made possible by the "customers also bought" links. These are a feature of Amazon that allows customers, when viewing, say, Robert Baer's "Sleeping with the Devil" to see what other books customers who bought "Sleeping with the Devil" bought. The map Krebs came up with, shown in Figure 9, was essentially a picture of a dialogue.

It is a picture of who valued what content, with no effort at changing those values. Those anti-Bush bought anti-Bush books, those pro-Bush bought pro-. The lack of middle ground is striking. Such a "dialogue" also affords a reality check to preconceptions. Many thought that Woodward's book *Plan of Attack* was be sufficiently neutral to attract both pro- and anti-Bush readers. Krebs' analysis shows this to be misguided.

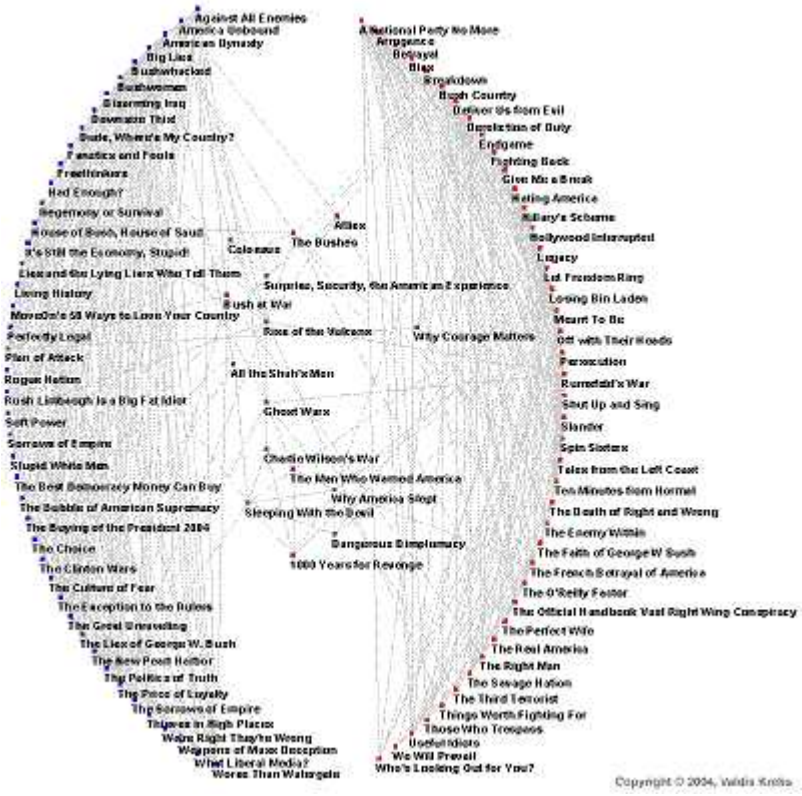

Figure 9: Valdis Krebs SNA of Political Bestsellers

The final way in which SNA can help us optimise the moderating layer is by allowing us to see who the key moderators are in a knowledge system (and so who, at least initially, it is most important to support). Returning to Krebs' analysis, his advice, from a social network analyst's standpoint, is to focus on the edge nodes and the bridges.

> "What can do you do with this information during the 2004 campaigns [for the American Elections]? See someone reading *Sleeping with the Devil* ? That is someone you can talk to about your candidate. If they are reading *Bushwacked* or *Dereliction of Duty* -- the most central books in each cluster -- then either give them a high-five or a sneer, you will not be successful in moving a reader from deep in one cluster, over to the other cluster. All you can do is focus on the edge nodes and the bridges." (Krebs, 2004b)

So, although the context-creating networks in knowledge systems may be technically unmanageable, along with the design perspectives, social network analysis lets us simplify the task by enabling us to identify where most contexts are being affected.

5.6.4 Context is consensus

Scholars studying organizational communication have long made theoretical and empirical distinctions between formal and emergent networks. Formal networks are those pictured in organagrams and organizational charts. These were presumed to represent, therefore, the channels of communication within an organization. On Weber's analysis (1947), orders were transmitted downward and information was transmitted upward. The notion of emergent networks rose out of the concern that such a model was too restrictive. It ignores the "unstructured" communication that happens around coffee machines, around cigarette lighters and the like. Social Network Analysis, as discussed in the next section, offers graphic representations of the way people *do* communicate, as opposed to how they *ought* to be communicating. For example, Stevenson and Gilly (1991) showed that managers faced with

a problem, more often than not, ignore the formal network and its channels from problem solution. Instead they make use of their personal contacts. Again, Albrecht and Ropp (1984) discovered that

> "Workers were more likely to report talking about new ideas with those colleagues with whom they also discussed work and personal matters, rather than necessarily following prescribed channels based upon hierarchical role relationships".

Now both types of communication network clearly exist (although arguably formal networks are in abeyance as companies try to make their organisational charts more fluid). Nevertheless, the fact that formal networks do not capture the richness of communication that exists organisations highlights a distinction that is critical in understanding how best to support context. People do not make decisions based on whom they ought to talk to, but people will act, for the most part, with on whom they should act with. As a result, the contexts people use to help them make decisions are not fully modellable through organizational charts.

Drucker offers a description of two different decision-making processes that is useful here. He explains that

> “the Westerner and the Japanese man mean something different when they talk of making ‘a decision’. In the West, all the emphasis is on the *answer* to the question … To the Japanese, however, the important element in decision making is *defining the question* . The important and crucial steps are to decide whether there is a need for a decision and what that decision is about. And it is in that step that the japans aim at attaining consensus. Indeed, it is this step that, to the Japanese, is the essence of decision. The answer to the question (what the West considers the decision) follows from its definition. During the process that precedes the decision, no mention is made of what the answer might be … Thus the whole process is focused on finding out what the decision is about, not on what the decision should be.” (Drucker,1974)

These two views of decision-making can be usefully combined. Projects, it seems, involve two discrete phases: consensus, or aiming at attaining the best initial context for the decision through emergent networks; and action, which is based on that consensus which flows through formal networks.

6 Social Tools

So far, though, the argument has been firmly theoretical. To balance this, Templeton has reviewed a number of different possible software options for supporting the moderating layer, the results of which can be seen in Appendix A. Happily, there are concrete tools and software being used by individuals and organizations that "fit" with this model. That they *are* being used serves to ground this complex systems model, and also serves to help focus how a methodology to support this model might be implemented. These tools have been labelled "social tools".

6.1 Definition

The name "social tools", like "knowledge management" is almost so vague as to be useless. To some a table is a social tool, to others it is not. Broadly, though, and certainly for the purpose of this paper they are:

- software (not tables, phones or conversations)
- network aware (in that they encourage rather than order collaboration), and
- amenable to the notion that meanings are socially negotiated, rather than out there.

A number of different social tools exist, wikis and weblogs. In this section, these are briefly reviewed, indications of how they map onto the complexity model are shown, and examples of their working (and not working) in business and other communities are given.

6.2 Example: Weblogs

6.2.1 What they are

Despite their names, these are not children's toys. The term "weblog" was first coined in 1997 to describe an easy means of publishing to the web. A weblog, or simply blog, is perhaps easiest understood as an online, linkable diary. The content in blogs can vary from personal musings to rants to bookmarks to fuller think pieces. However, common to them all at the moment are the two underlying metaphors of journal and conversation.

Entries into the blog journal are typically ordered and presented in date-order. Over and above that, those keeping the diaries ("bloggers") can categorise their posts. These categories and classifications are personal – it is perfectly conceivable that one man's "idiocy" classification is another man's "genius".

The second underlying metaphor is that of conversation. Although some blogs will be personal and private, the majority are freely commentable on and linkable to. People seeing a post of interest on a blog can leave remarks, helpful or not, on the post. If they have a blog themselves, they can link to the post that interests them and comment on it in their own "diaries". As such conversations take place round content online. Additionally, many blogs will have a "blogroll". These typically give a list of either other people whose blogs the blogger reads, or sources of information the blogger values, or a combination of the two.

Now, although at heart, these online journals seem to be fundamentally personal, group weblogs exist. Here, the journal is in effect a compendium of different journals. One increasingly common use for such compendiums is the project log (also known, unattractively, as a "plog"). There is arguably a value in such group blogs in that they allow those who are unfamiliar with posting, or worried about having their essentially personal takes on events and ideas published online, to garner comfort from a group.

There are various descriptions of blogs. Two perhaps stand out. The first comes from Dave Winer of Harvard University's Berkman Centre. Winer sees blogs as

> "often-updated sites that point to articles elsewhere on the web, often with comments, and to on-site articles. A weblog is kind of a continual tour, with a human guide [whom] you get to know. There are many guides to choose from, each develops an audience, and there's also comraderie [sic] and politics between the people who run weblogs, they point to each other, in all kinds of structures, graphs, loops, etc." (Winer 2002)

The second description came at a recent Microsoft summit. Bill Gates (2004) neatly summed up RSS and blogging as an effective middle ground between email and web site publishing.

> "Whenever you want to send e-mail you always have to sit there and think who do I copy on this. There might be people who might be interested in it or might feel like if it gets forwarded to them they'll wonder why I didn't put their name on it. But, then again, I don't want to interrupt them or make them think this is some deeply profound thing that I'm saying, but they might want to know. And so, you have a tough time deciding how broadly to send it out.
>
> Then again, if you just put information on a Web site, then people don't know to come visit that Web site, and it's very painful to keep visiting somebody's Web site and it never changes. It's very typical that a lot of the Web sites you go to that are personal in nature just eventually go completely stale and you waste time looking at it …
>
> And so, getting away from the drawbacks of e-mail -- that it's too imposing -- and yet the drawbacks of the Web site -- that you don't know if there's something new and interesting there – this [blogging and RSS] is about solving that."

6.2.1.1 Really Simple Syndication (RSS)

It is difficult to understand blogging without understanding RSS. RSS – or Really Simple Syndication – is effectively a web syndication protocol that is primarily used by news websites and weblogs. RSS allows a web developer to publish content on their website in a format that a computer program can easily understand and digest. This allows users to easily repackage the content on their own websites or blogs, or privately on their own computers. RSS repackages the content as a list of data items, such as the date of a news story, a summary of the story and a link to it. Streams of these packages are called "feeds", and in many ways these are a modern version of the old ticker-tape service used in news companies.

A program known as an RSS aggregator or feed reader can then check RSS-enabled web pages for the user, and display any updated articles that it finds. Importantly, this is more convenient than having the user repeatedly visit their favourite news websites, because it makes sure that the reader only sees material that they haven't seen before. These aggregators come both as software to be installed on the client machine, or as web-based services, thus making the user's feeds available on any computer with Web access.

At the top level, a RSS document is a <rss> element, with a mandatory attribute called version, that specifies the version of RSS that the document conforms to. Subordinate to the <rss> element is a single <channel> element, which contains information about the channel (metadata) and its contents. Behind the success of RSS is the notion that people are managing feeds, not content.

To paraphrase Bill Gates, blog software allows users to post to the web quickly and easily knowing that only those who are interested will subscribe. RSS aggregators allow users to opt in or opt out of receiving certain feeds easily, giving you more control of your inbox.

6.2.1.2 Post Pings and Trackback

If enabled, post pings allow the blogger to inform content aggregators every time they add a new post to their blog. For instance, should a blogger have a category in their blog on knowledge management, it is possible to set up their software so that for each new entry in that category, a service such as KMPings (http:// www.highcontext.com/**kmpings**/), which aggregates blog posts on Knowledge Management is updated. Those subscribing to KMPings as an RSS feed then receive all the posts from interested (and pinging) parties on knowledge management.

Trackback is essentially a bi-directional ping. It allows a blogger to see who has seen the original post and has written another entry concerning it. The system works by sending a 'ping' between the blogs, and therefore providing the alert.

Trackback typically appears below a blog post and shows a summary of what has been written on the target blog, together with a URL and the name of the blog. It provides a useful way of tracking conversations between blogs, and watching developing communities.

6.2.2 Examples of Blog use

Blogs are rapidly being adopted by organizations in the technology, media and research sectors. They are used both “in front of” and “behind the firewall”.

Microsoft, Sun (including their CEO Jonathan Schwartz), IBM and Dell all have developer community blogs[4]. The spirit behind all of them is similar to that described on the Sun site, where they say that

> “This space is accessible to any … employee to write about anything.”

Disney - who have interestingly called blog posts “public email” - are integrating blogs and wikis into their intranet, as are the BBC. And the business research firm Jupiter Research was one of the first to go public with analyst blogs in January 2003. Lenn Pryor, Microsoft's director of platform evangelism explained the attraction of this sort of blogging as "a way to scale our ability to communicate with customers in an open and honest way." (Festa, 2004)

Behind the firewall, Michael Schrage (2004) of MIT indicates that corporate plogs – project logs tracking the status of projects - while by no means the norm are becoming increasingly common in the Fortune 500 companies.

Amazon has approached them in a different manner through their introduction of plogs, which in this case stands for personal logs.

> “The Plog Service is a personalised blog.... Your Amazon.com Plog is a diary of events that will enhance your shopping experience, helping you discover products that have just been released, track changes to your orders, and many other things. Just like a blog, your Plog is sorted in reverse chronological order. When we think we have something interesting or important to tell you, we'll post it to your Plog.” (Amazon Website, 2004)

6.2.3 Key features

In general, then, the key features of blogs (and aggregators) are these:

- The metaphor underlying blogs is that of a conversation
- Aggregators allow you to manage feeds not content; blogrolls allow easy
- RSS integration allows easy on-line publishing and user choice of “public email”

[4] For details and URLs of the commercial blogs mentioned in these paragraphs, please see the References section under “Commercial Blogs”

- Blogs contain personal points of view
- Blogs and sharable bookmarks allow group selection of relevant content.

6.3 Example: Wikis

6.3.1 What they are

A *wiki* is one of two things: either a specific type of hypertext document collection or the collaborative software used to create it. The root of the term is the Hawaiian phrase for "fast". Indeed, it is the speed of creating and updating pages that is one of the defining aspects of wiki technology. Generally there is no prior review before modifications are accepted, and most wikis are open either to the general public or to all persons who at least have access to the wiki server. Often, even registration of a user account is not required.

Those who have access to the server can create new wiki pages, edit current wiki pages, and create new sub-categories. Although there is often some editorial control (such as, for example, making contact details static and unchangeable), most of the content contained within a wiki is effectively open source.

Each page typically contains a large number of links to other pages; hierarchical navigation pages often exist in larger wikis, but do not have to be used.

6.3.2 Examples of Wiki use

Wikipedia (http://en.wikipedia.org) is an on-line encyclopaedia that runs on a wiki. Although it was only founded in 2001, it is already a serious competitor to Encyclopaedia Britannica. As of July 2004, it contains 310,000 articles in English, and 530,000 articles in other languages ranging from French to Arabic to Korean. And it is free.

Anyone can contribute to the encyclopaedia, though there is some basic editorial control. 6,000-odd volunteers covering a huge range of subjects, even though it does better on science and technology than on arts and culture. Not surprisingly, the articles are of uneven depth and quality. If you find an error, you are welcome to suggest a correction. And if you find a topic that isn't covered, you are welcome to create a new article. (An editorial group decides which corrections and contributions merit posting.)

There are, predictably, quality issues. In June 2004, the Virginian-Pilot newspaper discovered that "the online encyclopaedia 'Wikipedia' created a version of Chesapeake's history that was literally a bunch of bull." (Virginian Pilot, 2004) The edit in question was put in on May 2 2004 and not removed until June 3 2004. More fundamentally, there is the issue of conflicting points of view on the same subject, such as, for example, justifications for the "war on terror". A different, but related problem stems from contributors trying to game or vandalise the system. For example, links to own pages can be added, allowing for spam entries to show up.

Quality is ensured in a number of different ways. First, there is the editorial control. Second, there are the guidelines for the users, indicating that articles should be written from a neutral point of view (NPOV). Third, there are facilities for those with Points of View to discuss them with other contributors with conflicting Points of View. Fourth, and arguably the most effective, is that the system is designed in such a way as to make it difficult to ruin (Ciffolilli, 2003). Previous wikipages are archived in such a way that it is far easier to revert to the last entry than it is to deface the current one. A fifth and subtler means of ensuring quality is Wikipedia's presentation of itself as a work in progress. It claims to be an aggregate of contributors' points of view, and one that is never finished. Readers using the encyclopaedia can use this to help them make their own quality judgements.

Currently, though, perhaps the best measure of quality is the number of references Wikipedia is beginning to garner from traditional news journalists and researchers. 72 news outlets cited, quoted or referred to Wikipedia in the English language "press" (though many were

Internet and television sites) from January 2003 to March 2004. (Lih, 2004) Perhaps as importantly, a slowly increasing number of publications looked to be using the wiki as a more regular resource.

Publication	Citations
Daily Telegraph online	41
Guardian	5
Sydney Morning Herald	5
Slashdot.org	4
NewScientist.com	3
ChessBase News	2

Table 2 : Publications citing Wikipedia multiple times

Interestingly, part of the reason Wikipedia is becoming a useful search resource may be due to the aggregation it facilitates. Whereas with Google, and other search engines, results matching your criteria can number in the thousands, one of the premises behind wikis is that those thousand references can in some fashion be aggregated into a single wiki page. Take for example the entry on MK-ULTRA, a cold war CIA mind control program). Google returns 63,900 documents; Yahoo returns about 346,000 documents; Wikipedia returns one. And the one page it returns is fuller, and more detailed than the entries on either the Encyclopaedia Britannica or Microsoft's Encarta.

Wikis are also used successfully in many corporate intra- and extranets, though mainly by the more IT literate employees. TWiki (http://www.twiki.org), one of the leading wiki software platforms, cites various success stories.[5] It is used by all SAP's development departments (TWiki SAP, 2004); British Telecom (TWiki BT, 2004) use it in their development department, and have recently launched the UK Telco B2B forum (www.telcob2b.org.uk) to provide information on and develop B2B standards and architectures for eBusiness integration of communications service providers in the UK; at Disney, it "has quickly become *the* central resource for ideas, notes, "how to's", specs, and brainstorming" (TWiki Disney, 2004); and Motorola use in their Systems-On-Chips Design Technology division across 7 different webs extending their 'team' from an on-site project team to "a virtual team including members in Germany, UK, France, Australia, Russia and the US, with about 60 regular contributors (and growing)." (TWiki Motorola, 2004)

While this paper is by no means meant to eulogise TWiki at the expense of other wiki solutions, two further quotes from their users are of interest given the model being proposed. The italics in both are inserted.

The first is from Martin Roberts of BT.

> "TWiki has been *introduced from the ground roots up* and the response has been extremely positive. We have had times when the *openness* of the TWiki has been questioned, but with only a few exceptions we have managed to persuade most users that it is safe to use. The ability to show old versions of documents is crucial. Within the company we have other tools for document management yet the TWiki *ability to be used by anyone for anything has proved very persuasive* ." (TWiki BT, 2004)

Wikis, as the model would suggest, have been adopted "bottom up". Equally, it is usable by anyone – there is no profiling, modelling or GroupWare style fragmenting of identities across different applications.

The second quote is from Crawford Currie of Motorola:

[5] For details of TWiki and its users as mentioned in the following paragraphs, please see the References section under "TWiki".

> "The way these different webs have *grown* is interesting. Each web has an owner, usually a project team leader. Some owners have customised their pages for their project, others have just taken the default. Looking through the webs, I can see different uses in each. Examples:
>
> Requirements capture
>
> Newsgroup
>
> Cooperative authoring environment
>
> Rapid production of web pages [for subsequent publishing]
>
> Issues lists
>
> Meetings calendar
>
> There is strong interest in using the knowledge base for customer support, but here it is *competing against a "corporate authorised" mechanism* so we're not going to push this." (TWiki Motorola, 2004)

Again, these wiki webs "grow" in a bottom up fashion, with groups modelling their wikis for their own needs. Lastly, they aggregate: requirements, issues, authorial opinions, and news. While users (and employees) appear to be finding the wiki approach useful, corporate "authorities", perhaps understandably, seem to be concerned about the lack of control they have over such an approach.

6.3.3 Key Features

In general the key features of a wiki are these:

- The metaphor underlying wiki development is that of a document
- Wikis make it quick and easy to update content
- Wikipages are open source, shared content
- Wikis contain aggregated points of view (POV's)
- Recent changes to pages can be turned into RSS feeds
- Editorial control is minimal, content is developed "bottom up", and structure emerges

6.4 Example: Social Bookmarks

Social bookmark tools (such as del.icio.us, pluck, and furl.net) offer a public, collective link-list; personal link-list; gives easy access to the personal pages of others, and the means to subscribe to anyone's link-list (thus creating one's own collective link-list consisting exclusively of contributions by selected individuals).

As a "bookmarks manager," these tools offers the ability to add metadata to collected links, such as short and extended descriptions, and "tags"--words used to classify and categorise a link.

6.5 Example: OhmyNews

A technologically supported, but less "geeky" example of the ethos behind social tools is OhmyNews (http://www.ohmynews.com). OhmyNews is a South Korean online newspaper that attracts an estimated 2 million daily readers. It is widely credited with helping to elect South Korea's new progressive president, Roh Moo-hyun. And it has done this in just 4 years. What is most striking about the newspaper is its staff: it has more than 26,000 registered citizen journalists, and anyone can sign up to be a reporter

Its founder and CEO, Oh Yeon-Ho says that

> "The main concept is that every citizen can be a reporter... We changed the concept of the reporter... a reporter is the one who has the news and who is trying to inform others." (Gillmoor, 2003)

Some of the key features of the process involved are: that OhmyNews edits and posts nearly 150 stories a day with the help of less than 50 professional editors and reporters; that it is still South Korean citizens doing most of the contributing; that writers get paid in relation to the "newsworthiness" of their story and must use their real name; and that stories enter an editing queue and are fact-checked before being published.

In many ways the production process does still resemble that of a daily newspaper with editors, freelancers and deadlines. There is a significant difference, though, in that this production process takes place in public and anyone is free to contribute. There are clear overlaps with the dynamics that seem to be behind the success of wikis, blogs and other social tools.

6.6 "Harvesting" the emergent order

An obvious area of concern with all these social tools is that of taxonomy. As already seen, if we cannot make the transition at some level to ontological mark-up and semantised metadata, then we are likely to miss out both on the power IT Agents might afford us. While at first sight, and especially if one takes the strong social constructivist view, this stamping may not be possible given the fluidity and dynamism of the small pieces, progress is being made that looks to enable the development of taxonomies from blogs, link blogs (blogs with no comments, simply links) and wikis.

Blogging, as we have seen, is essentially personal. What they do afford us are what have been called "semi-formal" representation of concepts in a domain and their relations. Tools in development using language analysis software are now able to extract concepts and relations between these concepts from blog posts (with single and multiple bloggers) (e.g. Anjewierden et al, 2004). The key point here is that while these emergent taxonomies are initially personal, as other bloggers post, there is the ability to extract shared concepts and shared relations, a consensus of views, and taxonomies. Users, though, are making their own classifications, and as such modelling themselves.

Given this, and given that social bookmarks are effectively "link blogs", the tags that individuals give when they classify their own blogs can be aggregated and analysed in a similar way. Again, taxonomies emerge from consensus. Again, it is the users who are modelling themselves.

Wikis perhaps allow for the easiest emergence of taxonomies in that they follow a document metaphor rather than a conversational one. The documents created will usually be structured in some way, similar to a contents page, and there will be cross-references between pages in many cases.

6.6.1 Personal Task Taxonomies.

User models focus on user, system and task. So far, the argument has been concerned more effective ways of modelling the user's context and the system, or systems, within which the user is acting on content. Nothing has been said concerning the user's tasks. Intriguingly, a suggestion of William Jones may allow us to use many of the same dynamics apparent in social tools for modelling tasks.

Jones' suggestion (2004) is the use of "Personal Unifying Taxonomies" or PUTs. He describes these as follows:

> "The categories of the PUT represent tasks (goals, projects, activities) that the person wants to accomplish. As [Figure 10] illustrates, portions of a PUT may do double duty as both a task decomposition and as a way to organize information relating to the tasks in this decomposition. When information is organized in this way, discrepancies between the information needed and the information currently available become more apparent."

Attend Interact'03 conference

Submit paper

...

Get authorization for international travel

E-document containing travel authorization form

Get plane reservations

Web reference (Favorite) for flight schedules

E-mail to travel agent

E-mail from travel agent with confirmation – pointing to web site

E-document on hard drive containing itinerary (also printed out)

Get hotel reservations

Web references to information on hotel location and pricing

E-mail to desired hotel

E-mail from hotel with reservation confirmation

Directions to hotel saved as file on hard drive

Arrange for substitutes at work

Submit expense report

...

Figure 10: Example of a PUT as a means of classifying information

Essentially, then, PUTs are Personal Task Taxonomies, structures off which information can be "hung". In Figure 10, under the "Get Plane Reservations" category, reference is made to the users "favorite". In the social tools world, this could well be navigated by, for example, some form of social bookmark. Perhaps more tantalizing, though, is the idea that these structures could be shared – made "social" as it were – so that some form of consensus could be found for how certain tasks were best completed. Indeed it would address the dark clouds that Jones spies on the horizon. For despite arguing from what has been termed above the positivist viewpoint, Jones raises a number of interpretivist concerns.

> "Into this sunny picture of PUT potential drift a few dark clouds of disclaimer. A PUT may not work for many people for any of a number of reasons. PUTs may turn out to be too difficult to develop, or to use and maintain on an ongoing basis. Many intended PUT owners, even if willing, may never accept the notion of a PUT and the changes it brings to information organization and habits of information management. People may never develop the discipline to make consistent use of a PUT. People may feel that a PUT is too constraining for their uses of information. And so on ...."

There is no reason, in theory, why personal task taxonomies, structures outlining individual approaches to tasks, could not be shared in such a way as to allow the emergence of group or organizational task taxonomies.

6.7 Features of the Model that Social Tools Support

Returning to the proposed model, and with the Boland and Tenkasi analysis of organizational communication in mind, it can be seen where these social tools are supporting the process of content contextualisation. (See Figure 11)

Blogs, social bookmarks, s, and aggregators facilitate conversation and the development of personal perspectives. As such they aid with “introspective” communication.

Wikis, with their emphasis on the document, encourage negotiated, system neutral perspectives. As such they encourage, and their end product is (if they have one) negotiated communication.

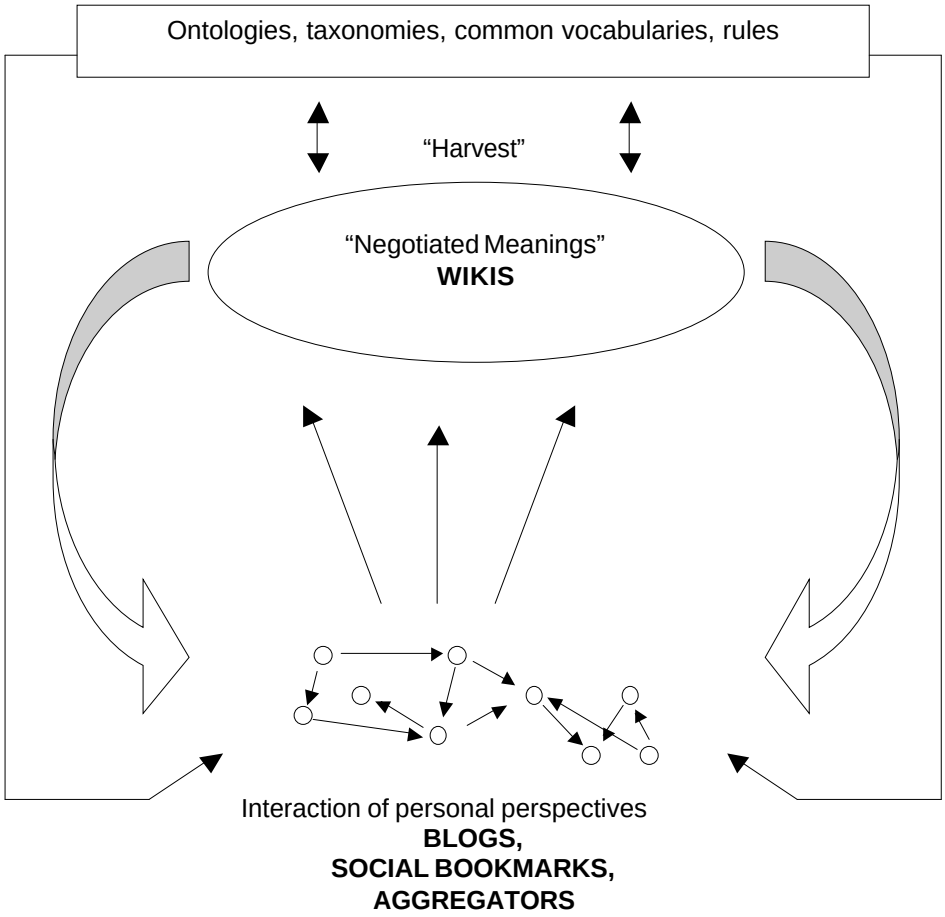

Figure 11: Role of Social Tools in Content Contextualisation

7 Conclusions

So far, then, the argument can be summarised as follows.

- We need to optimise the moderating layer in knowledge systems to improve knowledge worker productivity
- Optimizing this layer is essentially a context problem
- This context problem is typically addressed in one of two broad ways.
- The first way – the positivist approach – looks to model actor, task and system to embed context into the moderating layer. This plays to the strengths of the IT Agents in the moderating layer
- The second way – the interpretivist approach – argues that context cannot ever be successfully embedded into a system because individuals and their environments are complex and ever changing. This acknowledges the needs of the human agents in the moderating layer.
- Non-linear dynamic systems provide us with a model which might offer a valuable compromise between the positivist approach that supports much of context-aware computing, and the interpretivist approach, which makes sense of how we as human agents work with and develop context.
- Such a model offers some valuable insights as to how, and when best, to model user, task and system in such a way that helps improve the performance of the moderating layer.
- Consequently, using this model to help support knowledge systems should improve knowledge worker productivity.
- Happily, tools already exist and are being developed that support implementations of this model.

The argument from the theory and practice outlined above indicates various lessons that need to be taken to heart in any effort to optimise the moderating layer of a knowledge system.

The key lessons identified are:

- IT agents need semantic metadata to be able to effectively moderate content. Economic considerations mean that the only type of metadata that can be supported is automatically created. Ideally this will be transparently created, rather than heuristically.
- Taxonomies and ontologies are snapshots of consensus (the system, user and tasks they depict are fluid, fragmented and political). Treating them as anything else affects the ability of the system to be context-sensitive, hence there is a need for some ontological oscillation.
- Where used, taxonomies and ontologies inherently affect the behaviour of the system.
- External modellers also potentially bias the model. The best modeller of the actor is the actor; the best modeller of the group is the group; the best modeller of the organization is the organization. This is not to say that there is no place for external modellers. They may provide very valuable foils and “reality checks” to actor, group and organisation. Nonetheless, the less observer bias there is, the more useful the model.
- To be relevant, context must be based on consensus.

- Knowledge systems that are context-aware, and so consensus enabled, are more resilient if developed from the bottom up rather than managed from the top down (Snowden's Children's Party metaphor)
- Projects and plans form a natural frame for a dynamic environment, as shown by Weick's work on Sensemaking.
- Effective decisions are based on consensus among perspectives and are action-oriented.
- The mechanisms that work for individual content moderation should work for group content moderation thanks to the fractal nature of complex systems.
- Tend to openness where possible (Open source content, "given enough eyeballs all bugs are shallow", scale-free and bell-curve distributions of networks)
- Broader, open aggregation allows for more resilient systems,
- Overt or forced user profiling narrows possible aggregations of terministic screens. The context cycle is: conversation, document, structure conversation, document, and so on.
- Using "dialogue" tools, such as Social Network Analysis, to identify network patterns allows you view emergent structures. These can then be encouraged.
- The practical lessons of current social tools, suggest that the system should be only as resilient as it needs to be, and as reliant on self-policing as it can be, such as shown by the minimal editorial input into Wikis.

The practical experience from aggregators and communications networks suggest that individuals manage channels, not content.

From these lessons the following proposed Knowledge Service Methodology has been developed.

8 Knowledge Service Methodology

STEP 1: MAP SYSTEM – "Landscape"

IDENTIFY CONSENSUS POINTS

The first stage in the methodology is to identify the consensus points in a system.

The project, or plan, is a useful first frame for mapping the knowledge system. Each actor in the system may be involved in a number of different projects or plans, involving differing numbers of other actors/moderators. Again, the lesson from the complexity model is to map one project initially. Guidelines for sensible choices will hopefully become apparent during the course of the METOKIS project. For the moment, we have taken one facet of each organizations overall model.

Examples of mapped projects and consensus points can be found in the MCM Business Models (METOKIS Deliverable 09).

The three use cases – Executive Support, eLearning Publishing and Clinical Trials – were all mapped using the situations design methodology. In each of the cases, there are some clear consensus points:

Alternatives to situations design methodology are soft systems methodology and the like

	Project Frame	
	Consensus	Action
Clinical Trials (Ymega)	Protocol Design	Trial Enactment
ELearning (Klett)	Concept Development	Production
Executives (Templeton)	Agenda Setting	Arranging Meetings

IDENTIFY ACTORS

The next "landscaping" stage is the identification of actors in the system being supported. This is a two-stage process.

The first stage is to "roll call" those actors directly involved in the project scoped when identifying the consensus points. These are the individuals who will be playing the dual role of actors and moderators in the knowledge system.

The second stage is to establish boundaries for the system. Because of the recursion implicit in the moderating layer, care needs to be taken that the actor/moderator dual role does not result in an infinite regress of potential influences in the project.

Boundaries could be: explicit project members, people "behind the firewall", departments involved in the project.

Although as argued above openness is beneficial, these boundaries should hopefully indicate who has a vested interest in the system, as opposed to those who have an unpredictable interest in the system.

IDENTIFY COMMUNICATION CHANNELS ("Formal" and "Emergent")

Having identified the actors, the next "landscaping" task is to identify the communication channels. This needs to work on two different levels, reflecting the formal and emergent networks of communication.

a) *Identify "formal" communications channels between actors*

.

These are the channels that stem, effectively, from the organagram, group membership structure or project outline. Models of these have been made for Klett, Ymega and Templeton in the METOKIS project, again using MCM St Gallen's situations design methodology. This highlights the roles and rights of different actors in the system "as they ought to be"

b) *Identify "emergent" communications channels.*

These are the channels that have grown up around the organagram. While there may be a close, or even exact, correlation between these and the "ordered" communications channels, it is by no means certain. Social Network Analysis is a promising tool for mapping the actual communication channels within a system. As well as highlighting communications issues such as content bottlenecks, such a map also helps both to highlights natural "moderators" of content within a system (the edge nodes and the bridges), and give a first view of the "dialogue" of the system.

CONTENT & TOOLS AUDIT

The last phase involved in landscaping the system is to identify the content types existing electronic media tools commonly used by the actors (e-mail, Microsoft Office) and the content sources used by them, both paid for (e.g. Reuters), and free (e.g. internally produced documents).

This aims to establish:

- *Tools preferences*

  This is so as to be able to encourage adoption of systems through the tools that actors currently use, as opposed to introducing a raft of new tools. The thinking behind this is that, while more technically enthusiastic actors will willingly experiment with new tools, the majority of actors are time-poor and prefer more functionality from their existing tools to more functionality from more tools.

- *Costs of content*

  Without establishing the costs of the content bought in to the organisation, metrics for productivity improvement cannot be worked.

- *Content for quick wins*

  Identifying possible content feeds with which to pump prime the knowledge system is an effort at encouraging early wins. These might be paid for or free, within or outside of the organisational "firewall".

STEP 2: SEED SYSTEM – "Sow"

At this stage, a small reality check is in order. There is a tacit assumption behind many social tools that they foster communities. However, people join communities for different reasons. Not everyone will have the time to be an active participant; moreover not everyone will want their details publicly available to other members of the community. Those who sign up, or pay

for membership, will not necessarily be vocal contributing members. Many, in fact, may join for passive reasons: they want moderated quality content which has a high relevance "hit rate", the may join because someone from their company has to join. In short, the "free them and they will contribute" is not necessarily as rosy as some of the pictures the social tools paint. Lurkers and free riders are a familiar phenomenon in online and offline communities.

The aim of mapping the system (Step 1) is to try to help decide:

- first, who the most likely moderators are, so that their needs may better be supported;
- second, where, if at all, they are involved in the consensus stages of a decision; and
- third, who, of the potential "passive customers" may most easily and most willingly be involved.

One should stress that the Knowledge Service methodology aims to be firmly pragmatic. While the goal, of course, is a perfect optimisation of the moderating layer, a stepping stone to that is seen to be an improved moderating layer. The contention is that the more "eyeballs" content has the less "shallow" its context. More is better, but all is not necessarily achievable

As highlighted both by the Children's party metaphor above, and the fact that complex systems are unpredictable and difficult to manage, there is a need for staged implementation of tools to improve the moderating layer. Paradoxically, "selling" the system is likely to ensure its underperformance. The metaphor here, then, is a logical extension of the notion that organizations are ecosystems rather than complicated machines.

Having mapped the consensus points, communications channels and tools, Step 4 is to choose the "fertile ground". Actors or groups in the project scope who are already using social tools are an ideal target, since their involvement in a pilot stage heightens the chances of a quick win. As noted by communications networks workers, tool usage seems to follow a contagion dynamic. By extension, if we successfully identify the key "transmitters", then we can speed up the adoption rate.

Equally, edge nodes and bridge nodes should show who they key individuals involved in consensus are, and these provide natural targets.

Nonetheless, coaching, seminars and one-to-one sessions are critical at this stage. Explanations of what the social tools do and how they can help the individual knowledge worker need to be ongoing. Again, the focus needs to be on end benefit rather than details of the tools working and why they are clever.

Seed taxonomies are also needed. The lessons from the Positivist thinkers indicate that semantic metadata is critical to meaningful machine-to-machine communication, but that the economics of such considerations restrict the way such metadata can be authored. Metadata needs to be authored automatically but seeded by human authored structures. Initially any such metadata should be as broad and as unconstraining as viable. Project scoping may give cues as to what are sensible frameworks. Rhetorical structures may also give insights as to possible options.

The table below shows the structure of the approach, and where in the consensus forming stage different tools are seen to be most applicable. This can provide a useful guideline for what seeds to sow where.

	PROJECT FRAME				
		Consensus			**Action**
Clinical Trials (Ymega)		*Protocol Design*			*Trial Enactment*
ELearning (Klett)		*Concept Development*			*Production*
Executives (Templeton)		*Agenda Setting*			*Arranging Meetings*
Metaphor	Conversation	Conversation	Document	HARVEST	
Individual / Group	Individual	Individual/Group	Group		
Communication Type	Internal	Internal/Negotiated	Negotiated		
Tools	e.g. Blogs, Instant Messenger	e.g. Group Blogs, Forums, Discussion Boards	e.g. Wikis		KCO, KCCA, Semantic Web

STEP 3: ENCOURAGE EMERGENCE – "Grow"

Following the table above, three stages in consensus are outlined, all of which ideally need support. The first two have as "conversation" as their underlying metaphor, the third has document. This aims at capturing the shift from negotiation to persuasion. Ideally, all three levels will be supported, though again, this is dependent, first and foremost, on the actors in the network.

ENCOURAGE CONVERSATION

The first of these is "personal conversation". While encouraging this is not intended to mean nurturing a group of individuals who mutter to themselves, it is not far off. Personal conversation corresponds to Boland and Tenkasi's "introspective communication". It is the ordering of one's own view of the world, Weick's identity construction.

Conversation can happen at the group and individual level. Blogs, whether individual or group, may be a useful tool for this.

Where possible, given the constraints of the social network sizes, these conversations should be as open as possible. Disney's "public email" approach is helpful.

Tools supporting these conversations are still being developed. Two major considerations are relevant here, though. First, RSS aggregators enable easy management of feed subscriptions, just as they enable easy modelling of users preferred feeds. Second, easy online publishing allows for broader access than simply those subscribing to a particular feed.

ENCOURAGE COLLABORATION

Collaboration, on this model, is collaboration around a document. The lessons from the success of wikis indicate minimal editorial control, and ensuring that access is as open as can be. Again, these documents allow for easy aggregation of consensus and Neutral points of View. As important though, is that users can subscribe to sections of interest as RSS feeds allowing them to be alerted when key changes of interest to them are written.

Both of these need minimal taxonomies to get started.

STEP 4: REMOVE OBSTACLES – "Prune"

There are various different potential obstacles to emergence – ease of use, organizational rigidity and behavioural norms, and the perceptual stumbling blocks that, on the iceberg metaphor, lie below water level.

Equally, patterns of negative behaviour – such as endless debates over sensitive issues – need to be addressed in such a way that they do not affect the emergence of consensus. One possible solution for this is ZMET analysis, which is based on research that people think pictorially rather than verbally.

STEP 5: DEVELOP TAXONOMIES – "Harvest"

Taxonomies, it has been argued are snapshots of consensus. Modelling taxonomies is therefore an ongoing process. Tools exist to model the emergence of taxonomies, though some are still in nascent form. With projects as a frame, there are inevitable time-constraints to developing these.

STEP 6: BRIDGE CONSENSUS & ACTION – “Plough”

Ongoing discussions throughout the project are clearly useful, though, as above, there is a time that work actually needs to be done. Personal Task taxonomies and Group Task Taxonomies offer the potential to bridge consensus and action.

9 References

Albrecht, T. L., & Ropp, V. A. (1984). “Communicating about Innovation in Networks of Three U.S. Organizations”. *Journal of Communication*, 34, 78-91 quoted in Monge P. and Contractor, N. “Emergence of Communication Networks”, *Handbook of Organizational Communication*, Second Edition, ed. Jablin and Putnam (Thousand Oaks, CA: Sage, 2000)

Amazon Website, (2004) <http://www.amazon.com/gp/community/plog/about.html/ref=cm_plog_learn/002-6495458-6328061> [Accessed: June 2004]

Anjewierden, A., Brussee, R. & Efimova, L., (2004), *Shared Conceptualizations in Weblogs,* Presentation given at BlogTalk 2.0, Vienna <http://sharedcon.notlong.com> [Accessed: August 2004]

Back, K.. (1997), *Chaos and complexity: Necessary myths*. In R. A. Eve, S. Horsfall, & M. E. Lee (Eds.), Chaos, complexity, and sociology (pp. 39-51). Sage. p. 43-44

Barabasi, A, (2003) *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*, Plume Books.

Berners-Lee, T., Hendler, J., Lassila, O. (2001), The Semantic Web, Scientific American, May Issue. <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&pageNumber=1&catID=2> [Accessed: Jan 2004]

Boland Jr., R. & Tenkasi, R, (1995), "Perspective Making and Perspective Taking in Communities of Knowing", *Organization Science*, Vol.6, No.4.

The Brains Trust (2004), “Garbled Translation”, *Unreliable Facts from The Brains Trust* < http://www.thebrainstrust.co.uk/unreliable/History/Warfare/>[Accessed: Apr 2004]

Burke, K., (1945), *A Grammar of Motives*, Berkeley: University of California Press, 1969.

Burrell, G., & Morgan, G. (1979), *Sociological paradigms and organizational analysis*, London: Heinemann, quoted in Weick, K., (1995), *Sensemaking in Organisations,* Thousand Oaks: Sage Publications

Ciffolilli, A., (2003), “Phantom authority, self–selective recruitment and retention of members in virtual communities: The case of Wikipedia”, *First Monday*, Volume 8, Number 12. <http://firstmonday.org/issues/issue8_12/ciffolilli/index.html> [Accessed: June 2004]

Cooley, C., (1992), *Human Nature and the Social Order*, Transaction Publishers

Cross, R., & Parker, A., (2004), *The Hidden Power of Social Networks*, Harvard Business School Press, p.120.

Dueck, G. (2001), *Views of Knowledge are Human Views*, IBM Systems Journal, Volume 40, Number 4. < http://www.research.ibm.com/journal/sj/404/dueck.html > [Accessed: Feb 2004]

Drucker, P., (1974), *Management: Tasks, responsibilities, practices*, New York: Harper & Row

Dunbar, R. I. M. (1993). “Coevolution of Neocortical SIZE, Group SIZE and Language in Humans”. *Behavioral and Brain sciences* 16 (4): 681-735.

Feldman, S. and Sherman, C. (2001), The High Cost of Not Finding Information, IDC Whitepaper. < http://www.kmworld.com/publications/magazine/index.cfm?action=readarticle&Article_ID=1725&Publication_ID=108 > [Accessed: Mar 2004]

Festa, P. (2004), *Blog’s the Word in Big Business*, News.com, July 23, 2004 <http://news.com.com/2100-1038_3-5281894.html> [Accessed: July 2004]

Follet, M., (1924), *Creative Experience*, New York: Longmans, Green, quoted in Weick, K., (1995), *Sensemaking in Organisations,* Thousand Oaks: Sage Publications

Gates, W. (2004), Speech given at the Microsoft CEO Summit 2004, Redmond, Washington: May 20, 2004 <http:// http://www.microsoft.com/billgates/speeches/2004/05-20CEOSummit.asp> [Accessed: May 2004]

Gergen, K.J. (1994). *Towards transformation in social knowledge*. 2nd Ed. London: Sage.

Gillmoor, D. (2003) “A new brand of journalism is taking root in South Korea”, *San Jose Mercury News,* May 18. <http://www.siliconvalley.com/mld/siliconvalley/business/columnists/5889390.htm> [Accessed: May 2004]

Gray, D., (2000), *Brain SIZE, Group and Language: Are Human Group Sizes Evolut ionarily Constrained by Our Cognitive Capacity? What is the Right SIZE for a Group?* , Boston Consulting Group . <http://www.bcg.com/strategy_institute_gallery/gorilla2.asp> [Accessed: June 2004]

Harré, R. (1993). *Social being* . Oxford: Blackwell.

Haugaard, M. (1997). *The Constitution of Power.* Manchester University Press: Manchester.

Heidegger, M. (1978), *Being and Time,* Blackwell Publishers.

Jones, W. (2004) “Finders, Keepers, The Present and Future Perfect in Support of Personal Information Management”, *First Monday* , Volume 9, Number 3 <http://www.firstmonday.org/issues/issue9_3/jones/index.html> [Accessed: Apr 2004]

Kauffman, S., (1996), *At Home in the Universe: The Search for Laws of Complexity* , Harmondsworth: Penguin

Kelly, G. (1955). *The Psychology of Personal Constructs* . New York: Norton.

Kelly, G. (1970) *A brief introduction to personal construct theory* . Bannister, D. Ed. “Perspectives in Personal Construct Theory”, pp.1-29. London: Academic Press.

Kirsner, K. (2000), quoted in “UWA Researchers Get Serious About 'Chinese Whispers'”, *Media Statement* , University of Western Australia, May 11, 2000. <http://www.uwa.edu.au/media/statements/2000/05/2000511b.html> [Accessed: Feb 2004]

Krebs, V. (2004a), “Divided We Stand (Political Books – Polarised Readers)”, *Orgnet.com,* <http://www.orgnet.com/divided.html> [Accessed: June 2004]

Krebs, V. (2004b), “Divided We Stand … Still (Political Patterns on the WWW)”, *Orgnet. com,* < http://www.orgnet.com/divided2.htm> [Accessed: June 2004]l

Lanir, Z., Fischoff, B., & Johnson, S. (1988) “Military Risk-taking: C3I and the Cognitive Functions of Boldness in War”*, Journal of Strategic Studies, 12.*

Lewin, R. (1993) *Complexity* , Phoenix.

Lieberman, H. and Selker, T. (2000), *Out of context: Computer systems that adapt to, and learn from, context* , IBM Systems Journal, Volume 39, Numbers 3 & 4. < http://www.research.ibm.com/journal/sj/393/part1/lieberman.html> [Accessed: May 2004]

Lih, A (2004), *Wikipedia as Participatory Journalism: Reliable Sources? Metrics for evaluating collaborative media as a news resource* , Paper for the 5th International Symposium on Online Journalism, University of Texas at Austin, (April 16-17, 2004). < http://journalism.utexas.edu/onlinejournalism/wikipedia.pdf> [Accessed: May 2004]

Lilienfeld, R. (1978),*The Rise of Systems Theory* , Jon Wiley, pp. 191-192

Makkreel, R. (1992), *Dilthey: Philosopher of the Human Studies* , Princeton University Press

Mayfield, R. (2003), “Ecosystem of Networks”, *Ross Mayfield’s Weblog: Markets, Technology and Musings, February 12 < http://radio.weblogs.com/0114726/2003/02/12.html > [Accessed: Jan 2004]*

Mazzocchi, S., (2002) *Enabling Semantic Search ing*, Presentation given at Open Source Content Management Systems Conference 2002- Zurich, Switzerland. <http://www.betaversion.org/~stefano/papers/semantic-searching.ppt> [Accessed: Mar 2004]

McAndrew, F.T., & Milenkovic, M.A. (2002) “Of Tabloids and Family Secrets: The Evolutionary Psychology of Gossip”, 32, 1064-1082.. *Journal of Applied Social Psychology* , 32. <http://faculty.knox.edu/fmcandre/99P0282R3.pdf> [Accessed: Apr 2004]

Mead, G.H., ed. Morris, C.W., (1968), *Mind, Self and Society (Works of Geo rge Herbert Mead)* , University of Chicago Press.

Monge, P. and Contractor, N., (2000), *Emergence of Communication Networks* , Handbook of Organizational Communication, Second Edition, ed. Jablin and Putnam (Thousand Oaks, CA: Sage) <http://www.tec.spcomm.uiuc.edu/nosh/HOCNets.html> [Accessed: June 2004]

Munro, R. (1997) *After reality: Rubbing the talk onto the numbers of the future* . Keele University, Mimeo.

Newman, D., (2004*), Sociology, Exploring the Architecture of Everyday Life* , Sage: Pine Forge Press

Norman, D.A., (1983) "Some Observations on Mental Models," *Mental Models* , D. Gentner and A. L. Stevens, Editors, Lawrence Erlbaum Associates, Hillsdale, NJ.

Perelman, C. & Olbrechts-Tyteca, L. (1969), *The New Rhetoric: A Treatise on Argumentation* , Notre Dame: UND.

Schrage, M., (2004), *The Virtues of Chit Chat* , CIO Magazine, May 15, 2004. <http://www.cio.com/archive/051504/work.html> [Accessed: May 2004]

Shotter, J. (1975). *Images of man in psychological research* . London: Methuen.

Shutz, A. (1964), "The Stranger: An Essay in Social Psychology", *Collective Papers* , Vol. 2, The Hague, Netherlands: Martinus Nijhoff.

Sieloff, C. (1999), "Is Knowledge Draining Our Attention?", *Knowledge Management Magazine,* July issue*,* <http://www.destinationkm.com/articles/default.asp?ArticleID=113> [Accessed: Feb 2004]

Sillince, J.A.A, (1992), "Literature searching with unclear objectives: a new approach using argumentation"*. On-line Review* , 16 (6), 391-409.

Snowden, D., (2004) Proceedings of the Association of Knowledge Workers, Post 1795 <http://finance.groups.yahoo.com/group/AOK_K-Net/message/1795> [Accessed: June 2004]

Stacey, R.D*.* (1996),*Complexity and Creativity in Organizations* , Berrett-Koehler.

Stenner, P. (1998). Heidegger and questioning concerning psychology. *Theory & Psych ology*. *8*(1)

Stevenson, W.B., & Gilly, M.C. (1991). "Information processing and problem solving: The migration of problems through formal positions and networks of ties". *Academy of Management Journal* , 34*,* 918-928.

Varela, F, Thompson, E., & Rosch, E. (1991), *The Embodied Mind: Cognitive Science and Human Experience.* Cambridge: MIT Press

Virginian-Pilot (Norfolk, VA), June 11, 2004, (quot. Wikipedia Press Coverage, English, June 2004 at http://en.wikipedia.org/wiki/Wikipedia:Press_coverage). The edited Chesapeake entry is at *< http://en.wikipedia.org/w/wiki.phtml?title=Chesapeake%2C_Virginia&diff=3860769&oldid=34 14810>* [Accessed: June 2004]

Warnecke, H.J.. (1993), *The Fractal Company: A Revolution In Corporate Culture* , Springer-Verlag.

Weber, M. (1947). *The Theory Of Social And Economic Organization* (A.H. Henderson & T. Parsons, Eds. & Trans.). Glencoe, IL: Free Press.

Weick, K., (1995), *Sensemaking i n Organisations,* Thousand Oaks: Sage Publications

Welsch, W.(1997) Undoing Aesthetics London: Sage

Wheatley, M.J., (1992), *Leadership and the new science: Learning about organization from an orderly universe* , Berrett-Koehler.

Wilson, T.D., Ford, N., Foster, A., Ellis, D., and Spink, A. (2002), *"* Information seeking and mediated searching. Part 4: cognitive styles in information seeking", *Journal of the American Society for Information Science and Technology* , Volume 53 Issue 9

Winer, D. (2002), *The History of Weblogs,* <http://newhome.weblogs.com/historyOfWeblogs> [Accessed: May 2004]

Winograd, T. & Flores, F. (1995), *Understanding Compu ters and Cognition: A New Foundation for Design.* Addison Wesley

Commercial Blogs:

Microsoft Community Blogs (2004) –

<http://www.microsoft.com/communities/blogs/PortalHome.mspx> [Accessed: June 2004]

Sun Bloggers (2004)

<http://blogs.sun.com/roller/> [Accessed: June 2004]

Jonathan Schwartz's Weblog (2004)

<http://blogs.sun.com/jonathan> [Accessed: Aug 2004]

IBM's developerWorks Blogs (2004)

< http://www-106.ibm.com/developerworks/blogs/> [Accessed: June 2004]

Dell Linux Blog (2004)

<http://linux.dell.com/blog/> [Accessed: June 2004]

Jupiter Research Analyst Weblogs (2004)

http://weblogs.jupiterresearch.com/

TWiki Success Stories:

TWiki BT (2004)

<http://twiki.org/cgi-in/view/Main/TwikiSuccessStoryOfBritishTelecommunications> [Accessed: May 2004]

TWiki Disney (2004)

<http://twiki.org/cgi-bin/view/Main/TwikiSuccessStoryOfDisney> [Accessed: June 2004]

TWiki Motorola (2004)

<http://twiki.org/cgi-bin/view/Main/TwikiSuccessStoryOfMotorola> [Accessed: June 2004]

TWiki SAP (2004)

<http://twiki.org/cgi-bin/view/Main/TWikiSuccessStoryOfSAP> [Accessed: June 2004]