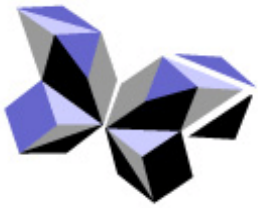

Digital Business Ecosystem

Contract n° 507953

Workpackage 34:
Exploitation & Sustainability

Deliverables D34.4.3:

Market Watch - The Social and Semantic Web

Project funded by the European Community under the “Information Society Technology” Programme

Contract Number: 507953

Project Acronym: DBE

Title: Digital Business Ecosystem

Deliverable N°: D34.4.3

Due dates: 12/2006

Delivery Date: 03/2007

Short Description:

Author: IBM

Partners contributed: IBM

Made available to: Public

Versioning

Version	Date	Author, Organisation
0.1	31/01/2007	Peter Stanbridge, Korora Ltd
0.2	31/01/2007	Tim Romberg, FZI

Quality check:

1st Internal Reviewer : N/A

2nd Internal Reviewer: N/A

The Social and Semantic Web

Market Watch

Abstract. This market watch looks at the current trends in the areas of the Semantic Web (which was a well-known concept at the beginning of the DBE project) and the new Social Web and Web 2.0, which have recently received attention and its implication for the DBE.

1 Introduction

There has been considerable attention given to the semantic aspects of the DBE, both from within the official tasks and deliverables and within a wider discussion amongst members of the different work streams, including a Blog that has included discussion in this area.[1]

In this market watch, the experiences of working with service discovery requirements will be explored within the context of Web 2.0, the social web and associated developments. The first part of the report suggests that a more delineated demarcation between service discovery and service execution is needed. This may seem a little disappointing given an early commitment to provide a uniform semantics to describe all aspects of the digital eco-system[2], but it is felt that (certainly given present technology capabilities) the different areas requiring semantic description are sufficiently different to warrant alternative technology paradigms.

The report then discusses the developments of Web 2.0 and social tagging and contrasts this as a movement with the semantic web. While many see these as two completely opposing forces being applied to the same application or problem domain, we see that both perspectives have insights to offer business service description. However, the emphasis here will be a promotion of social construction as a more suitable paradigm for describing businesses within the context of a digital business eco-system, but constrained by appropriate structure and technology support. That is, we suggest that the requirements of business description, service description and discovery are completely different to the requirements of software service aggregation and integration designed to support the business service execution. The relationship between these two modelling paradigms still remains elusive and we suggest that this aspect of the problem remains one of the most important research agendas in this area.

[1] http://opensoa.blogspot.com/2006/05/bml-treemaps.html

[2] Not withstanding the fact that the BML as delivered doesn't provide a uniform modelling capability.

2 Basic concepts

2.1 B2B automation and integration

Business-to-Business (B2B) applications, such as DBE, support interactions between business partners by making them more efficient or even automating them. Typical modern B2B applications enable some form of machine-to-machine communication (in contrast with earlier applications which only provided a human-readable web interface between business partners).

Such applications therefore require common data structures, encodings, protocols and so on. It must be clear to the engineers of the applications involved which meaning each field of a data structure (or a similar element) has, or at least how the application which he or she is responsible for should treat it.

Semantics and Pragmatics in B2B automation

The first aspect, namely the association between a syntactic element (sign) and either an object or phenomenon of the real world or a shared mental concept (often both) is called the *Semantics*. The second aspect, the consideration of how a sign influences the action of agents involved is called the *Pragmatics* (following [Morris 1938]).

2.2 Service-oriented architecture

A service-oriented architecture (SOA) is a particular, currently popular architectural style for implementing both internal and external application integration. At the core, there is the concept of a software service provided by a server or provider. As an example, a warehouse application may provide a software service to other applications that need to know about stock levels, or make orders from the warehouse. These other applications are called the clients. For our discussion, it suffices to remember that SOA is always about either company-internal or B2B automation (i.e. machine-to-machine communication), and that the software services referred to do not necessarily correspond to a “business service” as an economic good or legal obligation. There is frequent confusion regarding this point because especially during the introduction of “Web services” and the associated SOAP standard suite, they were sold by technology providers primarily as a way of generating new business; or even as the way all business would be conducted in the future. Nowadays service-oriented architecture is most often found inside one company as an Enterprise Application Integration architecture. But B2B applications also exist. But even then, it is important to distinguish the business level and the software level. In fact, it is not uncommon to find the client in a business relationship act as the provider of a software service to its supplier.[3] A specific service-oriented architecture (with additional elements of a Peer-to-Peer architecture) is also implemented as the DBE’s Execution Environment.

[3] Introductions to service-oriented architecture are provided by the Oasis consortium (SOA Reference Model, cf. http://www.oasis-open.org/), IBM (http://www.ibm.com/soa) and many others.

2.3 Business (service) discovery

The terms *business discovery* and *service discovery* emerged in the context of service-oriented architectures. They underline the loose coupling in such an architecture in that the (software service) client has the possibility to look up a suitable software service provider on the fly, when the need arises.

This was transferred from the level of software services to the level of business and the term “business discovery” therefore refers to scenarios where a suitable business partner registers with a directory of some sort, and then clients can look up suitable business partners quickly and at low cost. Of course, there are many established, traditional means of identifying new business partners in each industry – The Yellow Pages, Trade Shows, craftsmen of one trade grouped together in the same street, etc. – but the term “discovery” is especially associated with the new electronic directories made possible by the Internet. The Internet, by lowering the costs for establishing new business relationships, and by making markets more transparent, is often expected to lead to a higher percentage of business being conducted *ad-hoc*. On the other hand, however, the Internet is an anonymous medium with many possibilities of identity theft and fraud which raises the importance of brand and reputation. Also, ad-hoc business usually has to rely on some standard way of doing things, hard to change once established; while long-term partnerships can make investments into tighter process integration and innovation.

Typical examples of ad-hoc scenarios are:

- Sending flowers to someone in a different city. This usually works by the aggregator or directory looking up a local flower shop who can perform the delivery;
- Ordering a taxi (in cities where each taxi is in fact an independent business)

We would not speak of an ad-hoc or “business discovery” scenario when we are in fact dealing with many other actors in an anonymous market place, e.g. the stock market. There, interactions are even more standardized, even regulated, and we may not ever know who we sold to or bought from.

Today, parallel to what has happened to the other “Web service” standards, we can see that the “service discovery” technologies developed a few years ago (e.g. UDDI) have found their main application area inside a corporations application integration landscape (the new fashionable term being the *Enterprise Service Bus*); while the actual “business discovery” happens in very specific ways depending on each industry.

2.4 The Semantic Web

The Semantic Web is an initiative by the World Wide Web Consortium, specifically its founder Tim Berners-Lee, and supported by a community with a strong Artificial Intelligence research focus. The original vision, as described in [Berners-Lee et al 2001], pursues the following objectives:

- To improve the current web architecture (specifically, a better separation between human-readable presentation and machine-readable information)
- Enable interoperating applications (described in a health care scenario)
- Enable automated, intelligent combination of distributed information (reasoning)

In order to achieve these objectives, a set of concepts and technologies is proposed. Many come from the AI discipline of knowledge representation, but are designed in a more scalable way:

- Global identities for things in the real world: URIs
- Organize (shared) concepts and relationships: Ontologies
- Describe Ontologies and Information (metadata) in standardized, structured form: RDF (Resource Description Framework), OWL
- Store and access everything in a decentralized way, just like the WWW
- Derive new knowledge from existing knowledge: Reasoning based on formal logic
- Decentralized processing on behalf of a user: Agents

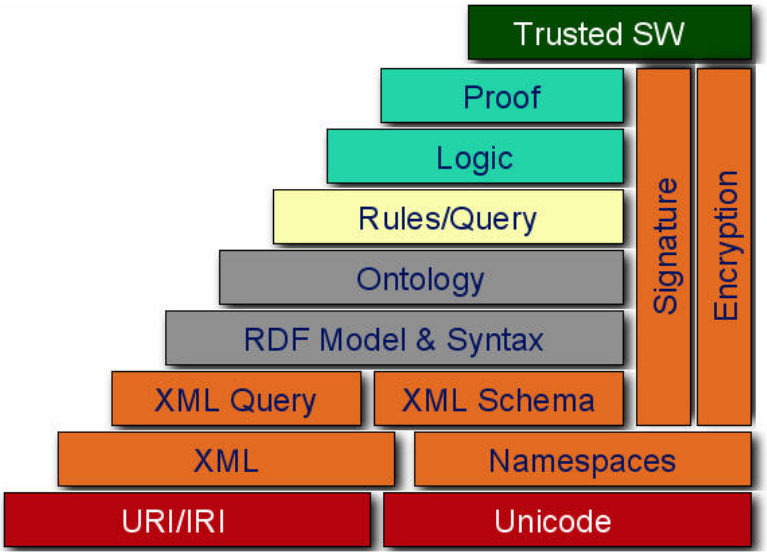

Figure 1: The Semantic Web stack (W3C)

An important question for the subsequent discussion and for contrasting the “Semantic” to the “Social” web is: Who delivers the ontologies, structures, etc. that are so important for this approach? It is difficult to pinpoint this in the vision described by [Berners Lee et al. 2001]; they are somehow assumed to be “there already”. When looking at the individual scenarios, one can hypothesize that likely sources of Semantics are:

- application software developers who engineer their software to act as RDF sources and sinks
- industry consortia, e.g. in the Health industry, who establish standards, e.g. for describing patient records, classify diagnoses, operations and procedures, etc.
- software integrators who need to bridge between different applications and standards.

This could be summarized as an “expert-based approach”. Not only are domain experts required, but also experts in the concepts and technologies of the Semantic web.

2.5 The role of XML

An important, currently popular alternative for some of the goals mentioned above is XML, the Extensible Markup Language. XML is also standardized by the W3C, but was pushed for by a different group of people, notably Microsoft and IBM. We assume that the reader is familiar with its basic concepts. When exchanging data between a well-known set of applications, it is often sufficient and more accessible to most programmers to define

standardized document formats, e.g. using XML schema, than formalizing the concepts referred to in an ontology. XML's advantages over the binary or textual document formats of the previous generation are:

- it is easier to make applications tolerant to slight changes in the format definition;
- there is a large set of tools and libraries available for editing, querying and transforming XML documents
- the data is self-documenting to a limited extent for a human reader.

Compared to XML, Semantic Web technologies have had a difficult time in that they impose higher costs for the development of a given application, but their benefits – e.g. being able to combine the data with data from other applications – are often external. Today, Semantic Web technologies are often found either outside the context of the entire stack in Figure 1 (RDF being frequently used without reasoning, for example), or isolated from the public Web in very specific application domains (like text mining of medical databases).

A given XML-based document format may have very precise *Semantics*, in the sense that the community using this format agrees precisely on the concepts associated with each element. In XML, one simply does not strive to formalize the semantics; they are often described in plain text in an accompanying document. Of course, this description can contain ambiguities, and they may never be resolved until there is a *performative* problem of one application incorrectly dealing with data produced by another application. We can therefore say that an XML-based standard is usually associated with a well-defined *pragmatic* context – the set of applications using it.

Most current technologies for implementing service-oriented architectures are on the same conceptual level as XML; i.e. the technical, formal part of a service definition often consists of the specification of message formats, using XML Schema or similar means (often, the power of XML Schema is restricted for this purpose).

2.6 Semantic Web technologies in the DBE infrastructure

Semantic Web technologies are used for the purpose of business modelling in DBE. The objective of business modelling is for individual businesses to be able to describe themselves based on a formal model of their industry and general characteristics (such as geography). In the currently developed concepts, the business models are exclusively used for business discovery. Based on the models, very precise searches can be formulated to identify matching business partners or businesses who face similar challenges as oneself. The Business Modelling Language (BML) enables the creation of these formal industry models. For example, one can express in this language that seats on a certain type of train can be either first or second class, and either smoking or non-smoking.

The formal business models can refer to concepts of a formal ontology, which can be imported into the DBE Studio, and the run-time Knowledge base. The main purpose of ontologies is to define some either industry-specific or globally relevant concepts, and provide global identifiers for the concept instances. For example, without ontologies, one train service may define the "seat class" attribute as a numerical value of either "1" or "2", while another train service introduces it as a string value of either "A" or "B". (There may be good reasons why the two train companies cannot use the same BML model.) The

Ontology would be a place for defining a common concept of “seat class” with two or more well-identified instances.

The DBE infrastructure also provides the so-called Execution environment, at whose core is a framework for implementing software services as a means for B2B automation; there is also a User Interface framework to support human access to these services. These software services are defined using a service description language (SDL), on the same conceptual level as XML Schema.

2.7 The Social Web and Web 2.0

The services offered on the World Wide Web have evolved in recent years, even though the Semantic Web as originally envisioned has not played a major role in this evolution. The common traits of the development are slowly becoming clear, and have been given the name “Web 2.0” by publisher O’Reilly media for a series of conferences. [O’Reilly 2005] summarizes the “core principles” of Web 2.0 applications as follows:

- considering the Web as the platform (rather than a particular Operating System, for example)
- giving the user control over his/her data
- services instead of packaged software
- creating an architecture of participation
- cost-effective scalability
- providing a remixable data source and data transformation
- providing software above the level of a single device
- harnessing collective intelligence.

These principles cover a vast range of applications including P2P file sharing networks and advertisers. In this article, we would like to concentrate on a narrower category of new applications which are characterized by user-provided, user-structured, and often remixable content, such as:

- *Wikis* – structured by hyperlinking between well-identified topics
- *Blogs* – structured by chronology, discussion threads, inter-blog linking, syndication and aggregation (e.g. by Technorati)
- *Social tagging or bookmarking* sites such as Del.icio.us, Flickr, CiteULike, or last.fm
- *Collaborative databases* such as Open Record, Google Base

They represent a new kind of *Social Software* [Hammond 2005], differing from more classic *groupware* (such as mailing lists) in a few respects: They are Web-based, provide access to every interested participant by default, leverage network effects to benefit from a large global user base, and allow each user to participate in the structuring of the content, for example by creating new tags, new hyperlinks, new database schemas or new aggregated feeds, where this would have traditionally been an administrator’s task. They are also very different from the Semantic Web concepts: There is no systematic engineering of common structures or ontologies; structures only emerge in a kind of fuzzy way by groups of people using common tags; the only reasoning offered is one based on statistics and correlations, rather than discrete logic.

Simply said, the Semantic Web emphasized the physical decentralization of data sources while requiring centralized (standardized) definition of concepts which are referred to in the data. The Social Web allows the decentralized (individual) definition of concepts; it adds value to the data by centralizing its storage, and aggregating it.

3 Social construction

The social web or web 2.0 has not come into prominence by accident – there is an intellectual context in which the idea of *social construction* has become acceptable. There has been considerable interest in social construction within certain traditions in philosophy, social anthropology, sociology, biology, cognitive science and other disciplines.

In this section some of the areas of influence contributing to the acceptance of the social web are discussed. The purpose of containing such material is to motivate the suggestions of the web 2.0 community, primarily to suggest that it isn't a reactionary fad, but a serious option.

The basic premise behind the selection of contextual material for the web 2.0 is the insight that the imposition of rigid or strict syntactical and semantic structures to modelling a business for the purposes of discovery, presentation, representation to an eco-system of organizations and their people does not meet the needs and uses for which such representations will be put. One danger of such an imposition is that users will ignore the technology, but this could result if further problems such as the inability to genuinely incubate an eco-system phenomenon.

It is clear from some of the earlier scientific work in the DBE that the concept of emergence plays an important role in the development of eco-systems. Emergence is a high order phenomenon and shares some interesting properties with complex systems from which emergence frequently results. In considering the properties of emergence human factors are very important because the complex systems involved are not entirely the result of blind interacting parts (as are many naturally occurring complex systems) but form from the interaction of people with their desires, unpredictable interventions, multiple roles and even self-destructive tendencies (as some ecologists or conservationists might argue).

Dave Snowden in [Snowden 2005] has described an interesting scenario that illustrates some of the features of what he calls social complexity (in comparison to scientific complexity).

> *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? Would those objectives be 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 motivation 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 pocket money is linked to achievement of the measures to each milestone? Would you conduct an after action review at the end of the party, update your best practice database and revise the standard operating procedures for party management?*

> *No, instead like most parents you would create barriers to prevent certain types of behaviour, you would use attractors (party games, football, a video tape) to encourage their formation of beneficial largely self organising identities; you would disrupt negative patterns early, to prevent the party becoming chaotic, or necessitating the draconian imposition of authority. At the end of the party you would know whether it had been a success, but you could not define (in other than the most general terms)what that success would look like in advance.*

In this paper, Snowden describes what he calls multiple-ontology sense-making where different ontological and epistemological perspectives have relevance within their restricted domain. The basic argument runs that if an approach, method, process or management fad overruns its domain boundaries, then failure of applicability typically follows. Examples sited are business process reengineering (which was taken from its legitimate bounds of manufacturing within organisational silos and taken into the whole organisation) learning organisations (Senge 1990) and emotional intelligence, knowledge management and many others.

> *Multi-ontology sense making argues that different approaches are legitimate, but within boundaries and that methods and tools that work in one ontology, do not work in another. It is thus behoven on management to know which ontological domain they are operating in, and what transitions between domains they wish to achieve.* [Snowden 2005]

The ontological and epistemological model demonstrated in [Snowden 2005] is shown in Figure 2.[4]

The two rows represent two fundamentally different ontologies – order and un-order – the first story of the childrens' party representing order – the type of approach one would expect to adopt is working within an ordered system. This is fine if cause and effect are easily identified and controllable. Outcomes can be predicated and controlled. In the un-order ontology it is not possible to predict either cause or effect and relationships between them do not repeat.

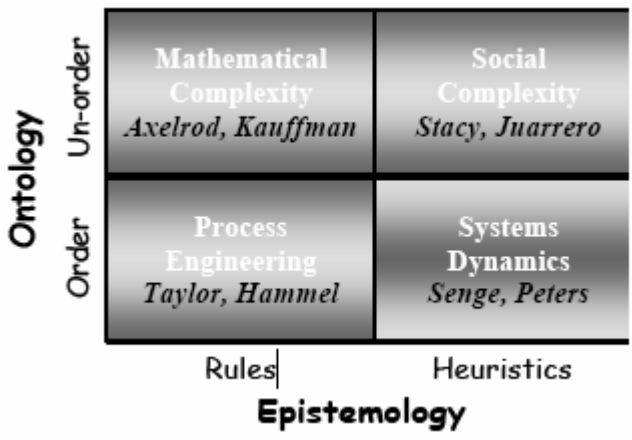

Figure 2 - Ontology and epistemology model

[4] But developed during a project produced by [Snowden, Shelton, Sage and Stanbridge] for the European Commission on knowledge management , but further described in [Snowden 2003] and [Stanbridge & Snowden 2004]

The two columns really represent

> *a continuum between low ambiguity rules that can easily be made explicit and the more ambiguous use of heuristics or rules of thumb which provide guiding principles but have high levels of ambiguity* [Snowden 2005]

Inside the blocks are representatives of typical management perspectives in each area including some key thinkers involved in each. These are reasonably well understood positions, the novelty in placing them within the perspective level model.

The important point to note for the purposes of organisational description and service discovery within the remit of the DBE is that the modelling efforts associated with the semantic web and the OMG MDA model are designed to work in the process engineering area (and in some cases the mathematical complexity area). Much discussion around the complex and non-linear aspects of eco-systems (within and outside the DBE) has focussed in the mathematical complexity area. But the business eco-system is a human system and shares many of the aspects of social complexity addressed by such movements as the social web. Social complexity shares the idea with mathematical complexity that in this domain the concept of un-order and emergence is relevant, but it shares with the system perspective that human systems are different. Some of these differences are discussed in [Snowden 2005] and will be briefly described here.

Humans make decisions based on patterns[5]. One of the key sources for this theory within organisational management is [Klein 1994] (as well as a host of work within the cognitive sciences – see [Rockwell 2005] and [Noë 2004] for a more philosophical perspective on this). People frequently rationalise their instinctive based decisions only in retrospective reflection. It is therefore rather surprising that we are asking them to dedicate such intellectual effort to model their organisations within the context of the active and complex business games and interactions that actually take place in the real business-world eco-system.

Humans create multiple identities. We take on a multitude of roles within the context of our various circumstances and activities. Much behaviour is context dependent and it is often impossible to take experiences and behaviour from one context to another, as is illustrated by the profound surprise many family members (even wives) of the 7/7 bombers felt when realising that their relative had completed a suicide bombing raid in London. The implications for the semantic descriptions within the DBE are that within the context of an organisation there really isn't "the" or a single "a" in the descriptions.[6]

Humans ascribe intentionality and cause even where none necessarily exist. This is a type of corrective to the previous two points. One of the key insights in social complexity is that "some things just "are" by virtue of multiple interactions over time" [Snowden 2005] but that we frequently wish to ascribe intentional causes to such things – we are always looking for causes. We need to take this insight into account as we wish to identify reasons

[5] While we think this is a genuine and important insight, it does not deny representational and logical thinking either, which is also an empirical fact. But the extensive research undertaken by Klein clearly indicates that in many critical situations practiced based expertise based on intuitions tend to not only rule much decision making, but when not hindered by over intellectualisation is typically considerably faster and superior to considered analytical approaches. But again one must be careful with context and circumstances.

[6] Also see [Berger & Luckmann 1969] pg 103, 104 for a sociological perspective (they talk about sub-universes).

and causes for all the different interactions that take place amongst SMEs in a business eco-system where for many it may simply be chance – a "just there at the time/place". It may also indicate that many of the models created using social network analysis may indicate constructed relationships, reasons and causes created in retrospect that may have no grounding in reality.[7]

The critical point from the above is to what extent does the semantic features of the DBE support this type of typical phenomena that is the real eco-system behind the digital business eco-system?

Humans have learnt how to structure their social interactions to create order
Humans, via social structures, myths, narrative, story, ritual, taboo etc. create stability and predictability in their systems, which also enable us to plan for the future. "Humans have learnt how to move between order and un-order". The implications for the DBE semantics is that in some cases, the social web is too chaotic. Experiences in Flickr for example, indicate the difficulty of finding photos posted of a specific event say and the question must be often asked about all the wonderful photographs loaded into Flickr but not categorised using the most popular language. Without some structure, the application becomes unusable.

[Clark 1996] pages 74 – 77 describes direct and indirect emergence. An example of direct emergence is the formation of traffic jams resulting from simulation of two rules – accelerate if there is space ahead and slow down when approaching another car starting with cars distributed randomly.[8] Traffic jams, which are a higher order phenomenon than the cars that constitute it form and whereas traffic moves forward, the jams move backward. The example given of indirect emergence is the nest building behaviour of some termites. When they construct arches from mud balls they initially do so randomly but each time they drop a ball they deposit a scent on the ball which is attractive to them and other termites – termites prefer to drop their mud balls close to a strong scent. As the mud balls are deposited the scent increases and multiple piles become enclosed as the scent from different piles becomes strongly mixed. The natural outcome is the arch shapes common to these termite homes. [Weissman 2000] Chapter 6 provides an interesting discussion on the ontological status of such complex systems, in particular noticing that the factors contributing to the stability of such systems reside in the relationships and not merely in the parts, thus indicating the importance of taking higher order phenomenon seriously.

But what are the implications for social systems? A simple example is suggested by Clark. If one was to construct a new university campus containing several buildings serving different functions and one was to consider the question of where to place footpaths to aid people walking through the campus: where would one place them? A solution sympathetic to emergence and complex systems would initially cover the whole campus in grass and planners would watch where the worn patches in the grass occur as indications of the best place to build the paths. An approach similar to this (luckily for drivers the approach was not identical) was undertaken in Swindon in the UK, which is home to an extremely sophisticated roundabout whose pathways were "constructed" via experimentation and not design based on expected traffic patterns. In a sense, the Swindon roundabout solution

[7] Hence this point being a corrective to the previous two, where the emphasis is on the vitality of social construction and the intuition involved in decision making.

[8] See also [Weissman 2000] chapter 1, [Clayton 2004] chapter 4 and many papers and texts covering complexity sciences.

indicates that new approaches developed via complexity and emergence do not always compete with traditional methods of engineering design (carefully planned scientific techniques were still required in the construction phases, for example) but certainly complement them.

Philip Ball in his book "Critical Mass: How One Thing Leads to Another", in a section on networks, states "Thanks to the new physics of networks, we can now see that the structure of the Internet reflects the ethos that produced it" [Ball 2004] p. 470, and quoting the sociologist Manuel Castells: "The culture of the producers shaped the medium"[9]. Ball elaborates with "That is to say, only a network that grows 'organically' accounting to no master plan, observing principles of free access and meritocratic choice will develop the kind of architecture that the Internet and the WWW display – with its strengths and pitfalls". In another chapter, on utopia, Ball warns against applying metaphors from physics into social systems. He targets his critique on William Petty, the Irish rationalist politician. Petty suggested that the labouring classes of Ireland could afford a tax rise by missing out on their Friday dinner. But as Ball states, "only someone deeply lacking in an understanding of the circumstances and psychology of his fellows would venture to make such a serious proposal" [Ball 2004] page 572. We should adopt the same warnings against overindulgence on biology and the mathematical complexity science (see above section on the two ontology model). While Ball's example of Petty is rather extreme within the context of the DBE, nevertheless, such a scientific oriented project can easily miss the point as to what people are really doing in their work environment and to miss the common motivational factors that inspire many small business owners in their work and work relationships.

An interesting critique of the type of correspondence thinking (in terms of correspondence theory of truth) is made by Jerome Bruner in "Acts of Meaning" [Bruner 1990]. In quoting and explaining Richard Rorty he makes the point that once one moves beyond simple statements such as 'the cat is on the mat' into statements with universals, hypotheticals or theories[10] the pairing of linguistic unit with the world breaks down. We think that there is a point to Rorty's pragmatism in certain contexts – in what Rorty calls 'pragmatic perspectival questions'. So, for example, statements such as 'history is the story of class struggle' are not to be judged by limiting oneself to questions like 'does that assertion get it right?'[11] To Rorty, questions such as 'What would it be like to believe that?' or 'What would I be committing myself to if I believed that?' are more appropriate. These types of questions are far removed from the typical Kantian essentialism which always asks for the real essence in things. Such pragmatism always questions the "*it*" in statements that contain it and similar words that assume a single ontology.

This commitment to essentialism is reflected in semantic models that assume that businesses have *a* description, *a* definitive structure to pricing, process descriptions and so on.[12] One serious (and to us sometimes justified) criticism of pragmatism is that it leads

[9] [Ball 2004] page 471 quoting M. Castells (2001) *The Internet Galaxy: Reflections on the Internet, Business, and Society*, p 207, Oxford University Press.

[10] And I would add counterfactuals and complex demonstratives and so on.

[11] Quoted in [Bruner 1990] page 26

[12] Sadly of course, there exists a management mentality in many circles that assume just that. One can think of extreme forms of process reengineering or even balanced score card methods that promote diversity in measurement but still assume a degree of singularity of perspective in a set of quantitative measures.

to absolute relativism. We consider that at least within the context of the social web, such pragmatism does not lead to an 'anything goes' mentality. "To ask the question 'how does this affect my view of the world or my commitments to it?'"[13] is not a commitment to complete relativism. Bruner quotes James Clifford *The Predicament of Culture* at this point, where Clifford notes 'That cultures, if they ever were homogeneoius, are no longer so, and that the study of anthropology perforce becomes an instrument in the management of diversity'[14].

We consider these comments to be of vital importance to a consideration of the semantic descriptions of organisations when made within the context of the social (which includes business) interactions engendered in trading and commerce. There is here a basis to facilitate multiple perspectives, to admit the difficulty of expressing an organisation through a modelling exercise that requires a degree of 'completeness' *prior to* the organisation interacting in the community. It certainly brings into question the type of representationalism quite explicit in our opinion in RDF specifications, which follow a subject/predicate model that comes from traditional logic and enlightenment thinking.

In any way, even if one were to acknowledge the existence of *a* science of reality (say ultimately in mathematical physics) we are interested in pursuing a sociology of our knowledge of reality because such realities we are addressing (business eco-systems) are created by the business people whose businesses they have created and the reality they have created by the engagements they have made.[15] From a social psychology perspective, [Hosking & Morely 1991] make similar points and provide interesting support for the idea of multiple ontologies advocated by Snowden. In a section on constructivist perspectives on networks they argue that the commercial world has now created a large number of different contexts, each involving networks and inter-networks of influences and communication. There is an emphasis in understanding the political nature of how people work and interact in the network and that their realities are frequently influencing the nature and shape of the network and the network is shaping their realities. There work also stresses the importance of power and power influences – that some people in certain power positions have more influence than others[16] but this is often because such people possess immense talent at networking through the network with an acute sense of the different worlds of discourse within them. They are able to achieve influence by their ability to exploit equivocal meanings in discourse. But they also recognise the influence of others through their interaction.

This political dimension is important to the DBE, in particular, to recognise that the multiple-universes, the complexity of the environment, the degree of social interaction, power games and politics is the world of multiple worlds. Our semantic descriptions must be capable of adapting and adopting to such worlds and most of all, the evolutionary environment must itself not be constrained by narrow definitions of the world we want to

[13] [Bruner 1990] page 27

[14] [Bruner 1990] page 27

[15] This point is made by [Berger & Luckmann 1969] e.g. pg 14, 15. Also see the following treatments attempting to understand and resolve the bifurcation between more analytical and continental styles of understanding and the so called 'Science wars' (these treatments are excellent because while each different they see the importance of a meaningful synthesis of approach). [Smith 2005], [Hung 2006] and [Norris 2000]

[16] We will add that anecdotal evidence suggests that people in power may not always have official positions of power, but this point is not important for the discussion here.

create. If the system developers define their world, especially if it is a narrowly rationalistic world, then the world will remain predominantly empty.

Our final selection of background works is Ralph Stacey, who is a management professor in the business school of the University of Hertfordshire in the UK. In his [Stacey 2001] introduction to part II he outlines a social conception of the human mind, but states clearly that his view is neither individualistic nor totally holistic. “It is a view in which individual minds and social relationships arise together, simultaneously”. It is a view of the individual mind in which neither the individual nor the social are prior or primary. This is the type of interaction one would expect when people within organisations come together to engage in business relationships, especially where knowledge creation and transfer are involved. It is not possible to support such interactions via a technical medium unless it is understood that knowledge is created within this context through the relating engagement – knowledge is not something that sits within the organisations independent of their web or network of relating.

In conclusion of this section: We have raised a number of different perspectives on the subject of social construction and recognise that given the reality of diversity, emergence, complexity and social interaction, there is a tremendous pressure to ensure that our supporting semantic meta-models, models and software artefacts support and not hinder the nature of reality we expect to utilize it. It is our opinion that the movement toward the social web and web 2.0 is an important direction in which future projects working in this area should travel. We are sceptical of our own achievements within the DBE in this area but accept the challenge that our own lessons and that of the web community in general are learning in this area too.

4 Examples

4.1 Flickr – an example of the social web

Flickr[17] is a hugely successful photo repository made available to both professional and amateur photographers. As of February 2007, roughly 400 million photos have been uploaded.[18] Photographs are described using keywords and then made available to either the public or made private. Private photographs are tagged as available to friends or family.[19] It is widely recognised as a site promoting principles of the social web. It provides facilities for users to rate photographs as well as to make comments. The site distinguishes professional photographers from non-professionals and provides facilities to organise your own photographs and create favourites from photographs within the site.

As one can see from Figure 3 a number of useful options indicate the community nature of the site. Key ones are groups, photos from friends, popular tags, creative commons, and search.

[17] www.flickr.com

[18] the exact number is not published, but Flickr numbers all uploaded pictures consecutively, and the most recent uploads are assigned numbers slightly above this figure. This number does not account for deleted photos or duplicates.

[19] There are a lot of photo sharing sites on the internet – for an interesting discussion on the differences and debates about who is in the lead see http://gigaom.com/2006/06/22/photobucket-rules/ but these discussions are not central to what is being discussed here.

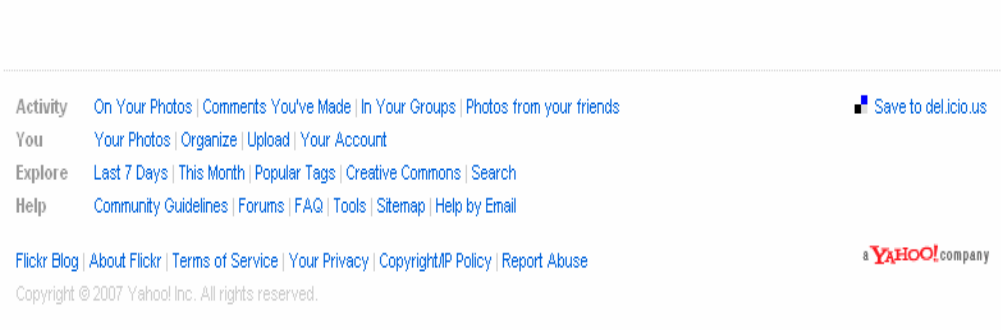

Figure 3 - Options at the bottom of the page

In Figure 4 below, options from the top of the page are shown. Here, a contact has been created in order to keep up to date with photographs added by aremac.

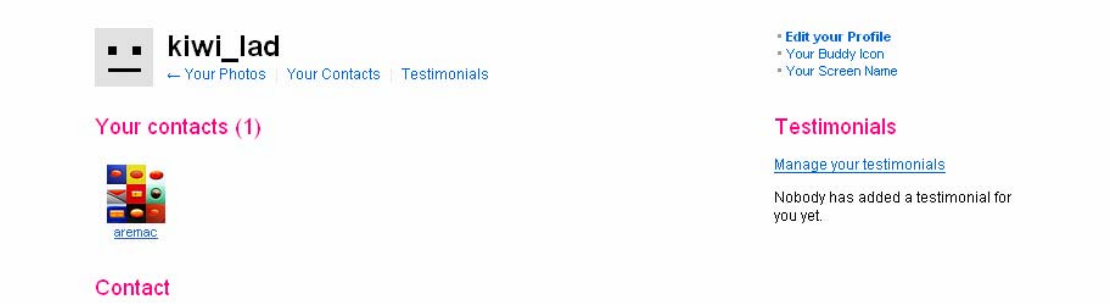

Figure 4 - Options on the top of the page

In Figure 5 below the favourites option is selected with some of the favourites displayed. Unfortunately no method currently exists to enable one to structure favourites – they sit in to a single bucket. As can be seen, one can display one's favourites as a slide show.

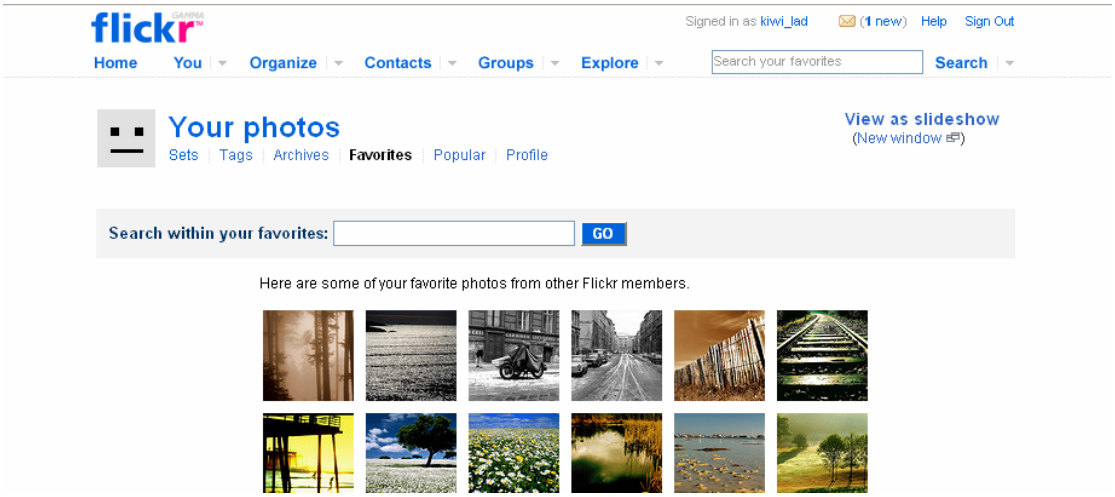

Figure 5 - Favourites

The explore option is very important – it contains a whole host of different access points into the photographs. We show the popular tags option in Figure 6 below, but different methods include by geography (via a map of the world), dates, most recent, and groups that one can join around a subject, style, camera types and so on (groups are the invention of the users). The tag size and boldness indicate the popularity size of the tag. Tags are created by users, they are not pre-determined; hence the tags here show a type of social construction. Sadly there is no way we know of to get to all the tags or to use search options to further restrict the returned photos except by the "clusters" option shown

in Figure 7 and Figure 8. One would expect that the clusters would be excellent because they go deeper than one level, but in experimenting, it would seem as though the second level clusters loose the original subject tag. For example, if you go “Birthdays” one will see “Cake” as a cluster. On selecting “Cake” one would expect birthday cakes, but one gets all cakes. This type of problem is fixable – it is not an inherent problem with social tagging, but it indicates the importance of using technology to provide the expected results or better, to enable users to alter the way the results work.

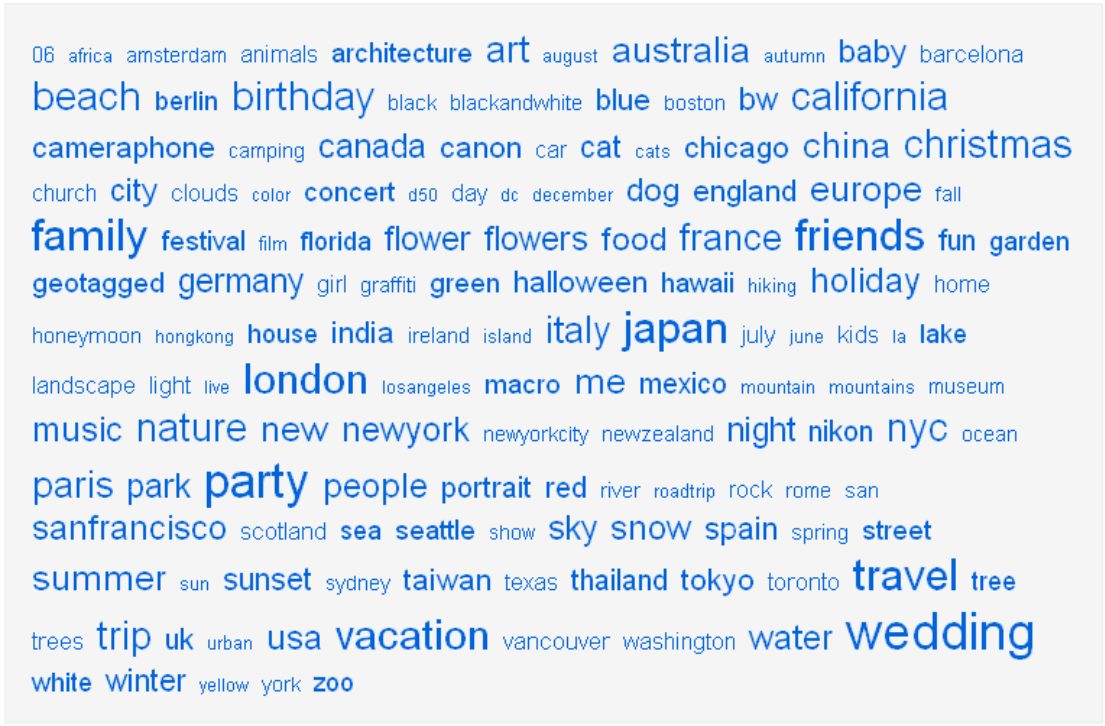

Figure 6 - The important most popular tags

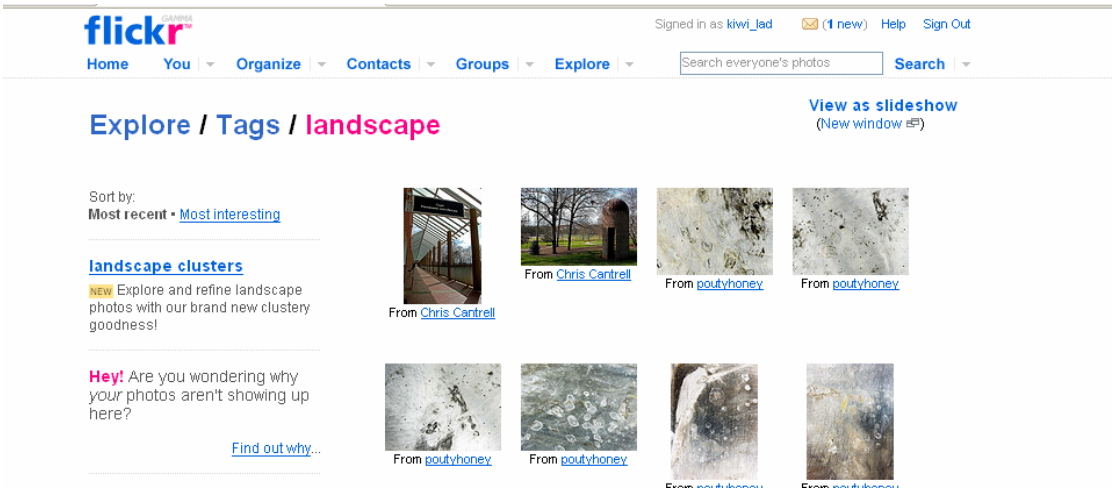

Figure 7 - Selecting landscape

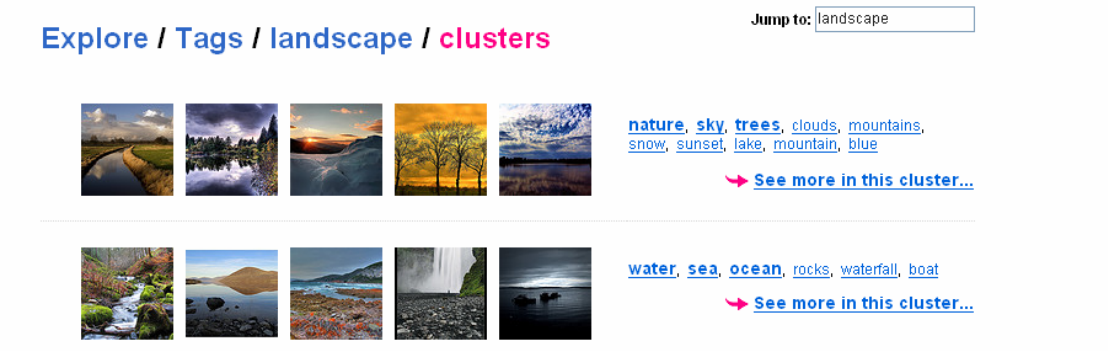

Figure 8 - Landscape clusters

Finally, Figure 9 below shows the upload and tagging process. The tags are entered with spaces to separate the tags and multiple words in quotes.

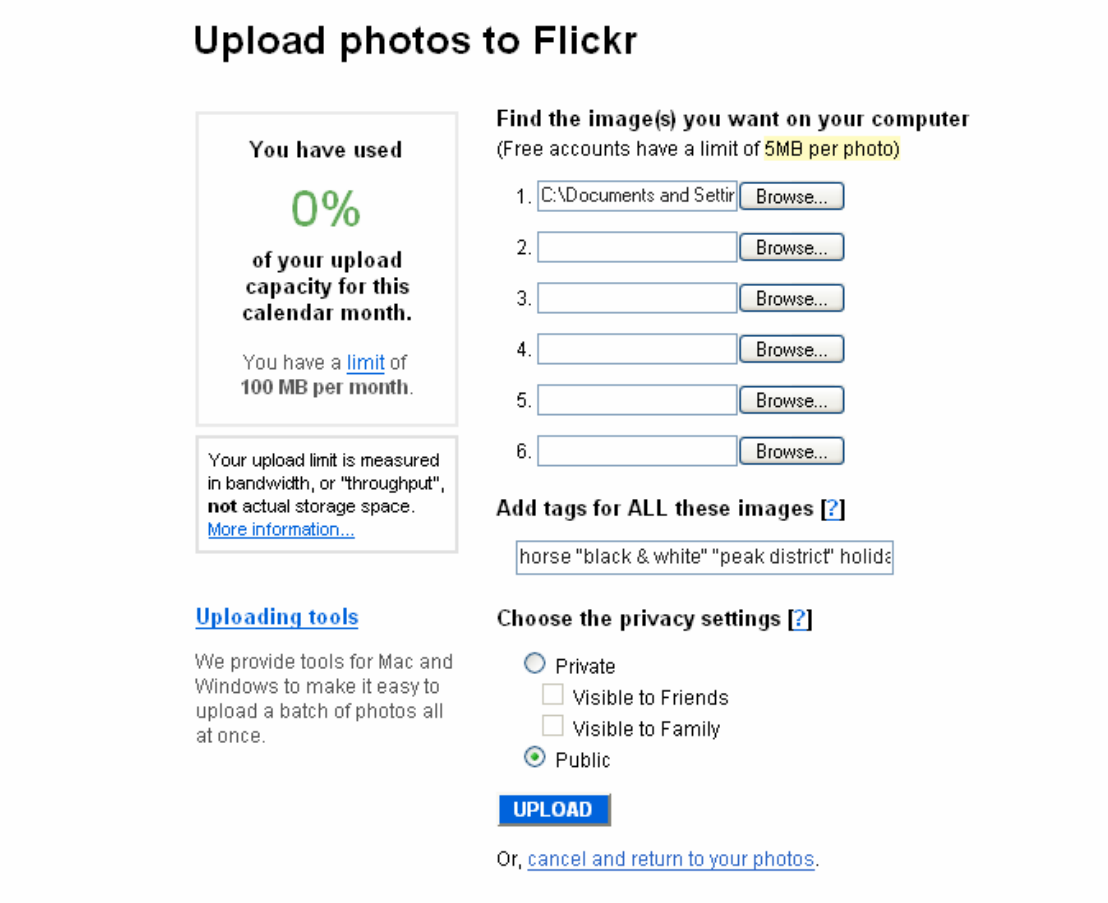

Figure 9 - Uploading and tagging

A few observations can be made with regards to the Flickr approach. Firstly, it is difficult to find photographs by a particular photographer if the photographer's nickname isn't known. At times, it has been difficult to find photos previously viewed. When tagging uploaded photographs the categories previously used may not be entirely suitable and without sufficient "support" from other photographers using the new tagging options they will not appear on the popular tags option. It is a pity that no facility is available to look at all tags, especially as modern linguistic based software and databases (such as thesauri) can create useful links (via topic maps, for example) between tags making all tags available somehow.

We were also a little disappointed by the search facility. Compounding words (adding new words to the search) increased the return size instead of decreasing it, indicating that the OR condition is being used by default (we did discover that one can put “and” between words or phrases in quotes, which fixed the problem, but the help is not extensive). But still, experiments with search came up with very poor and unexpected results.

The aspect of the site that has made the most impression is the effort required to get around the site and to find photographs of specific type or interest.[20] This is not a criticism; we think that in undertaking serious dialogue within a site as this (and even more so on a B2B discovery site) one needs to be patient and careful. Dialogue within the system and between members is likely to occur as a pathway to create trust and community, which is made the more possible by all the socialising features in the site. This therefore represents one advantage of this approach over the semantic web where one wonders whether the philosophy is quick or instant connection with 100% recall, 100% precision. We are not convinced that this represents a good model of reality.

4.2 Example of a BML Model[21]

In comparison with the Flickr example we provide an example of an instance model created for the DBE for web design companies in the West Midlands. We illustrate some problems and some advantages, but to begin, we note some major differences in the tagging process.

The tagging in Flickr initially shown in the bottom of Figure 9 is shown again below in Figure 10.

Figure 10 - Flickr tagging

This type of tagging is excellent because it is discrete (individual words or phrases that are identifiable) but it is unlikely to be sufficient for the purpose of organisational and business service discovery for an eco-system such as the DBE.

In the illustrations below, the firms that responded to the project are predominately web design companies. Within the respondent companies, however, are a few supporting companies such as graphic designers, web hosting companies and a couple of marketing consulting firms.

[20] It is not entirely clear as to what causes this difficulty or effort. One possible reason is that one is still required to remember user names for photographers if one is to find their photographs directly. Other reasons may be that less popular tags are never visible; that insufficient semantic structure is inferred from the metadata tags and hence not made available to the search engine or possibly that there are too many photographs (an expected problem with a system used on such a global scale).

[21] The tags shown here is not the complete set of attributes for this project – only selected ones!! The critique of the model is constructive in intent and if critical, as much self-critical. We think a lot has been learnt in this difficult area and we are hopeful that subsequent research will build on the experiences learned in the DBE project.

The first data block for this DBE model is shown in Figure 11 below. The types of questions and comments one can make of this block are:

- It is important to distinguish company name and basic address details from other attributes provided the search engine interfaces can support these details. How is a search engine going to include the right model elements given the multitude of models?
- What is the search engine going to return and in what circumstances will it return different sets of return values? We have the idea of data elements that are clearly used for input (searching) and others being what are returned from the search. So it is unlikely that anyone would search on the website address or email, but it is likely that they may need these returned from the search. Clearly some attributes could be both.
- The address details are probably an example of both searchable and returned, but it is likely that for searching one would use post code, county and city and not be too concerned with address 1 and 2.

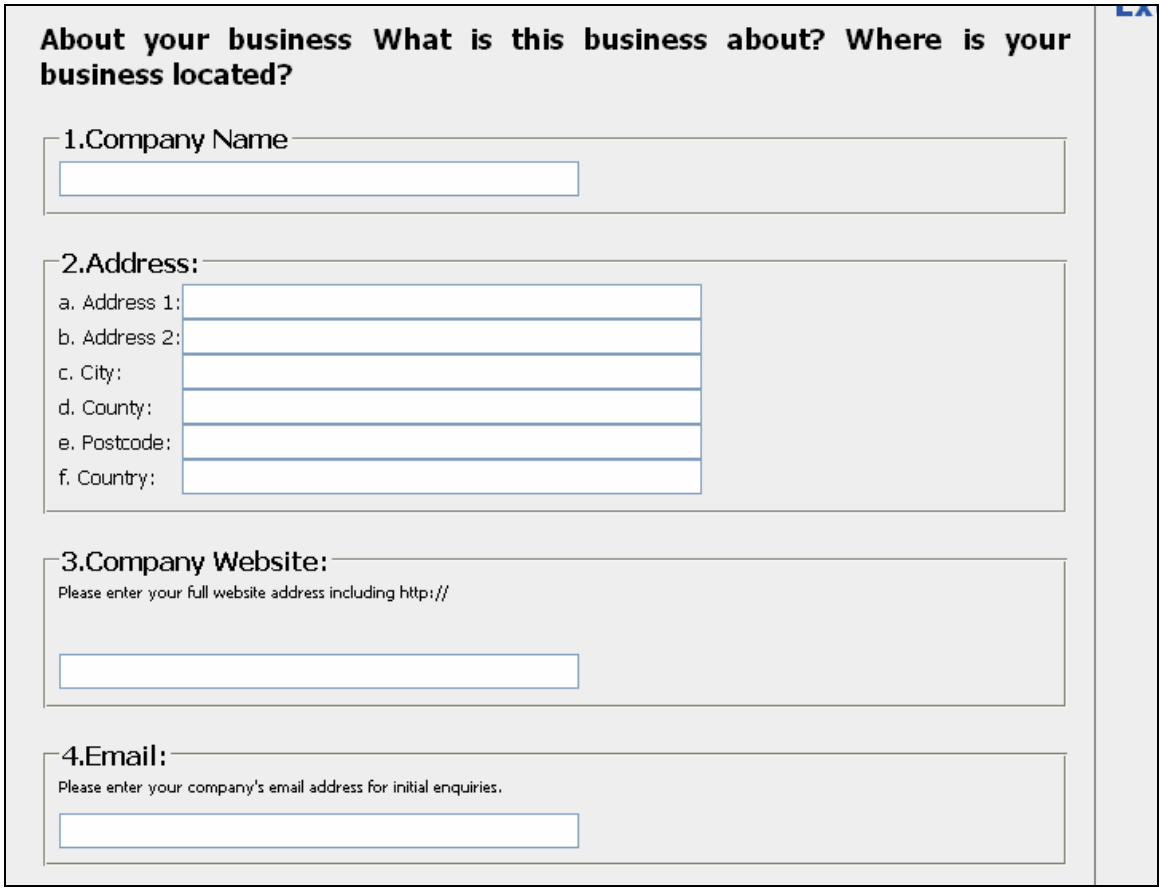

Figure 11 - BML model, First Block

In the second block shown in Figure 12, we have information about the type of business. These are typical search flags, but:

- How are they to be treated by a generic search engine?
- How is it possible to create a type of business list that is suitable for all the related organisations while simultaneously inheriting from a core definition?

- Any ontology sitting behind such descriptions might be more suitable if it enabled capabilities for logical inference (e.g. relationships between types of businesses)
- Our experience indicates that it was difficult to include the supporting businesses in this model – although this problem was more acute on other elements (following).
- It would be useful to use a single text box (doing away with the structure and the inference rules, ontologies etc.) and to follow Flickr, with the additional semantics that the list is "Type of business". Allowing a social construction of values we believe would have been more profitable (implemented here in the "other" although additional information needs to be given to the user – primarily a topic map (say) of the types already entered. This is something Flickr does not provide, which is an important omission on their behalf.[22]

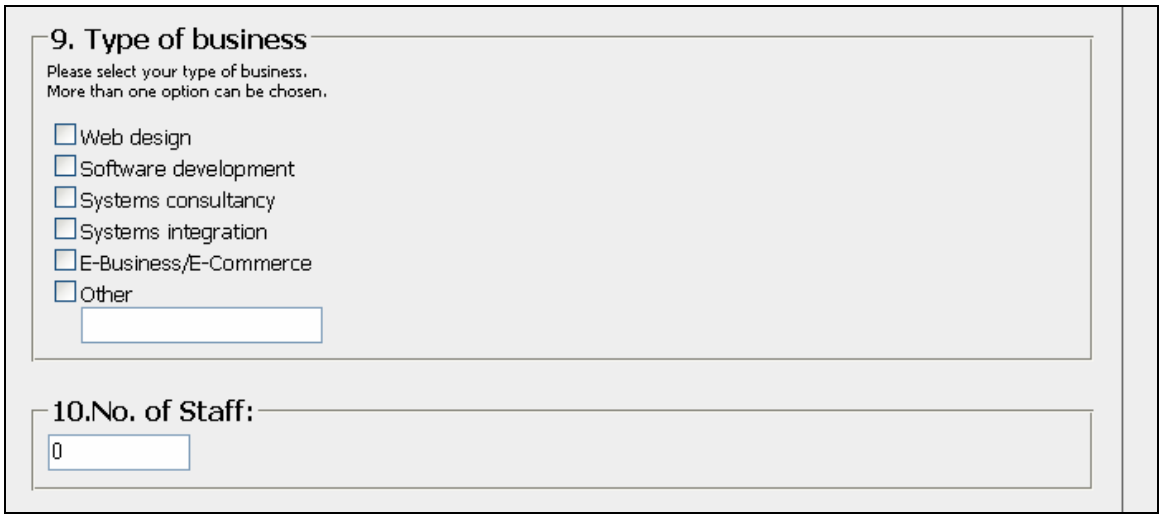

Figure 12 - DBE Model, Second Block

Figure 13 below shows how process data capture was required. Headed *unique or differentiating process*, it was expected that organisations would have such things that could be used for discovery. There are a number of problems:[23]

- It is too structured to have each differentiation as a triplet.
- None of the websites of the 70 organisations examined actually cared to show anything genuinely unique about process. : We can imagine several reasons for this; perhaps the project put an overly strong emphasis on process owing to the chosen MDA approach. However, it is almost certain that:
    - Each organisation has some unique things that they would like to say about themselves that may, nor may not directly tie to processes.
    - It is likely that other people (suppliers, partners, customers, employees) would have some interesting things to say about unique selling points.[24] It is

[22] In community based systems the "global" view of the community is important. Beta products such as IBMs Babble and the new Sugar operating system UI shows this in action. We should learn from these examples in our user interfaces for tagging – making it easier for people to adopt socially acceptable tagging standards or even enabling them to feel comfortable with breaking them and then subsequently monitoring for their success (or otherwise).

[23] Not all of these problems relate directly to the BML meta-model or even the BML model. Restrictions in the UI may have some part as well as expectations of the client (who may have other reasons for including some of the model elements).

possible that outsiders may have a better insight than insiders about these social tags and the system could provide for outsiders to contribute (even if it was against a different tag – such as outsider view).

- The model should not be limited to three values.
- Why would a unique and differentiating process necessarily have a name?

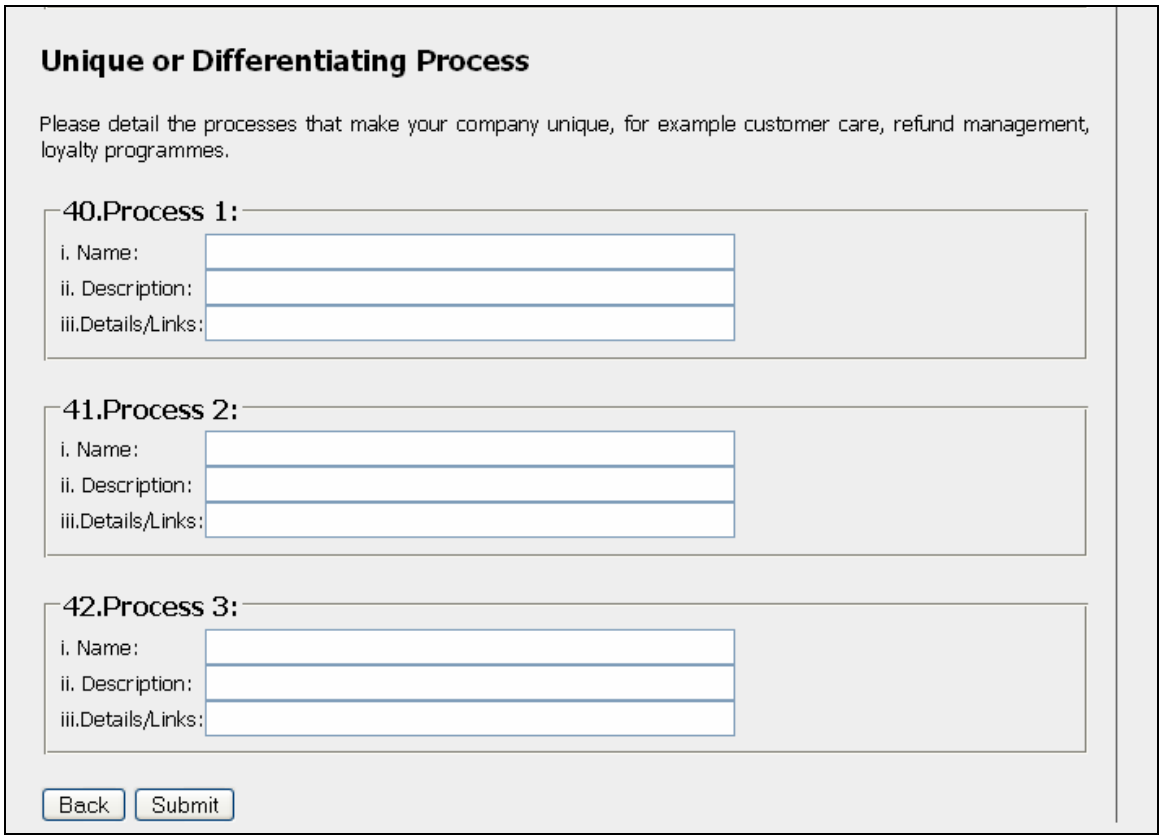

Figure 13 – BML Model, Unique Process

The tags shown in Figure 14 below are very problematic. They are to represent the products and services provided by the participating SME. With 10 sub-tags for each product and service with a restriction to only 2 products and services makes this nearly impossible to complete. Most organisations, even for a simple domain like SME website design companies, have a huge range of products and services and the pricing structures are quite complex and certainly varied. It would not be possible any of the 70 test cases to complete terms and conditions, references (what does this even mean), discounts and so on.

However, the attempt was very radical and brave, particularly given the fact that these are really important elements for any business arrangement. It is reasonably clear to us, however, that (especially for SMEs) pricing and service description is far too complex to describe in this way except for the most basic retail level product sales. Even in most retail cases, the retailer would stock too many items to even provide a basic price in the BML – and without some way to link prices directly with the supporting ERP system, it would be pointless. We expect product and service descriptions that are sharable, that can

[24] This suggestion is not the same as a rating system (such as ebay) or comments section. Trust isn't the issue here, but uniqueness.

aggregate and merge will remain an important research topic for some time yet, although we expect that the BML is not the place for it.

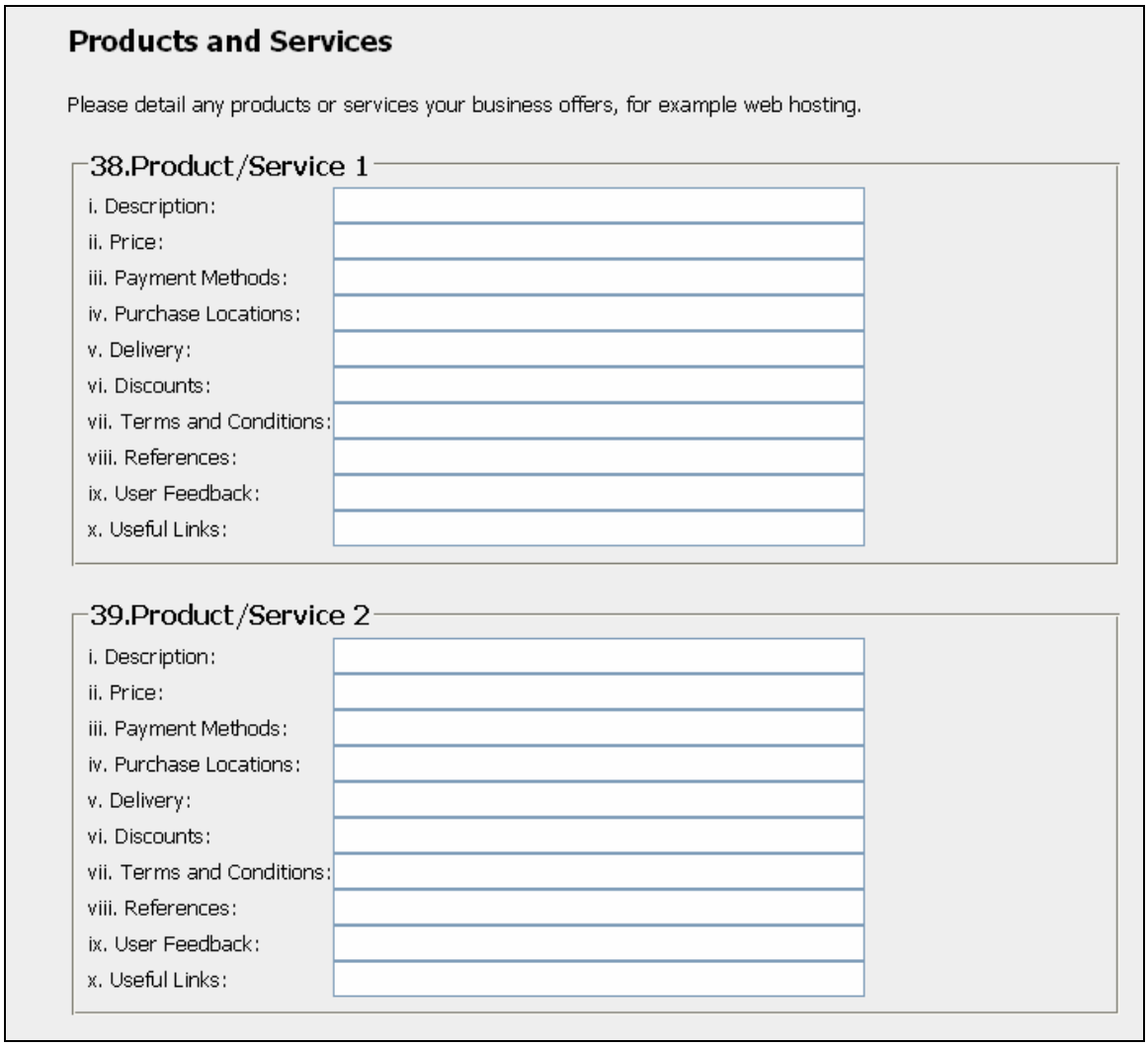

Figure 14 - BML Model, Products and Services

Figure 15 captures skill sets. This is an interesting one, particularly the “Quantity” attribute, which we expect represents the number of staff members with this skill.

How does one capture organisations skill sets? This is a big question in knowledge management and the non-financial measurement community, but we don’t think it is solved with the triple <description, quantity, details/links>. Again, it is an extremely interesting concept to capture and of considerable use to those searching for an organisation, but again we feel that more success at present would be felt with a more flexible way of describing it, provided sufficient feedback is given to the user to help them with the most common descriptors (so that they can align or distinguish themselves).

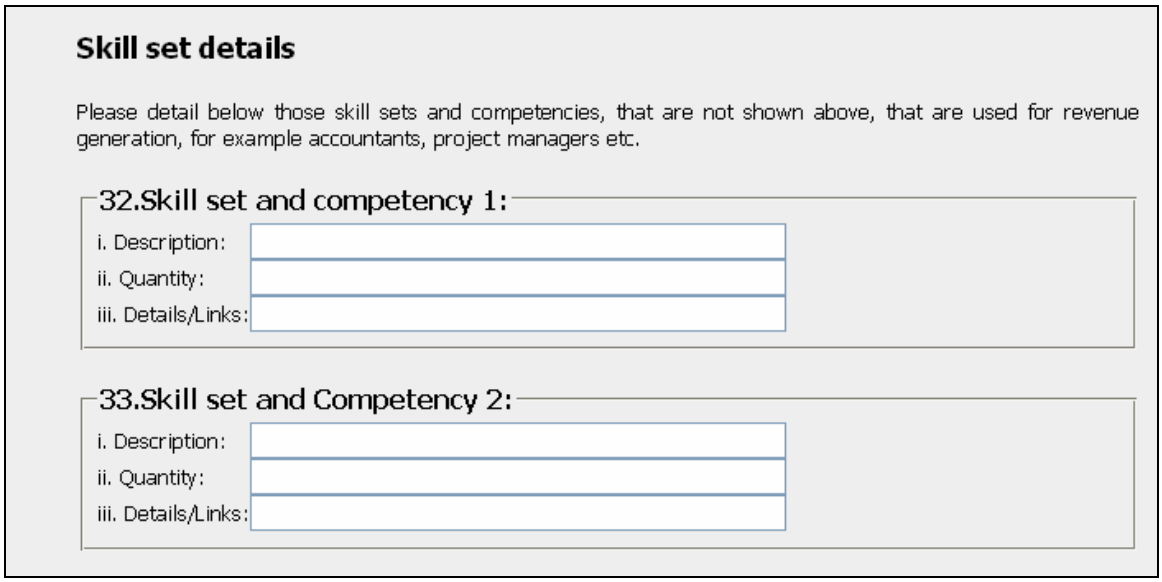

Figure 15 - BML Model, Skill set details

Our final example, shown in Figure 16, provides a list of clients. We think that it is interesting to have such a list, but it is difficult to know whether this list should be represented here or enabled via a link to the website of the company being modelled. If it is being used as a search attribute, then a simple list of names would suffice (again a free format list separated by spaces or commas). But the second attribute is useful because the person doing the search is likely to be interested in the work done for the client. Would that be a search term of a return value term or both?

For the 70 organisations examined, the number of clients listed ranged from zero (where a space for them was provided in the website) to nearly 100.

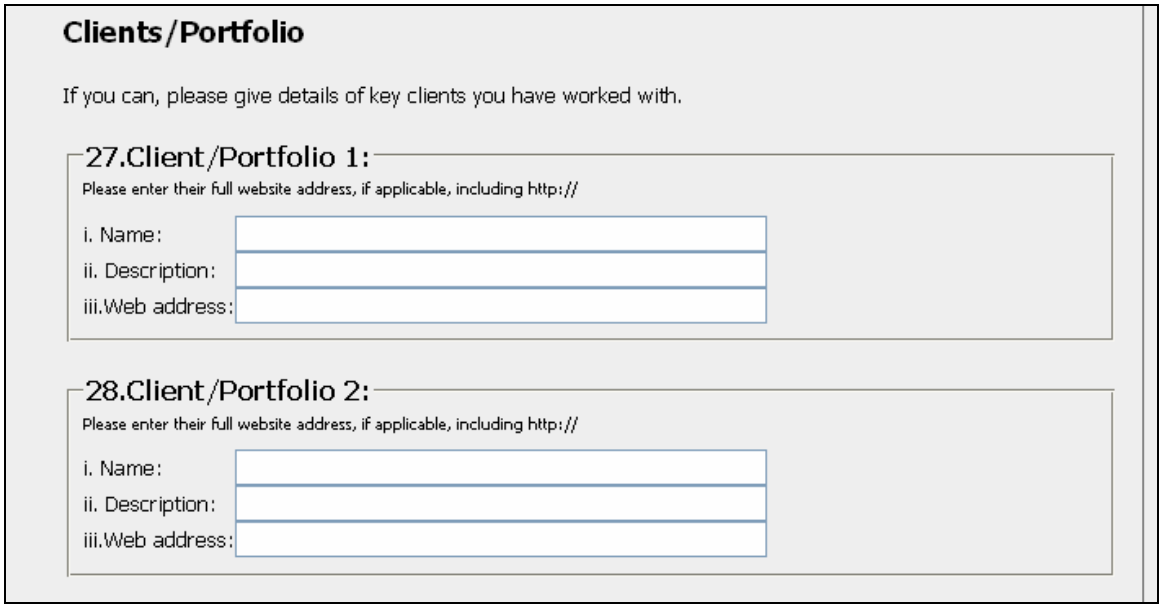

Figure 16 - BML Model, Clients/Portfolio

Regarding the BML modelling part of the DBE studio, the Hypermedia Laboratory of the Tampere University of Technology has performed an evaluation based on case studies carried out with local SMEs. Notably, a recruiting / project brokering system for freelance software experts was implemented together with a provider of software engineering

services. The results are reported in the technical report of Task C60 “Creation and enrichment of the BML and DBE ontologies, Part II: Implementation and evaluation”.

Some of the difficulties reported where:

- an unclear border between what parts should be modelled as a BML model and what part in an Ontology
- no reuse and versioning/updating mechanisms for BML models
- high technical and conceptual skills required for modeling, which even poses challenges to normal IT people
- no support for highly variable elements of the Ontology (such as programming languages known / required)
- No name spaces for attribute names – both for BML models and ontologies
- No list type attributes (often needed and supported by product catalogues)

In the case study, a workaround was implemented for the problem of updating the list of known or required programming languages, using a classic Web application. The DBE Knowledge base infrastructure is altogether not prepared for dealing with such frequent and decentralized changes by a large number of non-expert users.

Another issue identified was that all companies cannot use BML to deliver model-based descriptions to a restricted group of business partners. All models using BML are necessarily public. Altogether the current infrastructure was not designed to deal with issues of selective access rights, trust, and protection against malicious behaviour.

During our own work with the DBE Studio, we also observed that support for a *“BML modelling community”* is altogether lacking. Even if many SMEs/modellers publish their models into the same knowledge base, each company (SME ID) is a separate work space and therefore there is no awareness what other models might exist in the Knowledge Base for the same industry. Models lack metadata about what editing state they are in (draft / proposed / final), and if they are currently used at all, and by whom (this information would in principle be available in the Infrastructure, so it can easily be added in a later version).

5 Conclusion

It is beyond the scope of a market watch to recommend specific solutions, let alone design them. Our purpose is to identify relevant trends in the market place and reflect upon the approaches taken by the DBE project with regards to expectations in the market. Our conclusion can be summarized as follows:

1. Business discovery is a domain of human communication where no single ontological view exists and ever will exist. The Semantic Web technologies and concepts currently employed are problematic in being overly technical (geared towards machine-to-machine communication) on the one hand, and not being able to cope with the multitude of perspectives on the other hand.
2. Theories of social construction have been a very productive domain of 20th century philosophy, sociology and epistemology. They deliver important tools to understand

the problems inherent in large scale social systems and business ecosystems. They do not deliver readily applicable blueprints for alternative software implementations.

3. Web 2.0 applications, especially those in the *Social Software* category support, in a very rudimentary way, principles of social construction and, in the case of bookmarking applications, leave it to each user to find his or her personal structure for organizing content. Added value is generated through the centralization of the database and simple aggregation (e.g. tag cloud); common structures can emerge (but are not forced to do so) as a result of users being aware of what other users are doing. Current Web 2.0 applications are by no means *ideal* tools of social construction or communication.
4. A future solution for business discovery should depart from the strong epistemological assumptions (correspondence theory of truth) inherent in the Semantic Web approach, take into account the multiplicity of human perspectives, inspire itself by current Social Software, but first and foremost study the specific needs for social construction of discovery methods by a business community.

References

[Ball 2004] Ball, Philip; **Critical Mass: How One Thing Leads To Another, Being an enquiry into the interplay of chance and necessity in the way that human culture, customs, institutions, cooperation and conflict arise**; William Heinemann, 2004

[Berger & Luckmann] Berge, Peter L and Luckmann, Thomas; **The Social Construction of Reality: A Treatise of the Sociology of Knowledge**, Allen Lane, The Penguin Press, 1969

Berners-Lee, Tim; Hendler, James; Lassila; Ora. The Semantic Web. Scientific American May 2001, Vol. 284, Issue 5

[Brunner 1990] Bruner, Jerome; **Acts of Meaning**, Harva120

rd University Press; 1990

[Clark 1996] Clark, Andy; **Being There: Putting Brain, Body and World Together Again**; MIT, 1996

[Clayton 2004] Clayton, Philip; **Mind & Emergence: From Quantum to Consciousness**; Oxford University Press, 2004

[Klein 1994] Klein, G; **Sources of Power: how people make decisions**; MIT, 1994

[Hammond 2005] Hammond, Tony; Hannay, Timo; Lund, Ben; Scott, Joanna: Social Bookmarking Tools (I). A General Review. **D-Lib Magazine**, April 2005, Vol. 11 No 4. Available from http://dlib.anu.edu.au/dlib/april05/hammond/04hammond.html

[Hosking & Morley 1991] Hosking, Dian-Marie and Morley, Ian E; **A Social Psychology of Organizing: People, Processes and Contexts**; Harvester Wheatsheaf, 1991

[Hung 2006] Hung, Edward; **Beyond Kuhn: Scientific explanation, theory structure, incommensurability and physical necessity**; Ashgate; 2006

[Morris 1938] Morris, Charles W.: Foundations of the Theory of Signs, in: Carnap, R. Et al (eds.) **International Encyclopedia of Unified Science**, 2:1, Chicago: The University of Chicago Press, 1938

[Noë 2004] Noë, Alva; **Action in Perception** MIT 2004

[Norris 2000] Norris, Christoper; **Minding the Gap: Epistemology & Philosophy of Science in the Two Traditions**; University of Massachusetts Press, 2000

[Rockwell 2005] Rockwell, W. Teed; **Neither Brain Nor Ghost: A nondualist alternative to the mind-brain identity theory** MIT 2005.

[Senge 1990] Senge, P; **The Fifth Discipline: The Art and Practice of the Learning Organization**; Double Day Currency; 1990

[Shirky 2005]. Shirky, Clay: **Ontology is Overrated: Categories, Links and Tags.** http://www.shirky.com/writings/ontology_overrated.html

[Smith 2005] Smith, Barbara Herrnstein; **Scandalous Knowledge: Science, Truth and the Human**, Edinburgh University Press, 2005.

[Snowden 2003] Snowden, D; Innovation as an object of knowledge management: Part 1 the landscape of management; **Knowledge Management Research and Practice**; 2003, 00 1-7.

[Stanbridge & Snowden 2004] Stanbridge, Peter; Snowden D, The Landscape of Management: Creating the context for understanding social complexity; **Emergence** October 2004

[Weissman 2000] Weissman, David; **A Social Ontology**; Yale University Press, 2000