

Critical Success Factors of Semantic Applications

Marek Nekvasil¹, Vojtěch Svátek¹, Ota Novotný²

¹Department of Information and Knowledge Engineering
and ²Department of Information Technologies
University of Economics, Prague, Winston Churchill Sq. 4, 130 67, Prague 3, Czech
Republic
{nekvasim,svatek,novotnyo}@vse.cz

Abstract: Broader and broader areas of application have been covered by semantic technologies recently and in the meantime their scope has been constantly increasing. The possibilities of semantic applications are now so vast that they cannot be judged as one market segment any longer. The skepticism that arises due to the uncertainty of investments in such technologies is even augmented by these differences. This paper provides a possible approach to the categorization of semantic applications and subsequently sets up several critical success factors for the deployment of these technologies in a business environment. Last but not least we outline a possibility how maturity models of enterprises for preliminary assessment of the investments into semantic applications can be formed.

Keywords: semantic technology, critical success factors, process maturity

1. Introduction

By the end of the last century the web was viewed by both academic spheres and general public as a large set of hypertext documents, static or dynamic, more or less available. But then in 2001 this view was changed when Tim Berners-Lee, director of the W3 consortium¹, published his most known article [1] and so the concept of Semantic Web became popular. Consequently and promptly the semantic technologies became known widely and today they are applied in much broader areas than ever before: from applications that integrate data from different sources, support the search in a diverse range of data, derive new relationships across heterogeneous databases, including the application support of social networking, management decision-making, annotating and indexing of any content, up to such different tasks as information extraction from unstructured sources, and even so-called Business Intelligence 2.0 [5].

On the other hand this broad cover of semantic (also known as knowledge-based) technologies is not caused by their omnipotence but on the contrary by the fact that there is no generally accepted definition of what a semantic application is and what it

¹ <http://www.w3.org/>

is not. For example, according to [9] any application that stores data separately from the meaning and content files, and in the same time does not have the meaning hard-wired into the application code, can be called *semantic application*. This concept includes the use of ontology languages (such as RDF², RDFS³, OWL⁴, etc.) and rule-based systems. Therefore the concept of semantic technologies (or semantic web technologies) is far from being exclusive to the Web environment nowadays.

The W3C Semantic Web Education and Outreach (SWEO) Interest Group collects and publishes case studies of existing applications and potential use cases that take advantage of semantic technologies in praxis [8]. Thanks to this overview it is possible to gain insight into the current state of semantic applications and their usability in the production environment, however the extent of this catalogue is yet still rather limited (currently it consists of 27 case studies and 13 use cases).

Of course, for a semantic application to be feasible in the commercial environment it is necessary to justify the investments tied with it. Analogously to other areas of information technology, the costs of such investment can be identified with certain degree of endeavor but quantifying the gains is very difficult not to say impossible. It is caused by the fact that the main advantages of such systems are of non-monetary merits and there are many views on what the actual gains are. Much more essential by the time of assessing the investment is estimating (or defending) its feasibility and determining the necessary conditions under which the whole project will not be loss-making.

However, as we have already mentioned, the notion of semantic application is very diverse from project to project, hence the conditions of feasibility and potential profitability cannot be set generally (yet, some overviews have also been published, see [7]) but it is necessary to identify some *categories of knowledge-based applications* in the first place. Only once these categories are identified, it would be possible to formulate the requirements, because each kind of semantic application can be substantially different.

There are several goals of this work and identification of the most common categories of semantic applications is only the first one of them. It will be possible to isolate some of the substantial properties according to the categories after the applications have been categorized. On this level of abstraction it will be much more suitable to establish some possible sets of *critical success factors* (CSFs [6]). The second aim of this work is thus to identify some of the most important CSFs of deploying (and developing) the semantic applications.

Moreover, once the set of critical success factors has been verified, it will be possible to take this deliberation one more step further and formulate the *maturity models* for deployment of knowledge-based applications (in the sense of the maturity of enterprise processes, according to the original W. Humphrey's work [3]). However in this paper we will only outline the manner in which this is to be accomplished. Some of these ideas have already been published in [4] but this paper extends the classification part and supports it with an automated clustering approach (see the next section).

² <http://www.w3.org/RDF/>

³ <http://www.w3.org/TR/rdf-schema/>

⁴ <http://www.w3.org/TR/owl-features/>

2. Categories of semantic applications

We already mentioned in the first part that the semantic applications cannot be considered as a compact area of interest, because in fact they are really very heterogeneous uses of the appropriate technologies. The individual applications can differ between each other in many aspects (e.g. the scale of the used database, number of interested parties, kind of inputs and outputs or the subject of operation).

Thanks to this fact the categorization of knowledge-based applications is a multidimensional problem. The particular dimensions (i.e. categorization criteria) however had to be identified. This was accomplished by the preliminary analysis of the mentioned case studies published by the W3C interest group [8] that is working through them in detail. Thanks to the comparison of the individual cases the following aspects of differentiation of the semantic applications emerged (not all however have a direct impact on the forming of critical success factor – this will be considered in part 3). In our opinion the results are credible because of the fact that the SWEQ catalogue gathers together cases that represent, rather than a single piece of software, a distinctive kind of applications. These criteria are necessarily of two kinds: the first are specific criteria valid only for semantic applications (the first four) and the other are general valid for all kinds of applications (the other four). The categorization criteria we found follow:

- **Information sources.** The semantic character of considered applications directly implies that at least one knowledge model (ontology or taxonomy) has to be used. Some applications also use other knowledge models or even expect a variable knowledge base. Apart from that the applications can of course also use other data of various kinds. Knowledge-based applications can be divided according to whether they process structured knowledge, structured data or unstructured data.
- **Data source provenance.** Semantic applications can be distinguished according to whether the information they are working with arise in other systems (or are already available in a structured form) or whether they are created specifically for this system. If the data are created exclusively for the semantic system we can further distinguish the cases where this is done manually, automatically from other sources or as a side effect of other activities (such as normal user behavior).
- **Accuracy of inputs and outputs.** Considering the semantic applications we find different approaches of transforming inputs to outputs. Here the applications can be divided e.g. into those firmly relying on full precision of data, applications that expect that the data may be incomplete but do not expect them to be inconsistent and do not work with uncertainty, and finally, applications that include treatment of uncertainty.
- **Subject of operation.** From the analysis of case studies we managed to identify several main types of activities of semantic applications. These are data indexing, data integration and reasoning. These activities are, however, in most cases the means rather than the purpose of the activity (the exception is the integration of heterogeneous data). From these, we can derive several other

activities which support the main purpose of the application, for example, they are enabling better searching capabilities (indexing + integration), heterogeneous database browsing and navigation in the domain (integration + indexation), recommending new relations among entities (reasoning) and allowing the adaptability to change the systems' data structures (data integration).

- **Domain-specificity and reusability of applications.** Because of the separation of data from their meaning the semantic applications should be much less domain-dependent than conventional solutions, but even here there are exceptions, which include, for example, specific interfaces tailored to a specific domain or particular treatment of data on the application level.
- **Number and kind of users.** Users of semantic applications may constitute of unprofessional individual users, professional users (domain experts), knowledge experts and management. Applications can also be distinguished according to whether they are intended for individuals, working groups or thousands of users in social networks.
- **User × provider relationship.** Here we managed to identify several options for operating the applications: the user is an individual and operates the application for his/her own use; there are a few users and they are more or less equal subjects or form a social network and the operation is granted commercially, by the community or non-profitably; the users are the customers of the provider; and finally the users are the employees of the provider. For the last two possibilities we can distinguish cases where the operation of the application is the core business of the company and where it is only a supporting process and can therefore be considered as a possible target for outsourcing. The cases when the operation is ensured by the community can be broken down by whether the operation is centralized or decentralized.
- **Frequency of access to the application and its availability.** Applications may be used continuously (24/7), at random, regularly or by a single opportunity. Furthermore, a distinction must be also made by the availability of such applications: either the application must be available constantly, in defined intervals or on demand (e.g., the reactive manual start of the application).

By sorting the considered case studies we can find out how often some specific values of the proposed criteria are seen in the real-world applications. (The results were reviewed and normalized by one of the authors so they are consistent.) The relative frequencies of these values are shown below on Fig. 1.

Of course, one can imagine a semantic application that is classified by the mentioned aspects more or less arbitrarily, but in the presented case studies certain coincidences and clusters can be identified amongst the various aspects. This sorting thus enables us to identify and name intuitively some basic archetypes of semantic applications, based on examples:

“Improved search engine”. These applications focus on indexing the data, often associated with integrating data from various other systems where the data are generated automatically. These automatically acquired data are also often accompanied by manual annotation. Applications of this type work with both

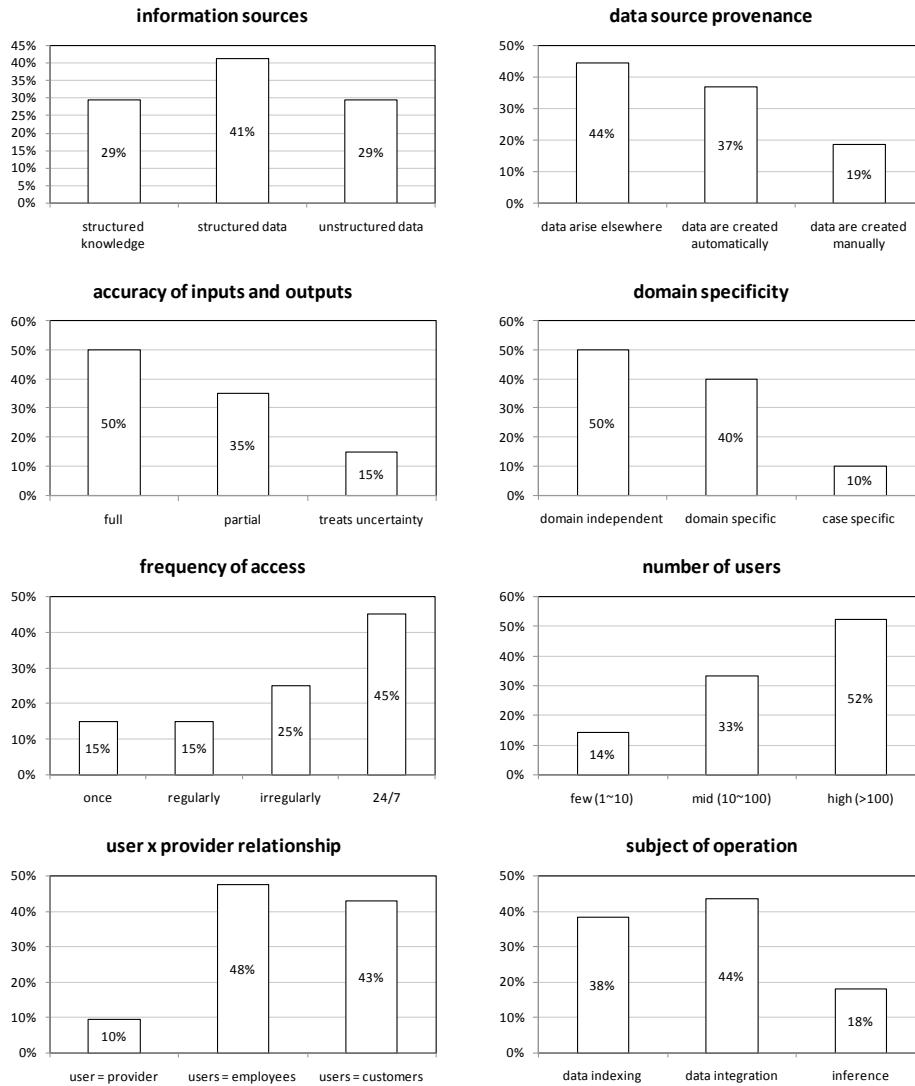


Figure 1 - Relative frequencies of certain criteria values

structured data and unstructured data (using automatic filters and wrappers). The main benefit of these applications is to enable searching in heterogeneous data base and the creation of complex queries without the need for a priori knowledge of data structures. In other aspects, however, they can vary greatly, so such different applications as support for annotating and searching of files on a personal computer⁵, a public portal for searching for findings of Chinese

⁵ <http://nepomuk.semanticdesktop.org>

medicine⁶ and management of sound recordings archives by a Norwegian radio station⁷ can be classified here.

- **“Data-browsing interface”**. These applications follow the abilities of the previous archetype, but enhance not only the possibility of displaying diverse content (videos, articles, chemical formulas, etc.), but also the possibility of visual data browsing, regardless of their structure. These applications are mainly focused on the use by professionals and are operated either commercially for internal use or non-profitably to support a professional community (and simultaneously promoting the technology). Examples of this archetype can be e.g. systems for the aggregation of medical data, whether in order to facilitate the treatment of patients⁸ or achieving savings in the development of new drugs⁹ or portal for the association of programming knowledge by Oracle¹⁰.
- **“Recommending system”**. The nature of these applications is the derivation of new relationships between entities. Moreover, apart from all other types of source data these applications often utilize data that are automatically generated as a side effect of normal user activity, which enables, inter alia, to propose new relationships on the basis of the current users’ context. The user is often the customer of the provider, be it either as a paid service, public service (e.g. designing of individual city tours in Zaragoza¹¹) or a commercial way to personalize advertisement targeting together with the provision of services (such as a system for recommending services to users of mobile devices¹²). Applications in this category often work with uncertainty, thus one can also include a variety of expert systems.
- **“Data interchange framework”**. Operations of applications in this category (because of their nature) are distributed, thus these knowledge-based systems “only” allow to unify the structure of data exchanged between the participants, regardless of their content. This content can evolve over time and be adapted to the needs of a particular bilateral exchange relationship and yet be transmitted in a standardized format. An example is the initiative for the establishment of semantic data interchange in the oil and gas industry¹³.

Of course this list of possible archetypes of semantic applications is not limited to these four; however, we consider these four to be the most usual. Of course there are also applications that cannot be assigned to any of these archetypes, as well as others which, on the contrary, lie in between two or more.

We also tried a more sophisticated way to find clusters amongst the examined case studies by means of the data-mining capabilities of Analysis Services in Microsoft

⁶ <http://www.cintcm.com>

⁷ <http://www.nrk.no/>

⁸ <http://www.pharmasurveyor.com/>

⁹ <http://www.lilly.com>

¹⁰ <http://otnsemanticweb.oracle.com/>

¹¹ http://www.zaragoza.es/turruta/Turruta/en/index_Ruta

¹² <http://www.w3.org/2001/sw/sweo/public/UseCases/SaltLux-KTF/KTF.pdf>

¹³ <http://www.w3.org/2001/sw/sweo/public/UseCases/Chevron/>

SQL Server 2008¹⁴. All the studied cases were assorted into the database manually and the categorization criteria were transformed to binary attributes to facilitate the possibility of multiple values per case. Even with such a small database it was possible to perform a cluster analysis with sensible results. The automated cluster analysis provided by Microsoft Analysis Services (the decision-tree clustering method based on Bayesian Dirichlet node scoring was used but changing to other clustering methods did not have any severe effect) brought up five distinct clusters. These clusters along with the overall attribute values relative distribution are depicted on Fig. 2. Let us discuss them in closer detail:

- **Cluster 1** – The first and biggest cluster is characterized by the following: The application is used by more than a hundred users who are mostly customers of the application provider. The application is based on data indexing and searching with results fully accurate, it is accessed 24/7 and the particular solution is not domain specific. This cluster resembles the first intuitively identified archetype “Improved search engine” in many ways and it seems that they really stand for the same kind of applications with the only uncanny exception – data for applications in cluster 1 are created mostly manually. Therefore this cluster should be better named something like “Data indexers”.
- **Cluster 2** – This cluster also encompasses mainly the applications based on data indexing and searching but they mostly also use some form of inference over the data. The data itself are created automatically and arise in other systems. Most of the users are providers employees who access the system regularly. The application is not only suited for a particular domain but in most case also case-specific. All these characteristics suggest that a typical application from this cluster will resemble both “Data-browsing interface” and “Recommending system” archetypes which of course is not undesirable because the archetypes are not disjoint.
- **Cluster 3** – Characteristics of this cluster are very similar to the first one, the typical application is used by hundreds of providers customers who access the system continually. The application does data indexing but the data are created automatically, in most cases from several systems, and the system performs some form of data integration. This cluster constitutes the purest “Improved search engine” archetype because it does not rely on the user-tagged input data.
- **Clusters 4** – These applications are typically used only by a few users (or even a single user) who are at the same time providers of the system. Main purpose of these applications is again in data integration and indexing but other aspects are not so definite and differ even amongst this cluster. This cluster constitutes of applications that can be characterized as “Personal desktop indexers”.

¹⁴ <http://www.microsoft.com/sqlserver/2008/en/us/data-mining.aspx>

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Size	30%	20%	20%	15%	15%
data are created automatically	17%	51%	75%	67%	66%
data are created manually	100%	75%	0%	34%	32%
data arise elsewhere	33%	76%	75%	33%	100%
uses structured data	67%	76%	100%	100%	100%
uses structured knowledge	50%	73%	0%	0%	0%
uses unstructured data	33%	26%	50%	34%	100%
full accuracy	100%	25%	50%	100%	69%
partial accuracy	0%	75%	50%	0%	32%
treats uncertainty	33%	0%	25%	25%	0%
data indexing	100%	76%	50%	100%	66%
data integration	33%	76%	75%	100%	100%
inference	0%	73%	25%	34%	0%
case specific	33%	73%	25%	0%	0%
domain independent	50%	0%	50%	0%	0%
domain specific	17%	27%	25%	100%	100%
1-10 users	0%	0%	0%	100%	0%
10-100 users	0%	51%	0%	0%	100%
100+ users	100%	49%	100%	0%	0%
user _ customer	100%	24%	100%	0%	0%
user _ employees	17%	75%	25%	34%	100%
user _ provider	17%	25%	0%	67%	0%
access 24/7	100%	0%	100%	0%	0%
access irregularly	17%	51%	0%	34%	100%
access once	0%	25%	0%	33%	33%
access regularly	0%	49%	0%	67%	0%

Figure 2 - Clustering of the SWEO Case Studies

This group is not reflected in the few aforementioned archetypes, but the small number of users and some accompanied characteristics are the only thing that distinguishes it clearly from the other clusters.

- **Cluster 5** – Applications in this cluster are typically designed for the integration of data from various sources, they use both structured but as the only cluster also mainly unstructured data which are processed automatically. The application function is accessed irregularly and the solution depends on the particular domain. This is generally given by the use of some parsers and specifically tailored interfaces and thus this cluster can be called “Wrapper-based integrators”.

The cluster analysis thus introduced two additional kinds of semantic applications not considered at the first sight; one being characterized by a very small number of users and the other by automatic processing of unstructured data.

3. Critical Success Factors

By the synthesis of the risks mentioned in the individual case studies [8] we can draft some of the most frequent critical success factors in the deployment of the semantic applications and their deployment into the production environment, as already indicated in the section 2. These CSFs are however not valid universally but more likely their relevance depends on the kind of semantic application we are dealing with. So far we identified these critical factors for success of semantic applications:

- **Correctness of the core ontology/taxonomy.** This factor holds for all knowledge-based applications and the more complicated and less volatile the used model is the more crucial is its correctness. Achieving this success factor entails the need for recruitment of high-quality analysts and knowledge engineers in the phase of development and deployment of the application, which involves considerable costs. The quality and reach of the used ontologies is not limitless; apart from the costs of creation it also has other more structural constraints (see [2]).
- **Sufficiently steep learning curve of end-users.** This applies to applications that have individual end users. Semantics used in this type of applications entails quite atypical method of control compared to standard applications and the learning curve is rising very slowly. Not only a comprehensive and intuitive user interface of the system is a must, but also clarity and accuracy of the outputs and results is vital for the users' work.
- **The potential of possible benefits to compensate the temporary reduction in productivity during implementation and learning** (as well as operating costs). The benefits of the applications are very diverse and often very vague (in contrast with conventional solutions) and thus can be hardly estimated (and quantified) at the time of the deployment of the system. Operating costs are mostly comparable to conventional applications, but in the phase of deployment, it is necessary to count with temporary decrease of productivity of the users (see previous item). For an application to be successful, this temporary decrease should not be so serious that it overshadowed its potential benefits.
- **Will and discipline of all parties to use the same knowledge model.** In case the operation of application is distributed, it is necessary that all interested parties use a central shared knowledge model. There is therefore a potential risk in terms of the need to negotiate on its form and content.
- **Synchronized distribution of central ontology.** Gradually, there may be modifications of the central knowledge model that arise subsequently and if the operation is distributed, it is necessary that these changes are properly disseminated amongst interested parties, or else this could lead to some inconsistencies. While these changes and modifications take place the previous item still holds.
- **Sufficient number of users.** If the respective semantic application is based on social data, its success is conditioned by the existence of a large enough number of users that produce this data. The risk in this case occurs in the form of necessary expenses for the promotion of an emerging system.

- **Users' motivation.** This critical factor occurs at two levels. The first is in the time of the introduction of a new application while the user experiences a negative stimulation in the form of the slow rise of the learning curve. The user thus lacks the motivation to learn to deal with the system in the first moment. Moreover the user that does the work is not always the one who benefits from it (discussed in [2]) which can be of a further burden. The second is in the actual phase of operation; a common source of data for the semantic systems of all sizes is manual annotation, whose creation is up to a certain point very labor-intensive for the users. Partial source of motivation may be a potential benefit of the better results or a facilitation of work in the future. In addition the user can be motivated by the possibility of using the experience gained elsewhere, while even the most different semantic applications use similar technologies (e.g. SPARQL querying).
- **Sufficient supply of data.** For applications that use some reasoning having a sufficient data source is very essential for providing beneficial results (i.e. utilizing the added value of semantics). Even for applications based on data indexing, having enough data is critical to the success, because for small volumes of data they give similar results as traditional methods but with higher initial costs.
- **Diversity of sources and forms of data.** The greater the richness of the knowledge-modeling language (namely, its part actually used in the application), the more beneficial results can be produced by applications based on the derivation of new relationships. Likewise, the greater is the diversity of data content the more useful are the results given by applications performing semantic integration. The use of semantic technologies on trivial systems will therefore likely not pay off.
- **Maintaining at least the same accuracy of results as the sub-systems.** Applications that integrate data of some source systems are at risk of finding an inconsistency in the aggregated results. The functionality of semantic applications itself is not subject to consistent data, but the possible inconsistency should be expected in the design. Good estimation of the reliability of data sources is thus crucial at this point.
- **Reliability of parsers and wrappers.** If the application handles unstructured data, it is dependent on the output of parsers and wrappers of various content and, where appropriate, the natural language processing systems. Here again the same applies as in the previous paragraph, namely that it is necessary to correctly estimate the reliability of the information obtained in such a way.

Of course, these critical factors will be weighted differently in the scope of different applications. If, for example, the source application collects data automatically and passes the outputs to the user in almost natural language and in an appropriate context, we can expect a relatively steep learning curve, so that the period of reduced productivity is quite minimal and as a result it will be compensated enough even by minor benefits. These universal critical success factors can only be taken as starting points when considering a particular case.

4. Future Work – Maturity Models

Maturity models [3] have developed over the past two decades in order to enable to assess the readiness of enterprise to implement some kind of structural investment. Most commonly they are used in the deployment of any IT applications such as CRM systems, ERP and Business Intelligence. In our opinion it should be possible on the basis of the above aspects of categorization and the associated critical success factors to establish enterprise maturity models for the deployment of a certain type of semantic technologies. Although these models would be without factual content and without the target statement, they could be formulated by concrete examples and at least for each archetype would thus make it possible to set a certain level of requirements for an enterprise, which should be met in order to consider the feasibility of the solution.

An example maturity requirement for the archetypal semantic search engine could look like this: *If an enterprise uses a single source of data and a proprietary data structure, then it is unprepared for the introduction of this kind of system. If it is using multiple systems with heterogeneous data structure, the introduction of search engines with semantic indexing can bring some improvements to the search results. The enterprise achieves next level of readiness if it uses more systems with a standardized data structure; in such case it can start thinking about the integration of these systems with a semantic data exchange, etc.*

Before formulating such exemplar maturity models the critical success factors need to be evaluated because their validity is vital. We plan on performing a detailed survey of the successful semantic technology projects (based on the SWEO catalogue) and test whether the critical factors hold. As a side effect such a survey will help quantify exact boundaries and limits in our dimensional categorization approach. Only after verifying the CSFs we can move onto formulating the maturity models.

5. Conclusion

In this paper we introduced a method of categorization of semantic applications that is both consistent (due to multidimensional set of criteria) and practical because certain clusters of applications emerge and thus they can be named, characterized and referred to. The main study was performed on the catalogue of semantic applications published by the W3C interest group and therefore certain degree of independence of the source materials is assured. The sample clustering was done intuitively in the first run when we came up with certain application archetypes and then the idea was supported by an automated cluster analysis. The results were similar but the automated approach brought in some details that were overlooked at first.

We then formed a set of some most critical success factors for the deployment of semantic applications in a commercial environment. There are of course certain limits of validity of these CSFs and the most important of them were discussed continuously. Last but not least we proposed a way of how these efforts so far can lead to establishing elaborate models for maturity of enterprises for semantics, which is however still left for future work.

References¹⁵

- [1] Berners-Lee, T., Hendler, J., Lassila, O.: *The Semantic Web*, Scientific American, May 2001. <http://www.sciam.com/article.cfm?id=the-semantic-web>
- [2] Hepp, M.: *Possible Ontologies: How Reality Constrains the Development of Relevant Ontologies*, in: IEEE Internet Computing, Vol. 11, No. 1, pp. 90-96, Jan-Feb 2007 (invited), doi:10.1109/MIC.2007.20
- [3] Humphrey, W.: *Managing the Software Process*, Addison-Wesley Professional, Massachusetts, USA, 1989
- [4] Nekvasil, M., Svátek, V.: *Towards Models for Judging the Maturity of Enterprises for Semantics*, In: Business Information Systems 2009 Workshops, Lecture Notes in Business Information Processing, ISBN 978-3-642-03423-7
- [5] Raden, N.: *Business Intelligence 2.0: Simpler, More Accessible, Inevitable*, <http://www.intelligententerprise.com/showArticle.jhtml?articleID=197002610>
- [6] Rockart, J. F.: *Critical Success Factors*, Harvard Business Review, 1979, p. 81-91
- [7] Sauermann, L.: *Benefits of Semantic Web .. for you*, DFKI GmbH, 2008 <http://www.dfki.uni-kl.de/~sauermann/2008/04/benefits/>
- [8] Semantic Web Education and Outreach (SWEO) Interest Group: *Semantic Web Case Studies and Use Cases*, <http://www.w3.org/2001/sw/sweo/public/UseCases/>
- [9] Staab, S., Studer, R.: Handbook on Ontologies - Preface. *International Handbooks on Information Systems*, Springer Verlag, 2004

¹⁵ The hypertext links are valid to the date of September 30th, 2009