e-swim: Enterprise Semantic Web Implementation Model
Towards a Systematic Approach to Implement the Semantic Web in Enterprises

Reinaldo Ferreira[1] and Isabel Seruca[2,3]
[1]*Planeta Virtual, Porto, Portugal*
[2]*DICT, Universidade Portucalense, Porto, Portugal*
[3]*Centro Algoritmi, Universidade do Minho, Braga, Portugal*

Keywords: Semantic Web, Linked Data, Implementation Model, Enterprise.

Abstract: The adoption of Semantic Web technologies constitutes a promising approach to data structuring and integration, both for public and private usage. While these technologies have been around for some time, their adoption is behind overall expectations, particularly in the case of Enterprises. This paper discusses the challenges faced in implementing Semantic Web technologies in Enterprises and proposes an Implementation Model that measures and facilitates that implementation. The advantages of using the model proposed are two-fold: the model serves as a guide for driving the implementation of the Semantic Web as well as it helps to evaluate the impact of the introduction of the technology.

1 INTRODUCTION

Over the last twenty years, the Web evolved from a laboratory where it consisted in merely an idea to a ubiquitous and universal environment, accessible from basically everywhere, with information from basically everything. Yet, the "Web of Documents" is far from being considered a fully structured version of the Web. The thinking underlying the Semantic Web, providing a common framework that allows data to be shared and reused across application, enterprise and community boundaries (W3C, 2013), puts forward the introduction to data access as well as to data structure and meaning, and enables a more integrated robust and practical environment. It is believed that the next two decades will provide enough time to spread the "Web of Data", being that the expectation of the scientific community.

The introduction of the Semantic Web involves the progressive transformation of the Web based on hyperlinks between documents, the base for its first generation, in the Web based on hyperlinks between data or information, giving place to a Web scale distributed database (Heath & Bizer, 2011). This very large database is still in its early days, although a rather significant number of examples that allow to evaluate its potential can be found, since some research efforts and government policies for data publication have already produced satisfactory results (Bizer, Heath, Idehen, & Berners-Lee, 2008). Despite these contributions, Semantic Web technologies usage within the enterprise community is still a rather unexplored theme and in an early adoption phase (Ahmed & Gerhard, 2010; Kuhn, 2010). The reasons for this reduced level of adoption may point to several difficulties, such as the homogenization and validation of data sources, the definition of knowledge rules and borders that allow to relate data in a uniform way, the analysis of too complex examples, the availability of low cost technological capacity to allow its implementation, the availability of development tools, the recruitment of experienced professionals, the diffusion of success stories and the adoption of a paradigm shift in modelling, design and development (Ahmed & Gerhard, 2010; Kang et al., 2008; Kuhn, 2010; Pollock, 2008).

Acknowledging these shortcomings, this paper presents a model for Semantic Web implementation in Enterprises with two main goals. The first goal is to facilitate the introduction of the technology in organizations with different characteristics and motivations, acting as a guide and providing a roadmap for a quicker and more intensive adoption. The second goal is to evaluate the impact of the introduction of the technology in the applications used in these organizations and in the tasks

Ferreira R. and Seruca I..
e-swim: Enterprise Semantic Web Implementation Model - Towards a Systematic Approach to Implement the Semantic Web in Enterprises.
DOI: 10.5220/0004972706530658
In *Proceedings of the 16th International Conference on Enterprise Information Systems* (ICEIS-2014), pages 653-658
ISBN: 978-989-758-028-4

performed by their users.

This paper is structured as follows. Section 2 discusses the present and future trends for Semantic Web implementation. In Section 3 we present the e-swim, the Semantic Web Implementation Model for enterprises proposed in this paper. In Section 4 we outline the research methodology used to validate the model and include some preliminary findings. Section 5 concludes with a summary of the project achievements and directions for future work.

2 THE SEMANTIC WEB: PRESENT AND FUTURE TRENDS

The economic dimension is one of the most important dimensions of the World Wide Web. It has been so in its first generation, with “long tail” enterprises approaching the global market of customers and products, until the targeting of the smallest segment. It has also been so with the growth of social networks, and these being explored as a competitive advantage to connect people to enterprises, brands and products. And it will eventually be so in its third generation, with enterprises actively participating in the construction of the Semantic Web, the Web of Data. It constitutes certainly a great and broad research opportunity, in many subjects, allowing the space for the creation of the Web Science (Berners-Lee, Hendler, Hall, Shadbolt, & Weitzner, 2006; Ferreira & Seruca, 2013a).

Semantic Web technology may transform enterprise software, contribute to the emergence of new business models and reduce costs in areas like data integration, master data management and enterprise information management (Pollock, 2008). Enterprise internal networks based on Linked Data principles constitute a particular subset of Semantic Web technologies that, amongst other benefits, may substantially reduce information integration costs, such as in the integration of information about a product, supplier, materials, legal, market and finance data or other internal and external data sources (Janowicz & Hitzler, 2010). However, the growing adoption of Semantic Web technologies and Linked Data principles raise the question of which applications may be developed to take advantage of this potential. The answer may be obtained from identifying the areas where these technologies and principles may constitute a distinctive contribution, when compared with traditional technologies (Heath, 2010).

Meanwhile, the compromise between the issues of computational effort and flexibility tends to favour the latter, in line with Moore’s law projection, since additional computing power leads to less concerns about simplification or optimization. Simultaneously, the economy has benefited from technological innovation and became a highly competitive environment where speed and flexibility often play a more important role than robustness and trust. The economic validity of current data normalization models can therefore be questioned in comparison to the flexibility and universality promises of the Semantic Web (Segaran, Evans, & Taylor, 2009).

The Semantic Web development has been guided by the way different communities envisage its evolution, considering their specific areas of research. One approach is that of semantic annotation, addressing the large volumes of data available on the Web and using different techniques to originate structured data. Another approach is that of data repositories, starting from pre-defined structures that are updated and interlinked with additional structures. Finally, the approach that puts forward the Semantic Web as an agent platform, with applications combining different data sources and, ultimately, executing actions in replacement of individuals (Domingue, Fensel, & Hendler, 2011).

Additionally, considering the diverse nature of organizations, it seems reasonable that the enterprise approach may be different from the academic one. Enterprises will look for more practical results, focusing on the short term and higher success rates, while academic research will typically look for theoretical results that are more ambitious and focused on the medium term. However, the fact that too much divergence may cause a fracture, with enterprises focusing on a small number of issues and academia in issues that will never be tackled (Cardoso, Miller, Su, & Pollock, 2008), suggests to take a balanced view between both approaches.

The debate over the possible commercial success of the Semantic Web is exhausted and has been replaced by the discussion of what changes in commercial software applications may occur with the introduction of the technology. Some authors suggest as possible evolutions the proliferation of highly distributed applications, agile development and dynamic integration of legacy applications, sensor networks and decision support systems (Domingue et al., 2011).

3 THE ENTERPRISE SEMANTIC WEB IMPLEMENTATION MODEL

Acknowledging the different issues to consider in the adoption of Semantic Web technologies in Enterprises, it is envisaged that a model for Semantic Web implementation in Enterprises will constitute a useful approach to guide enterprises in the introduction of the technology as well as in the evaluation of its' impact in the applications used and tasks performed by users in these organizations (Ferreira & Seruca, 2013b).

The model hereby presented is called "e-swim", an acronym for Enterprise Semantic Web Implementation Model, and is based in four dimensions, where each dimension considers different requirements for the technology implementation and, simultaneously, serves as a guide to identify the desired continuous evolution through subsequent steps.

Figure 1 illustrates the e-swim model and its four dimensions: Adoption, Provenance, Accessibility and Activities. The following sub-sections describe the rationale for each dimension.

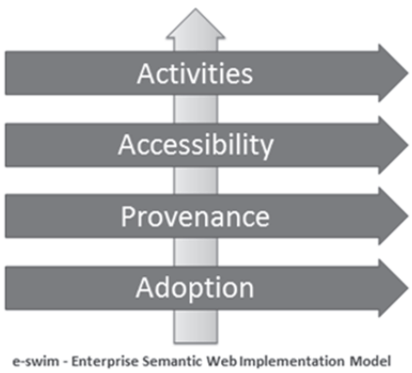

Figure 1: The e-swim Model.

3.1 Adoption

The purpose of this dimension is to determine the degree of preparation of the Enterprise to adopt Semantic Web technologies.

The main technological adoption models for enterprise application are the Innovation Diffusion Model and the Technology-Organization-Environment Model (Oliveira & Martins, 2011). According to the Innovation Diffusion Model (Rogers, 1995), organizational innovation is essentially dependent on Leadership, Organizational Structure and Openness and is adopted according to a normal distribution of organizations that includes Innovators, Early Adopters, Early Majority, Late Majority and Laggards. The Technology-Organization-Environment Model proposes three issues of the enterprise context that influence technological innovation, namely Technology, Organization and Environment. By combining both approaches, the e-swim Model considers the following features for the dimension:

- Technology, including opportunities of technology usage in the organization, in this case, Semantic Web technologies and its applications;
- Organization, referring to internal organization relations and including:
  - Leadership, attitude towards change from the top management
  - Structure, relations between people in the organization
- Exterior, about the external framing of the organization, including:
  - Interface, openness of the organization to the outside
  - Environment, players in the space where the organization is located

These features are summarized in a graphical form in Figure 2.

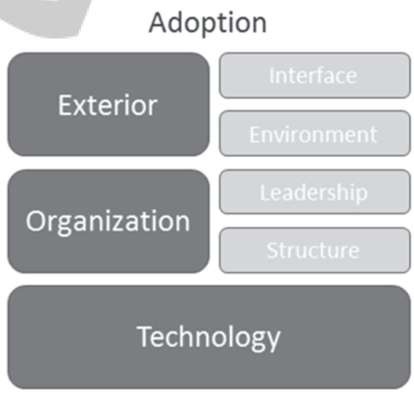

Figure 2: The technology Adoption dimension.

The planned intervention in enterprises takes these features into account and, through observation and interrogation, tries to quantify the position of each Enterprise in the path from a lower to a higher technology adopter.

3.2 Provenance

This dimension of the Implementation Model considers data provenance as a determinant factor. Enterprise innovation and competitive advantage depend entirely of its capacity to deal with a constant and always growing information flow. Consequently, information integration efforts must follow that growth. Semantic Web technologies usage in those integration efforts may increase substantially the return, reducing integration costs and increasing subsequent benefits (Janowicz & Hitzler, 2010). Opening public data to citizens represents an increasing democratic transparency, possible due to technological availability. In that area, the efforts of the American and British governments, among others, already led to a broad availability of data sources with wide usage possibilities. The need to explore these data sources reveals itself primarily as the possibility to explore wealth sources (Koumenides, Salvadores, Alani, & Shadbolt, 2010).

Therefore, the purpose of this dimension, as illustrated in Figure 3, is to identify those data sources, according to:

- applications topology, namely Web Sites, Extranets, Intranets, Web Applications or Web Services
- location: outside or inside the enterprise

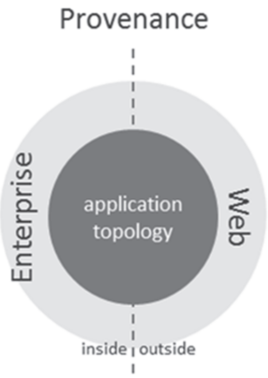

Figure 3: Enterprise Data Provenance.

The planned intervention aims to identify the data sources in the context of the enterprise Information Systems and to quantify their respective usage.

3.3 Accessibility

The common user perception about data available in the Web is largely influenced by the format of these data. However, the structure of these data can influence their subsequent usage. The path to the universal availability of data, as shown in Figure 4, was clearly identified by Tim Berners-Lee with a five stars classification or five steps of evolution (Berners-Lee, 2006):

- available: just having data available in the Web
- formatted: through proprietary formats
- open: with open formats
- semantic: via semantic standards
- interconnected: with data hyperlinks.

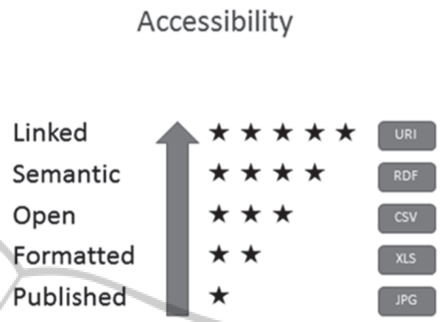

Figure 4: Data Accessibility, adapted from (Berners-Lee, 2006).

The planned intervention will determine the degree of openness of data sources and consequently quantify their suitability towards Semantic Web standards.

3.4 Activities

Information Systems have been classified according to different approaches, some with a broader scope, others with more specific purposes (Lopes, Morais, & Carvalho, 2005). Despite the high number of efforts, the issue is not exhausted and, in this particular case, it is important to find a classification that positions the Web as a base environment, in alternative or in complement to more traditional classifications. With several studies related with Web usage, Tom Heath's work introduces important clarifications and a purpose oriented classification (Heath, 2010). According to this classification, user activities may be instantiated in the following categories:

- Locating: look, find
- Exploring: gather, research
- Grazing: navigate, browse, follow
- Monitoring: monitor, check, detect
- Sharing: distribute, collaborate
- Notifying: state, inform, communicate
- Asserting: opinion, suggestion
- Discussing: comment, respond
- Evaluating: assess, analyse
- Arranging: combine, negotiate
- Transacting: transfer, pay

Figure 5 shows the space for Semantic Web technologies implementation according to the degrees of diversity and conceptualization of the tasks performed.

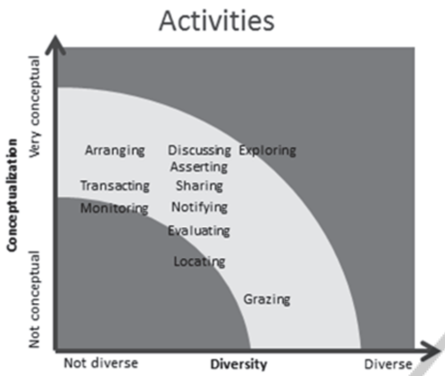

Figure 5: Activities over Conceptualization and Diversity, adapted from (Domingue et al., 2011).

The conceptualization and diversity degrees of the tasks performed influence technology implementation feasibility. Simple and repetitive tasks are those naturally already satisfied by traditional software applications and that will less benefit from Semantic Web technologies. As task conceptualization increases, opportunities arise for the technology, but in the conceptualization threshold, when tasks are hardly typified and demand a person's creative intervention, the effort for their implementation is simply higher than that of just executing the tasks. Low diversity tasks are once again easily supported by traditional software and in the threshold of diversity it will rarely be worth systematizing tasks to be supported by the technology. Hence, the ideal space for Semantic Web technology implementation will be that of some conceptualization or diversity that make tasks too complex or too diverse for traditional software (Domingue et al., 2011). By positioning the previous classified activities in this space, the opportunity for this technology can be visualized.

This dimension addresses the wide area of the enterprise Information Systems and introduces a fit to action approach adequate to an Action-Research effort. The planned intervention will determine what are the tasks performed that may better incorporate Semantic Web technologies and quantify their usage through the Enterprise.

4 MODEL VALIDATION

In order to validate the proposed model, an Action-Research based intervention was planned in several enterprises. Previous work conducted identified the low awareness for Web technologies and lack of collaboration between enterprises, academia and government (Ferreira, 2013), which justifies the need for an immersive approach, where all parts involved work for a common goal. The approach to take would have to consider, within the scope of the proposed implementation model, the need to provide technological leadership, develop awareness for Web technologies, facilitate access to training where appropriate, identify and quantify available resources, qualify the tasks performed and involve the participants in a collaborative effort to, whenever possible, implement change recommendations.

The planned intervention is based on two cycles, where the first cycle aims to identify recommendations for each individual enterprise and the second to share and reapply the most relevant recommendations with all the participant enterprises. The first results of this intervention are promising and challenging, with several recommendations being currently developed. Some of these recommendations are clearly adequate for wide reapplication, benefiting the outcomes of this research work, while others are more specific to each enterprise, contributing to the satisfaction of individual expectations.

5 CONCLUSIONS

Semantic Web constitutes an innovation opportunity for Enterprises. This research project identified implementation barriers, such as the organizational structure, as well as important issues to address, namely what are the enterprise software applications that may benefit with the technology introduction.

This paper proposes an Enterprise Semantic Web Implementation Model (e-swim) based in four dimensions:

- Adoption
- Provenance
- Accessibility
- Activities

The Implementation Model hereby proposed as well as the results obtained with the planned intervention in several enterprises should facilitate the measurement of the implementation degree of the technology in an Enterprise and serve as a guide for the Semantic Web technology implementation by providing a roadmap for a quicker and more intensive adoption as well as pointing possible evolutions in the process.

REFERENCES

Ahmed, Z., & Gerhard, D. (2010). Role of Ontology in Semantic Web Development. *Knowledge Management*.

Berners-Lee, T. (2006). Linked Data Desing Principles. Retrieved July 14, 2012, from http://www.w3.org/DesignIssues/LinkedData.html.

Berners-Lee, T., Hendler, J., Hall, W., Shadbolt, N., & Weitzner, D. J. (2006). Computer science. Creating a science of the Web. *Science*, *313*(5788), 769–771.

Bizer, C., Heath, T., Idehen, K., & Berners-Lee, T. (2008). Linked data on the web (LDOW2008). In *WWW2008 Workshop on Linked Data on the Web* (Vol. 2008, pp. 1265–1266). ACM New York, NY, USA. doi:10.1145/1367497.1367760.

Cardoso, J., Miller, J., Su, J., & Pollock, J. (2008). Academic and Industrial Research : Do their Approaches Differ in Adding Semantics to Web Services ? *Lecture Notes in Computer Science*, *3387/2005*, 14–21. doi:10.1007/978-3-540-30581-1_2.

Domingue, J., Fensel, D., & Hendler, J. A. (Eds.). (2011). *Handbook of Semantic Web Technologies*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-92913-0.

Ferreira, R. (2013). Applying Semantic Web technologies to City Tourism information. In *1st International Conference - Porto as a Tourism Destination*. Porto, Portugal: CEPESE.

Ferreira, R., & Seruca, I. (2013a). A Ciência da Web : oportunidades de investigação. In *13ª Conferência da Associação Portuguesa de Sistemas de Informação* (pp. 13–31).

Ferreira, R., & Seruca, I. (2013b). Modelo de Implementação da Web Semântica nas Empresas. In Á. Rocha, L. P. Reis, M. P. Cota, M. Painho, & M. C. Neto (Eds.), *Actas da 8ª Conferência Ibérica de Sistemas e Tecnologias de Informação (CISTI'2013), Lisboa, Portugal, 19-22 Junho 2013* (pp. 615–620). Lisbon: AISTI.

Heath, T. (2010). A taskonomy for the Semantic Web. *Semantic Web Journal*, *1*, 75–81. doi:10.3233/SW-2010-0003.

Heath, T., & Bizer, C. (2011). Linked Data: Evolving the Web into a Global Data Space. In J. Hendler & F. Van Harmelen (Eds.), *Proceedings of the 17th international conference on World Wide Web WWW 08* (Vol. 1, pp. 1–136). Morgan & Claypool. doi:10.1145/1367497.1367760.

Janowicz, K., & Hitzler, P. (2010). Creating knowledge out of interlinked data. *Semantic Web*, *1*, 97–104. doi:10.3233/SW-2010-0019.

Kang, S. S., Yang, J. Y., Lee, S. K., Gong, K. H., Myung, J. S., Park, S. C., & Lee, S. G. (2008). An Enterprise Strategy for Semantic Technology Adoption. In *The 5th International Conference on Information Technology and Applications*.

Koumenides, C. L., Salvadores, M., Alani, H., & Shadbolt, N. R. (2010). Global Integration of Public Sector Information. In *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line, April 26-27th, 2010, Raleigh, NC: US*.

Kuhn, W. (2010). Modeling vs encoding for the Semantic Web. *Semantic Web Journal*, *1*, 11–15. doi:10.3233/SW-2010-0012.

Lopes, F. C., Morais, M. P., & Carvalho, A. J. (2005). *Desenvolvimento de Sistemas de Informação* (FCA., p. 209).

Oliveira, T., & Martins, M. (2011). Literature Review of Information Technology Adoption Models at Firm Level. *The Electronic Journal Information Systems Evaluation*, *14*(1), 110–121.

Pollock, J. (2008). A Semantic Web Business Case. *W3C*. Retrieved July 14, 2012, from http://www.w3.org/2001/sw/sweo/public/BusinessCase/BusinessCase.pdf.

Rogers, E. M. (1995). *Diffusion of innovations*. (M. B. Salwen & D. W. Stacks, Eds.)*An integrated approach to communication theory and research* (Vol. 65, p. 519). Free Press. doi:10.1525/aa.1963.65.5.02a00230.

Segaran, T., Evans, C., & Taylor, J. (2009). *Programming the Semantic Web*. (M. Treseler, Ed.) (p. 300). O'Reilly.

W3C. (2013). W3C Semantic Web Activity Homepage. Retrieved from http://www.w3.org/2001/sw/