Semantic Web Tools and Decision-making

Francisco Antunes[1,2], Manuela Freire[2] and João Paulo Costa[2,3],

[1] Department of Management and Economics, Beira Interior University, Portugal
[2] INESCC – Computer and Systems Engineering Institute of Coimbra, Portugal
[3] Faculty of Economics, Coimbra University, Portugal
francisco.antunes@ubi.pt; maria-m-freire@telecom.pt; jpaulo@fe.uc.pt

Abstract. Semantic Web technologies are intertwined with decision-making processes. In this paper the general objectives of the semantic web tools are reviewed and characterized, as well as the categories of decision support tools, in order to establish an intersection of utility and use. We also elaborate on actual and foreseen possibilities for a deeper integration, considering the actual implementation, opportunities and constraints in the decision-making context.

Keywords: Semantic Web; Decision-making; Web evolution.

1   Web evolution

Web 1.0 is known as an early stage of the conceptual evolution of the World Wide Web, where users simply acted as publishers and consumers of content, as webpage information was closed to external editing. Rather than a specific technology update or specification, Web 2.0 core was a transformation in the way web pages were made and used, adding a multitude of users responsible for all information management activities.

Traditional group decision-making presents a "top down" approach, usually designed to deliberately guide the interactions of groups in decision-making processes, while in social software users, in the public internet, generate the content and define both the rules and reasons for usage [9], constituting a "bottom-up" approach.

The term Semantic Web [6], considered by many an evolution of Web 2.0 – hence the term Web 3.0 [18] – means a set of technologies that includes ontologies, software agents and rules of logic. These technologies can greatly improve the ability to connect and automatically organize the content of information spread across multiple pages or sites [17]. In this paper, we will make a brief initial review of the general objectives and technologies proposed with the implementation of the Semantic Web, which we will later combine with its actual implementation, opportunities and constraints within the context of decision-making.

2   Semantic Web technology

According to [6], the Semantic Web will enable machines to comprehend semantic documents and data, not human speech and writings. Moreover, the Semantic Web, in

naming every concept using a URI (Uniform Resource Identifier), should express, seamlessly, new concepts that people invent. The base of Web 3.0 for exposing data to applications is the Extensible Markup Language (XML), which lets everyone create their own tags. Scripts, or programs, can make use of these tags in sophisticated ways, but the script writer has to know for what the page writer uses each tag.

Meaning is expressed by Resource Description Framework (RDF), which encodes it in sets of triples that use URIs to name the relationship between things as well as the two ends of the link [6], allowing structured and semi-structured data to be mixed, exposed, and shared across different applications. The resulting linking structure forms a directed, labeled graph, which is the easiest possible mental model for RDF which is often used in easy-to-understand visual explanations.

With SPARQL (a recursive acronym for SPARQL Protocol and RDF Query Language), a query language for RDF data, applications can access native graph-based RDF stores and extract data from traditional databases [16]. SPARQL intends to integrate disparate databases (domain-limited or specific databases – relational, XML, HTML, etc.) so that one query spans (seamlessly and on-the-fly) through several datasets to deliver targeted results [18], also referred as Linked Data.

On the Semantic Web, vocabularies or ontologies define the concepts and relationships (also referred to as "terms") used to describe and classify terms that can be used in a particular application, characterize possible relationships, and define possible constraints on using those terms. The most typical kind of ontology for the Web has a taxonomy and a set of inference rules, which defines classes of objects and relations among them [6]. Web Ontology Language (OWL) and RDF are much of the same things, but OWL is a stronger language with greater machine interpretability than RDF. OWL is built on the top of RDF but comes with a larger vocabulary and stronger syntax than RDF [24], being the basis for implementing inference techniques on the Semantic Web.

Inference rules in ontologies can be characterized by discovering new relationships among terms. Although the computer doesn't truly "understand" any of these relationships, it can manipulate the terms much more effectively in ways that are useful and meaningful to the human user [6]. Inference is also intended to improve data integration and handle possible data inconsistencies on the Web, by seamlessly analyzing data content.

In spite of the earlier vision for a future with Web 3.0 [6], the problem is that a complete re-annotation of the Web is a massive undertaking. As an alternative, many researchers take a very different approach to the Semantic Web. Rather than calling for an overhaul of Web formats, and the massive effort of using Semantic Web tools (not to be expected), they are building software agents that can better understand web pages, as they exist today. Instead of waiting for additional information and for more "machine-understandable" web pages, the alternative is developing improved software agents for Information Retrieval and Natural Language Processing.

Natural Language Processing/Programming (NLP) is a field of computer science, artificial intelligence, and linguistics that regards the interactions between computers and human (natural) languages. NLP and Information Extraction (IE) seek to deduce rules or a domain model out of texts. The knowledge base they hope to extract is frequently designed to drive an expert system or case-based *reasoner* [10] or knowledge-driven decision support systems.

Information extraction identifies specific pieces of information (data) in a unstructured or semi-structured textual document (e.g. a webpage) and transforms unstructured information into a corpus of documents or web pages into a structured database [1].

In artificial intelligence, an intelligent agent (IA) is an autonomous entity, which observes through sensors and acts upon an environment using actuators and directs its activity towards achieving rational goals [27]. Intelligent agents may also learn or use knowledge to achieve their goals, ranging from very simple or very complex (a thermostat is an intelligent agent, as is a human being, as is a community of human beings working together towards a goal, as described in [12]). Nevertheless, the effectiveness of such software agents can only achieve its full potential when more "machine-readable" Web content and automated services (including other agents) become available

3 Decision-making and Semantic Web

The Semantic Web has implications for decision-making support, namely filled and unfulfilled promises derived from the earlier vision of the Semantic Web and research opportunities.

We can accept the categories of decision support tools as [4; 5] established, based on their main objectives: *Personal Decision Support Systems* (PDSS); *Group Support Systems* (GSS); *Negotiation Support Systems* (NSS); *Intelligent Decision Support Systems* (IDSS); *Knowledge Management-Based DSS* (KMDSS); *Data Warehousing* (DW); and *Enterprise Reporting and Analysis Systems* (ER)[1].

[7] stands that the Semantic Web data can be utilized in several ways to process and share information, namely in DSS context: (1) *Information integration*; (2) *Information filtering and selection*; (3) *Information extension, exploration, and explanation*; (4) *Information interpretation, event detection, and prediction*; (5) *Information tracking and post-event analysis*; (6) *Models and model evolution*; and (7) *Sharing decisions*.

Table 1. Intersection of the Semantic Web and Decision Support

<table>
<tr><td rowspan="2"></td><td rowspan="2"></td><td colspan="7">Decision support tools</td></tr>
<tr><td>PDSS</td><td>NSS</td><td>GSS</td><td>IDSS</td><td>DW</td><td>KMDSS</td><td>ER</td></tr>
<tr><td rowspan="7">Semantic Web</td><td>RDF</td><td>1, 3, 7</td><td>1, 3, 7</td><td>1, 3, 7</td><td>1, 3, 7</td><td>1, 3</td><td>1, 3, 7</td><td>1, 3</td></tr>
<tr><td>XML</td><td>1, 3, 7</td><td>1, 3, 7</td><td>1, 3, 7</td><td>1, 3, 7</td><td>1, 3</td><td>1, 3, 7</td><td>1, 3</td></tr>
<tr><td>Ontologies</td><td>3, 5, 6</td><td>3, 5, 6</td><td>3, 5, 6</td><td>3, 5, 6</td><td>3, 5, 6</td><td>3, 5, 6</td><td>3, 5, 6</td></tr>
<tr><td>Inf. rules</td><td>4</td><td>4</td><td></td><td>4</td><td></td><td>4</td><td>4</td></tr>
<tr><td>Query</td><td>2, 3</td><td>2, 3</td><td>2, 3</td><td>2, 3</td><td>2, 3</td><td>2, 3</td><td>2, 3</td></tr>
<tr><td>NLP</td><td></td><td>2</td><td>2</td><td>2, 4</td><td>2, 4</td><td>2, 4</td><td>2, 4</td></tr>
<tr><td>Agents</td><td></td><td>2, 3, 4, 7</td><td>2, 3, 4, 7</td><td>2, 3, 4, 7</td><td>2, 3, 4</td><td>2, 3, 4</td><td>2, 3, 4</td></tr>
</table>

The feature categories span across the different DSS and Semantic Web tools (presented in section 2), even though they are not always present or bear the same

---

[1] Which include enterprise focused DSS, namely executive information systems (EIS), business intelligence (BI), and more recently, corporate performance management systems (CPM).

importance. We can match the utility of each semantic web tool to information processing and sharing against each decision support tool category and the intersection of tools and feature categories is depicted in Table 1.

DSS can be viewed from several different perspectives [4; 25; 29] and we can trace them to Web evolution, according to their intrinsic purposes, as represented in Table 2. It is easy to realize that PDSS are much more related with producing content than disseminating such content, while NSS and GSS naturally involve a multitude of users (even though bearing different objectives). Knowledge-driven and Data-driven DSS can benefit the most from Semantic Web features, as it provides enhanced content relationships with the possibility for greater retrieval accuracy.

Table 2. Web stages and their adequacy regarding decision-making tools

[25]	[4; 5]	Web		
		1.0	**2.0**	**3.0**
Model-driven	PDSS	+ + +	+	+
	NSS	+ +	+ + +	+ +
Communications-driven	GSS	+ +	+ + +	+ +
Knowledge-driven	IDSS	+ +	+	+ + +
	KMDSS	+ +	+ +	+ + +
Data-driven	DW	+ +	+	+ + +
	ER	+ +	+	+ + +

+ + +   *Excellent fit*
+ +   *Adequate fit*
+   *Poor fit*

Regarding the creation of information, and contrarily to the traditional group decision-making "top-down" approach usually involving small groups, Web 2.0 stands for a "bottom-up" approach where information is produced by mass collaboration of people that create, update and share knowledge on a regular basis [14], which constitutes a very distinct approach from PDSS. The use of *folksonomies*, ontologies, software agents and social classification of information relevance (through registered classifications performed by past information users, according to their perceived relevance) provide an opportunity for a larger spectrum of possibilities in searching and recovering relevant information [2]. Compared to ontologies, *folksonomies* offer greater flexibility and adaptability in organizing information and users do not need to agree on a detailed tag hierarchy and taxonomy, though it implies that each user can create a separate set of tags that would then need to be disambiguated, using specific ontologies to be created or a combination with existing ontologies. *Folksonomies* may also suffer from ambiguity regarding the meaning of the tags and lack of semantics, for example, synonyms. Moreover, a coherent categorization scheme when using *folksonomies* can become difficult to achieve, because their contributors do not operate under a centralized controlling vocabulary, though empirical work shows the emergence of stable collective consensus around the categorization of information driven by tagging behaviors [26].

According to the Technology Acceptance Model and its extensions [31], and in spite of the fact that people seamlessly create and disseminate information through social media, the intention to individually add any further annotations to contents seems compromised (at least until they have better tools to do so). Thus, the use of software agents and NLP seems appropriate to perform an automatic processing of the dynamic

and massive amount of information encompassed in social media at least until technology takes full advantage of *folksonomies*.

The unstructured nature of decision-making, especially its early stages [30], is very well suited for the *ad-hoc* nature of social networking. During the intelligence phase, in spite of the fact that Semantic Web cannot be viewed as a decision-modeling technology to improve decision *per se*, it can be seen as an enhanced possibility for integrating data [20] and revealing implicit information than usually would remain undiscovered, thereby resulting in sub-optimal decisions [21].

Semantic Web technologies can be exploited to the advantage of DSS, namely by applying Information Extraction (IE) to populate Semantic Web datasets and to perform the automatic detection of arguments within group discourse (and from external data), for later analysis by a DSS. Interconnecting users' contributions would enrich and produce a much more accurate information to be used in the intelligence phase. Nevertheless, the creation of *folksonomies* lack tools that can make this a seamless work (or at least very simplified or intuitive), making it a time-consuming task.

During the design phase, structured versions of a group discourse allow a better understanding of the expressed points of view. However, Social Media does not favor this latter type of structure or the generation of tags that can explicitly define used concepts, applied values or any types of quantitative or qualitative parameters. Accordingly, the use of XML/RDF to structure the produced meeting content could alleviate this problem. Nevertheless, the generation of *tagged* content, which would be of enhanced utility in decision-making, requires computer skills that cannot be expected from all participants in all decision meetings and, as in the intelligence phase, the creation of such structured content also constitutes a time-consuming task.

There is a need for ontologies that are suitable for representing *informal* Social Web arguments and ontologies that map between the social world and the argumentative world [28]. Nevertheless, Social Media are understood as failing the criterion of "argumentative discussions", as the argumentation support of general Web 2.0 tools is considered to be peripheral [28]. The writing style commonly used in these platforms has a pattern out of the ordinary that sometimes makes it incomprehensible to those who are not part of the conversation and/or culture/context, thus making it very hard to make it "machine-understandable"[8; 15]. Another problem (described in [19]) is the fact that a dialogue can be written in more than one language (code-mixing[2]). As users can also omit much of the speech, this means that data is possibly tangled, incomplete and sometimes error-prone. Even harder to grasp, are the artifices of language, which help to define how these interactions and respective arguments do come out. Herein lays the challenge to achieve its capture in order to be used by "machines". One way of doing this is by using formal models that capture arguments and convert implicit (concealed in discussions) to explicit knowledge [22].

In spite of the earlier considerations and knowing that RDF triples consist of text encompassing relations between described entities, we can argue that Semantic Web tools will be able to transform the representation of a simple (unstructured) text into a representation that follows or is supported by one or more argumentation models. Such process would follow: (1) the establishment/extraction of a taxonomy of elements contained in the text/speech; (2) the development of a specific ontology or use of

---

[2] Where lexical items and grammatical features from two languages appear in a sentence.

existing ontologies to relate the elements included in the taxonomy; (3) the development of ontologies according to the intended argumentation models; (4) the combination of steps (2) and (3).

The revelation of implicit attributes or argumentative properties could be achieved by IE/NLP techniques that could also build and associate different ontologies containing rules of the argumentative association derived from semantic terms (e.g. terms such as "in support of", "against", "in favor of", etc.). These processes combine the ease of use of Social Media for presenting, discussing and narrowing ideas (Intelligence and Design phases), while using Artificial Intelligence (AI) tools (IE/NLP in particular) to structure the produced content[3] and, thus, leading to the choice phase. This would be done by enabling a richer and more structured visualization of the speech[4], namely by presenting the information according to different models of argumentation.

In the choice phase, many DSS applications use ontologies and rules as a means for making the DSS "intelligent" by adopting the emerging Semantic Web standards for knowledge representation [7]. According to [13], the use of ontologies can facilitate collaboration, by providing a unifying multiple-criteria decision analysis/aiding (MCDA) decision knowledge skeleton that can be used as a common and shared reference for a collaborative process. In addition, the deployment of Service Oriented Architectures (SOA), enhanced by Semantic Web technologies for sharing and accessing data, can apply Semantic Web technologies in peer-to-peer networks, for facilitating offers in negotiation scenarios [11].

Semantic Web tools, namely ontologies, could also be applied to provide a follow-up on decisions after they are taken. This could become an excellent source for decision reconstruction [3] and evaluation of the implemented choice. Unfortunately, the pervasiveness of ontologies in the Web, is not yet a reality, as their creation involves a top-down process, which constantly requires disciplinary experts checking the evolution of the ontologies [23].

4   Final remarks

We concluded the greatest obstacle to actual arrival of the Semantic Web mostly relies on the technologies that have to come together in order to make it a seamless. Nevertheless, Semantic Web and decision-making possess many connection points, namely regarding decision-making phases, which are worthy to explore and develop (Table 3).

Some argue that it is unrealistic to expect busy people and businesses to create enough metadata to make the Semantic Web work. The simple tagging used in Web 2.0 applications lets users spontaneously invent their own descriptions, which may or may not relate to anything else. However, the solution to this problem may simply rely on better tools for creating metadata, like the blog and social-networking sites that have made building personal websites easy.

The first step towards a real Semantic Web-based decision-making environment is making data accessible through queries. The second step towards Semantic Web-based

---

[3] Even though manual/human intervention is expected at some extent.
[4] For which visualization analytics and tools are complementary to Semantic Web tools.

decision-making seems to be ontology mapping, as the amount of public available ontologies increases steadily and as the Semantic Web grows (even some argue its rhythm is not fast enough).

Table 3. Semantic Web and the decision-making process

<table>
<tr><th colspan="5">Decision process</th></tr>
<tr><th></th><th>*Intelligence*</th><th>*Design*</th><th>*Choice*</th><th>*Implementation and evaluation*</th></tr>
<tr><td>*Opportunities*</td><td>- Data integration and interoperability</td><td>- Enhanced structuring and argument representation of collaborative discourse</td><td>- Collaborative MCDA
- Facilitating offers in negotiation scenarios</td><td>- Follow-up on decisions</td></tr>
<tr><td>*Constraints*</td><td>- Requires specialized computer skills
- Needs seamless tools</td><td>- Requires specialized computer skills
- Perception of utility
- Needs seamless tools</td><td>- Requires specialized knowledge</td><td>- Ontologies are not web-pervasive
- Requires specialized knowledge</td></tr>
</table>

It is easy to understand that a broader or generalized Semantic Web integration in the decision support community is still a work in progress.

5   References

1. Acharya, S and Parija, S, The Process of Information Extraction through Natural Language Processing. International Journal of Logic and Computation (IJLP), 1(1):40-51, (2010)
2. Antunes, F and Costa, J P, Integrating decision support and social networks. Advances in Human-Computer Interaction, 2012:1-10, (2012)
3. Antunes, F and Costa, J P, A decision reconstruction support model. International Journal of Information Technology & Decision Making, 12(2):309-333, (2013)
4. Arnott, D and Pervan, G, A critical analysis of decision support systems research. Journal of Information Technology, 20:67-87, (2005)
5. Arnott, D and Pervan, G, Eight key issues for the decision support systems discipline. Decision Support Systems, 44:657–672, (2008)
6. Berners-Lee, T; Hendler, J and Lassila, O, The semantic web. Scientific American:34-43, (2001)
7. Blomqvist, E, The Use of Semantic Web Technologies for Decision Support - A Survey. Semantic Web, Pre-press:1-25, (2013)
8. Bodomo, A B, Computer-Mediated Communication for Linguistics and Literacy: Technology and Natural Language Education. Information Science Reference, Hershey, New York, (2010)
9. Boyd, D M, The significance of social software. In: Burg T N and Schmidt J (eds) BlogTalks Reloaded. Social Software - Research & Cases., Books on Demand GmbH, Norderstedt, Germany, pp 15-30, (2006)
10. Cowie, J and Lehnert, W, Information Extraction. Communications of the ACM, 39(1):80-91, (1996)
11. Du, T C, Building an automatic e-tendering system on the Semantic Web. Decision Support Systems, 47(1):13-21, (2009)

12. Franklin, S and Graesser, A, Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In: Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Budapest, Hungary, Springer-Verlag, August 12-13, pp 21-35 (1996)
13. Gašević, D; Djuric, D and Devedžic, V, Model driven engineering and ontology development. Springer, Berlin, (2009)
14. Gehrke, N and Wolf, P, Towards Audit 2.0 – A Web 2.0 Community Platform for Auditors. In: Proceedings of the 43rd Hawaii International Conference on System Sciences (CD-ROM), Kauai, Hawaii, USA, Computer Society Press, January 05-8, (2010)
15. Georgalou, M, "Pathfinding" Discourses of Self in Social Network Sites. In: Taiwo R (ed) Handbook of Research on Discourse Behavior and Digital Communication: Language Structures and Social Interaction, IGI Global, Hershey, New York, pp 39-62, (2010)
16. Hendler, J, Web 3.0 Emerging. Computer, 42(1):111-113 (2009)
17. Kousetti, C; Millard, D E and Howard, Y, Study of Ontology Convergence in a Semantic Wiki. In: Proceedings of the WikiSym08 2008 International Symposium on Wikis Porto, Portugal, September 08-10, (2008)
18. Lassila, O and Hendler, J, Embracing "Web 3.0". IEEE Internet Computing, 11(3):90-93, (2007)
19. Muysken, P, Bilingual Speech: A Typology of Code-mixing. Cambridge University Press(2000)
20. Necula, S-C, Discussions on semantic-based in decision support systems. In: Proceedings of the 5th European conference on European computing conference (ECC'11), Paris, France, World Scientific and Engineering Academy and Society (WSEAS), April 28-30, pp 193-197 (2011)
21. Necula, S-C, Implementing the Main Functionalities Required by Semantic Search in Decision-Support Systems. International Journal of Computers, Communications & Control (IJCCC), 7(5):907-915, (2012)
22. Nonaka, I and Takeuchi, H, The Knowledge-Creating Company. Oxford University Press, New York, (1995)
23. Padula, M; Reggiori, A and Capetti, G, Managing Collective Knowledge in the Web 3.0. In: Society I C (ed) Proceedings of the INTERNET 2009, First International Conference on Evolving Internet, Cannes/La Bocca, French Riviera, France, August 23-29, (2009)
24. Pattal, M M I; Yuan, L and Jianqiu, Z, Web 3.0: A real personal Web! More opportunities & more threats. In: Proceedings of the Third International Conference on Next Generation Mobile Applications, Services and Technologies, Cardiff, Wales, UK, IEEE Computer Society Press, September 15-18, (2009)
25. Power, D J; Sharda, R and Kulkarni, U R, Understanding Decision Support Systems for Global Enterprises. In: Decision Support for Global Enterprises, Springer US, pp 3-14, (2007)
26. Robu, V; Halpin, H and Shepherd, H, Emergence of consensus and shared vocabularies in collaborative tagging systems. ACM Transactions on the Web (TWEB), 3(4):Article 14, (2009)
27. Russell, S J and Norvig, P, Artificial Intelligence: A Modern Approach. Pearson Education, Inc., Upper Saddle River, New Jersey, (2010)
28. Schneider, J; Groza, T and Passant, A, A review of argumentation for the Social Semantic Web. Semantic Web, 4(2):159-218, (2013)
29. Turban, E; Aronson, J E and Liang, T P, Decision Support Systems and Intelligent Systems. Pearson Education Inc., New Jersey, (2005)
30. Turban, E; Liang, T-P and J.Wu, S P, A Framework for Adopting Collaboration 2.0 Tools for Virtual Group Decision Making. Group Decision and Negotiation, 20:137-154, (2011)
31. Venkatesh, V and Davis, F D, A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science, 46(2):186-204, (2000)