

The Role of Linked Data and Semantic-Technologies for Sustainability Idea Management

Alain Perez¹, Felix Larrinaga¹, Edward Curry²

¹Mondragon Unibertsitatea, Arrasate-Mondragon, Spain

{aperez, flarrinaga}@mondragon.edu

²Digital Enterprise Research Institute, National University of Ireland, Galway, Ireland

ed.curry@deri.org

Abstract. Idea Management Systems (IMS) manage the innovation lifecycle from the moment of invention until ideas are implemented in the market. During the lifecycle the IMS supports collaboration, allows idea enrichment with comments, contextual data, or connected to other relevant ideas. Semantic technologies can improve the knowledge management capabilities of IMSs allowing relevant information to be easily linked to ideas.

Many Enterprises concerned with sustainability encourage employee's participation as a means to boost creative innovation within their Sustainability Initiatives. However little work has examined the role of an IMS within Sustainability. In this paper we analyse the impact of a semantic-enabled IMS within a sustainability innovation process. In particular, how ideas can be enriched with contextual Linked Open Data (LOD), especially Life-Cycle Assessment (LCA) data, to improve the understanding, implication and value of the idea from the sustainability perspective.

Keywords: Idea Management Systems, Semantic Web, Linked Data, Sustainability, Life-Cycle Assessment.

1 Introduction

Sustainability is the responsible management of resources encompassing the triple bottom line of environmental, economic, and social dimensions. Many organisations are starting to make serious commitments towards incorporating sustainability into their own organizational logics [1] to maximise profits in an environmentally and socially responsible manner. Sustainability is not only about Corporate Social Responsibility, Sustainability is an important business issue, affecting new products and services, compliance, cost reduction opportunities, the organization's reputation, and revenue generation often derived from technological innovation [2]. Porter recognises the role sustainability can play as part of an organization's Competitive Strategy with the concept of "innovation offsets" – where companies can "not only lower the net costs of meeting environmental regulations, but can lead to absolute advantages" over competitors [2].

Sustainability requires information on the use, flows and destinies of energy, water, and materials including waste, along with monetary information on environment-related costs, earnings, and savings. This type of information is critical if we are to understand the causal relationships between the various actions that can be taken, and their impact on sustainable performance.

Innovation is key to articulate knowledge management by means of effective processes and methodologies. The phase of the innovation process where idea management is developed is one of the most critical stages [4]. IMSs support this stage providing the necessary tools to collect, enrich, store, present and select ideas. IMS manage ideas through their life-cycle from the time of creation until they are selected for implementation. During this life-cycle it is crucial to gather as much relevant information as possible in order to collect quality-relevant ideas. Users can enrich ideas with opinions, other ideas and additional content. This task can be cumbersome and its automation is fundamental.

This paper aims to define how sustainability in enterprises can benefit from IMS. The hypothesis is that more precise relations among ideas and richer content can be automatically achieved if Semantic Web and Linked Data technologies are employed in IMS, linking sustainability ideas with data from different data sources.

Managing innovation for sustainability needs to address some major challenges; the emergence of radical new technologies and markets, constant shift in the regulatory conditions, the involvement or participation of many agents, the large volume of ideas for screening and evaluation, and in particular the need to acquire, assimilate and exploit new knowledge [5].

An increasing number of organizations worldwide have adopted innovation contests not only for innovation purposes, but also for other reasons such as promoting sustainability [6]. A proof to this can be seen in the annual reports and sites of several energy providers and enterprises. In some of these experiences IMS have proven beneficial. IMS provide the workflow tools necessary to launch and manage innovation contest or waves, a common platform where different agents can collaborate, a repository where ideas are gathered and tools for editing, commenting or voting upon ideas. IMSs also encourage collaboration among people and enterprises.

One of the biggest problems in IMS is the difficulty of enriching these ideas. It is a manual and time consuming task that includes the searching and gathering of additional knowledge in different sources. Most of the time ideas are not linked with other data the enterprises may have in their systems or data available outside the company [7]. This causes disinformation and generates duplicates or poor quality ideas. A system that links generated ideas to stored data automatically may be an improvement. Semantic Web and Linked Data technologies propose a set of good practices to publish and link structured data in the Web. Many datasets and repositories are already available adopting this philosophy enabling machines in the “understanding” of the data they store.

2 The Role of Knowledge Management & Idea Management in Sustainability

2.1 Reference Architecture

IMS are software platforms or applications that provide repositories and tools to gather, search for, edit, comment, and vote upon ideas. They accommodate different idea contests or campaigns and stakeholders; idea operators, experts, decision makers and administrators. Dependencies between ideas from different campaigns must be addressed. A standard IMS architecture is shown in Fig. 1.

IMS must provide workflows to manage the interactions between the different innovation process stages: Idea generation, analysis, enrichment, selection, development and implementation.

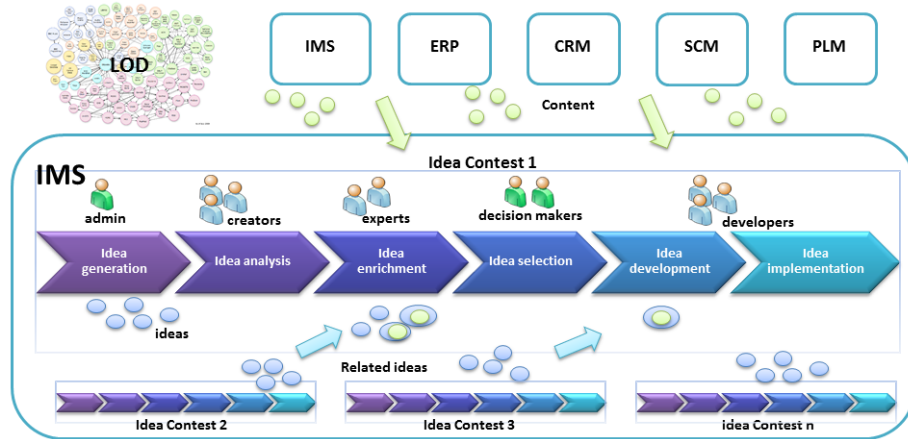


Fig. 1. IMS Architecture.

2.2 Linked Data and Linked Open Data (LOD)

Emerging from research into the Semantic Web, Linked Data proposes an approach for information interoperability based on creation of a global information space. Linked data leverages the existing open protocols and standards of the World Wide Web architecture for sharing structured data on the web. Linked data technology uses web standards in conjunction with four basic principles for exposing, sharing and connecting data. These principles are:

1. Use URIs as names for things.
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information using the standards.
4. Including links to other URIs so that people can discover more things.

Linked Data is facilitating the publishing of large amounts of structured data on the web. The resulting Web of Data can be considered as a web scale data space support-

ed by Semantic Web technologies. The Linked Open Data represents a large number of interlinked datasets that are being actively used by industry, government and scientific communities.

2.3 Sustainability IMS and Semantic Web/Linked Data

A common problem with IMS is idea assessment due to data overflow, noisy data, bursty nature of idea contests and difficult in rating innovation [8]. This translates into difficulties such as the detection of similar or duplicated ideas. Several studies propose the application of Semantic Web technologies on the innovation process and more specifically for IMSs to overcome these issues[9][10]. The GI2MO ontology, for instance, is a project that tries to improve current Idea Management Systems by offering an ontology that models the innovation process. The ontology lays foundations for knowledge management based on interlinking of enterprise systems and web assets to increase information awareness and help in innovation assessment [11]. Although GI2MO ontology provides coverage for most of the properties included in IMS, it lacks of explicitly capturing the contextual knowledge for the idea.

To boost interoperability among heterogeneous systems, some IMS platforms present ideas in RDF format and work as SPARQL endpoints so third parties can place idea or innovation process related queries [12]. In the reference architecture Sustainability datasets will be interlinked to IMS as external knowledge to enhance IMS content.

3 Examples of Ontologies for Use Cases

This section describes 3 different sustainability use cases that can benefit from the linking of ideas with external information. The first use case aims to enrich ideas for energy reduction. The second one addresses products life cycle and how data can be linked. The last use case shows how similar ideas can be identified helping administrators management tasks.

Energy Reduction: Imagine a user involved in an idea contest oriented to sustainability that proposes a new idea (graphically on Fig. 2.): “*I would change the incandescent bulb in desk #333 for a LED bulb in order to save energy.*”

The system could identify the concepts “*incandescent bulb in desktop #333*” and “*LED bulb*” find on the data space (external to the IMS) information about the bulb in that desktop and led bulbs and show it in a widget or block next to the idea. If the system is able to identify the domain of the idea and annotate it semantically, the idea can be linked to data stored in a data space [13] (results on grey and blue widgets in figure 2) or searched on some data sources [14]. That way, the user would create an idea with automatically added information. If someone reads the idea will know if it is worth the effort of changing it or not. The reader could comment on the idea and discuss about it.



Fig. 2. Energy Reduction and LCA ideas examples (automatically added data in blue and grey widgets).

Life-cycle Assessment: This use case links IMSs with LCA data stored in a data space. Imagine a user concerned about the Greenhouse Gas emissions (GHG emissions) discovers that new laptops are going to be bought. He could write the following idea (graphically in Fig. 2.):

“If we buy 13-Inch MacBook Air laptops instead of 13-Inch MacBook Pro, we can reduce the amount of GHG emissions in the manufacture of our devices.”

If we can identify that the idea talks about 2 different laptops and their GHG emissions, we could link the idea with that data and show the amount of GHG emissions each laptop has and the savings of the idea.

In order to link the data we have to annotate it semantically, for example using The Resource Description Framework (RDF). In RDF, the statement “LCA Idea mentions MacBook Air” is expressed in triple format as:

(Subject - “LCA Idea”) > (Predicate - “mentions”) > (Object - “MacBook Air”)

Using this semantic annotations some links can be found between different data in the system. That data can be found in the IMS or in some internal and external data spaces. Below, on Fig. 3, a graphical representation of those links can be seen.

Taking the LCA use case as an example, we can see the different links the data have (Fig. 3). Using these links we can extract some data and see what is the actual impact of the idea. For example, we can see how much greenhouse gas each mentioned laptop emits to the atmosphere on their manufacturing process.

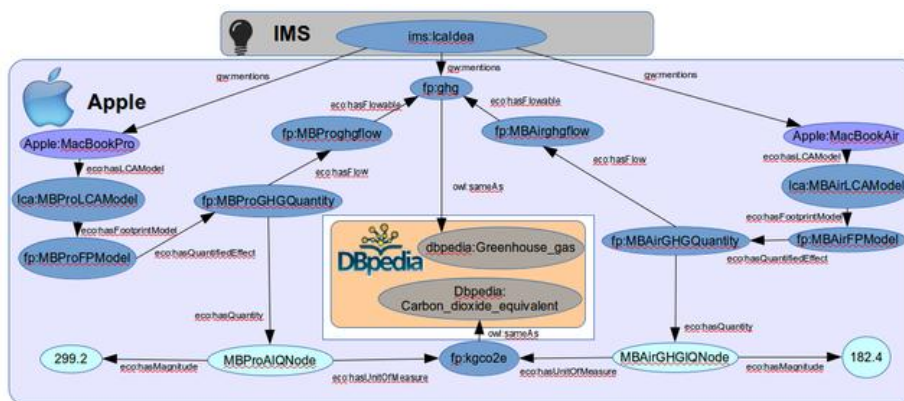


Fig. 3. Data links example.

Similar ideas recognition: We can imagine an enterprise has an innovation process that generates a wave every year in order to collect new ideas. Sometimes ideas can be repeated in the same wave or in past waves. Having a system that identifies similar ideas can help innovation administrators in identifying those relations or knowing why an idea was rejected in the past.

For instance in a previous wave the idea of changing an incandescent bulb for a led one was rejected because the led bulb was too expensive. If someone generates a similar idea it can be linked to the previous one, and see the reasons for the previous rejection. If now the led bulbs are cheaper, that idea may be interesting.

In order to perform that task the identification of main concepts of the idea is needed. If the system can find similarities in those main concepts, the ideas should be linked. The idea with mentioned concepts can be compared with the other ones and similar ideas can be identified and presented to the users.

4 Potential Benefits

Firstly, IMS ideas would be enriched automatically with relevant- specific data provided by LOD/SW repositories. Sustainability impact will be included among that content. This could help users and administrators to see how important an idea could be. Secondly, the tools should help users in order to understand the context of the ideas. With additional data and related information users could understand some issues that the idea itself does not explain explicitly. And finally, the tools are expected to help administrator to perform their task faster and in an easier way. On one hand showing them the relations of the ideas can help them identifying in what are the users concerned and see if there are repeated ideas. On the other hand, measuring their possible impact could help them recognizing the most important ideas and select the best ones, helping even more if there is a big amount of data.

5 Summary and Future Challenges

The main innovation strives in the application of semantic web and LOD technologies to interlink Sustainability repositories with IMS in such a way that ideas are enriched with relevant content. SW technologies are also proposed in the detection of similar ideas. Some use cases have been defined describing the advantages semantic annotations and linked data might bring. Moreover, some possible benefits for users and system administrators have been outlined. Linking ideas with related data can be a powerful approach, but it can be a difficult task for the user. A future challenge can be identifying the data automatically in order to help users linking the idea. The system could make some recommendations based on what the user has written. Another challenge could be helping administrators developing some additional visual tools in order to visualize data or idea relations. These visual tools can be very helpful, especially with large amounts of data. Identifying the domain of the idea can be helpful too if additional information is wanted to be added. Knowing the domain of the idea can help recognizing specific data sources where more data can be found and linked. Fi-

nally, some case studies should be implemented in order to obtain results that validate the hypothesis of this paper.

6 References

1. Curry, E., Guyon, B., Sheridan, C. & Donnellan, B.: Developing a sustainable IT capability: lessons from Intel's journey. In: *MIS Quarterly Executive*, vol. 11, no. 2, pp. 61-74. (2012)
2. Watson, R. T., Lind, M., and Haraldson, S.: The Emergence of Sustainability as the New Dominant Logic: Implications for Information Systems. In: *International Conference on Information Systems, ICIS 2012*. (2012)
3. Porter, M. E., Linde, C. V. D.: Toward a New Conception of the Environment-Competitiveness Relationship. In: *Journal of Economic Perspectives* (9:4)JSTOR, pp. 97-118. (1995)
4. Errasti, N., Santos, I., Lizarralde, O.: Social Software in Support of Collaborative Innovation. In: *Erima*. (2010)
5. Seebode, D., Jeanrenaud, S., Bessant, J.: Managing innovation for sustainability. In: *R&D Management*, vol. 42, no. 3, pp. 195-206. (2012)
6. Adamczyk, S., Bullinger, A.C., Möslin, K.M.: Innovation Contests: A Review, Classification and Outlook. In: *Creativity and Innovation Management*, vol. 21, no. 4, pp. 335-360. (2012)
7. Curry, E., Hasan, S., Hassan, U. ul, Herstand, M., O'Riain, S.: An Entity-Centric Approach To Green Information Systems. In: *Proceedings of the 19th European Conference on Information Systems (ECIS 2011) Helsinki, Finland*. (2011)
8. A. Westerski, A.: Semantic Technologies in Idea Management Systems: A Model for Interoperability, Linking and Filtering. In: *PhD Thesis*. (2013)
9. Riedl, C., May, N., Finzen, J., Stathel, S., Kaufman, V., Krcmar, H.: An Idea Ontology for Innovation Management. In: *International Journal on Semantic Web and Information Systems*, vol. 5, no. 4, pp. 1-18. (2009)
10. Lorenzo, L., Lizarralde, O., Santos, I., Passant, A.: "Structuring E-Brainstorming to Better Support Innovation Processes. (2011)
11. Westerski, A., Iglesias, C.A., Rico, F.T.: A model for integration and interlinking of idea management systems. In: *4th International Conference on Metadata and Semantic Research, MTSR 2010, October 20, 2010 - October 22 Springer Verlag, Alcala de Henares, Spain, 2010*, pp. 183. (2010)
12. Perez, A., Larrinaga, F., Lizarralde, O., Santos, I.: INNOWEB: Gathering the context information of innovation processes with a collaborative social network platform. In: *International Conference on Concurrent Enterprising (ICE)*. (2013)
13. Curry, E., Hasan, S., O'Riain, S.: Enterprise Energy Management using a Linked Dataspace for Energy Intelligence. In: *The Second IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT 2012) Pisa, Italy: IEEE*. (2012)
14. Curry, E., O'Donnell, J., Corry, E., Hasan, S., Keane, M., O'Riain, S.: Linking building data in the cloud: Integrating cross-domain building data using linked data. In: *Advanced Engineering Informatics* (27:2), pp. 206-219. (2013)