

LinksTo – A Web2.0 System that Utilises Linked Data Principles to Link Related Resources Together

Owen Sacco¹ and Matthew Montebello¹,

¹ University of Malta, Msida MSD 2080, Malta.
{osac001, matthew.montebello}@um.edu.mt

Abstract. Although social sharing websites are currently employing Semantic Web techniques to structure the data, these websites still stand in isolation since most of the data is not linked. Therefore, this paper proposes a prototype system called *LinksTo* that provides users the functionality to link resources from the Web at large. These links are described in RDF adhering to vocabularies recommended by the linked data best practices. The links described in RDF are transparent to the user. However, the system also provides functionalities to access the RDF data that can be utilised by the linked data community.

Keywords: Semantic Web, Linked Data, Web2.0, RDF, Ontologies, FOAF, SIOC, SKOS.

1 Introduction

With the advent of Web2.0, many social collaborative platforms have emerged providing users the functionality to share information in a personal and collective manner. Most of these social platforms consist of wikis, blogs, social bookmarking websites, photo sharing websites and video sharing websites. The majority of these community systems also provide the functionality to organise and describe the content by means of a lightweight knowledge representation called folksonomy that consists of describing the act of tagging Web resources. Although some of these Web applications are utilising Semantic Web technologies to add more meaning and structure to the data, most of the data is not structured using linked data principles and practices [1]. Moreover, even though Semantic Web technologies are being exploited in current collaborative knowledge sharing web sites, the data of such web systems is still not linked and therefore the data is still isolated limiting the aspect of collaborative knowledge sharing amongst web systems. Therefore, this necessitates for a system to allow users to link resources from diverse web systems and such links are described according to the linked data best practices.

The aim of this paper is to propose a prototype of a Web2.0 system called *LinksTo* that utilises Semantic Web technologies whereby various resources can be linked and such links are described in RDF conforming to vocabularies recommended by the linked data best practices. LinksTo provides an interface for users to collect and link resources that are related to a particular topic the user is searching on that would collectively form a collaborative sharing of information.

This research paper is organised as follows: in section 2, a brief discussion about related work is provided. Section 3 provides a brief explanation of the functionality and the technical aspects of the LinksTo system. In section 4 a concluding note is provided that summarises this research paper.

2 Related Work

Social resource sharing systems consist of web-based platforms that provide the users to publish, share and manage resources. Such systems have attracted a number of users because these platforms do not require any specific expertise. One of the most popular social resource sharing systems are the social bookmarking websites such as delicious¹ and BibSonomy[5] that assist users to save, share and tag URLs of resources for later retrieval. Resources which are bookmarked in these social websites are organised and indirectly linked by means of tags. Linking of resources is achieved by assigning the same or similar tags to each resource's URL. Therefore this implies that if users want to link resources, the users have to save related resources one at a time rather than as a bundle of resources and also have to assign the same tag(s) for each individual resource. This could be a daunting task if the user wants to link a large number of resources. Moreover, current social bookmarking websites do not provide the functionality to describe linked resources and their assigned tags in RDF adhering to the linked data best practices. Hence, LinksTo is designed to extend the idea of social bookmarking websites by providing users the functionality to save multiple resources at one instance, assign tags to the collection of these resources and describe in RDF the linked resources together with their assigned tags.

Apart from social bookmarking websites, Baeza-Yates and Tiberi [3] propose how to extract semantic relations from query logs. The authors explain how query logs can be presented as a folksonomy whereby user queries act as tags assigned to documents clicked by the same user after the query result was retrieved. This method shows positive results for information retrieval, however, this method does not take into consideration that clicked documents may not be relevant to what the user requires. It is the norm that the relevance of a document to the topic the user is searching on is known after the user has examined the document. Therefore, relying on clicked documents is not sufficient in order to link related resources. Another similar approach is proposed in [6] whereby the authors term logsonomy as a folksonomy for web search engines. Their approach is to define a folksonomy for the relation between a query, a clicked document and the user. However, since this approach is also based on clicked documents and disregards the relevance of the content of a document is to a query, this approach does not yield accurate results when linking resources.

An approach that is similar to the LinksTo application is the GroupMe! system [7]. GroupMe! provides a Web2.0 interface for users to group resources and also to tag such resources. The GroupMe! system also describes in RDF the grouped resources and their associated tags. Moreover, this system extends the tripartite folksonomy tuple defined in [5] by adding a concept called a group that denotes the grouped

¹ <http://delicious.com/>

resources. In order for the GroupMe! system to describe the grouped resources in RDF based on this extended folksonomy, a GroupMe! specific ontology is used. The problem with this specific ontology arises when external applications want to make use of the RDF data. This is because when applications use diverse system specific ontologies, application developers implementing systems that consume RDF data have to be aware of all these ontologies in order to exploit the RDF data. If the application is not aware of the specific ontology in use, then the application will not be able to parse such RDF data. In fact, in order to publish data on the Web, the authors in [2] recommend that vocabularies such as Semantically Interlinked Online Communities (SIOC), Friend-Of-A-Friend (FOAF) and Simple Knowledge Organisation Systems (SKOS) are used wherever possible in order to simplify client applications to process the data. Therefore, LinksTo models its data on the tripartite folksonomy model defined in [5] rather than extending it and uses the SIOC ontology to describe in RDF the linked resources. This allows the data to conform to the linked data practices allowing other applications to utilise the data without the need to conform to any system specific ontology. Furthermore, LinksTo provides other features that are not present in GroupMe! such as: users can follow other user's linked resources; users can edit tags when editing linked resources; users can link their user profile to other social network profiles of the same user; and the user profile is described in RDF using the FOAF vocabulary which can be exported to an external file or linked from other Web systems.

3 The LinksTo System

3.1 A Web Search Scenario – LinksTo Motivation

A user is searching on a particular topic and the web search engine retrieves many resources that are related to that specific topic. However, only some of the resources are important and relevant for the user. Once the user has decided which resources are relevant, the user would then desire to save all the significant URLs for later retrieval. Since the resources are all related to the topic the user was searching on, the user would require linking the resources and tagging this collection of linked resources with keywords that describe best that specific topic. Therefore, the user adds all the relevant URLs to a space that links these resources and such space also provides the user with the functionality to assign tags to the linked resources. The user or any other user(s) that eventually would want to search on that same topic can query the system by using tags; and the linked resources tagged with the exact or similar tags can be retrieved. With this approach, the user searching for resources will be presented with the most relevant resources and the user does not need to filter out unrelated resources. Thus, the user only searches for a desired topic and the linked resources are displayed as the search results.

3.2 LinksTo User Features and Functionalities

In a nutshell, the LinksTo system provides the following user functionalities:

1. Creation of user profile. LinksTo provides users to create a profile that can be linked to other profiles of the same user created in other social network systems. Moreover, if the user has a FOAF file, this can be uploaded by the LinksTo system. Once the FOAF file is uploaded, the system parses such FOAF file and retrieves any data that is required within the system.

2. Creation of Web resources *Links*. LinksTo provides users the functionality to create links amongst various Web resources and collect them in a *Link*. This functionality also extracts any available tags assigned to the resources found in delicious. Once the tags are extracted, the system adds them to a *Link* tag cloud which consists of a set of tags assigned to a particular *Link*. The system also allows users to add or remove tags from the *Link* tag cloud. Once a *Link* is saved, the Web resources links are described in RDF using the FOAF, SIOC and SKOS vocabularies and the data stored in persistent storage. Fig. 1 depicts a screenshot of the creation of *Links*.

3. Searching mechanisms for Web resources and *Links*. LinksTo provides searching capabilities for Web resources using the Google search API. When a Web resource is selected after retrieved by the search engine, *LinksTo* offers the option to preview the Web resource within the system. Moreover, LinksTo provides searching functionality using tags as queries to search for *Links* within the system.

4. Exporting RDF data. The system provides users to export RDF data to an RDF file. Moreover, the system also provides application clients to request for RDF data in the form of HTTP requests and the system sends back the linked resources structured in RDF. This is convenient for applications that utilise Semantic Web technologies since such applications can make use of the LinksTo's data without the need to re-format the data.

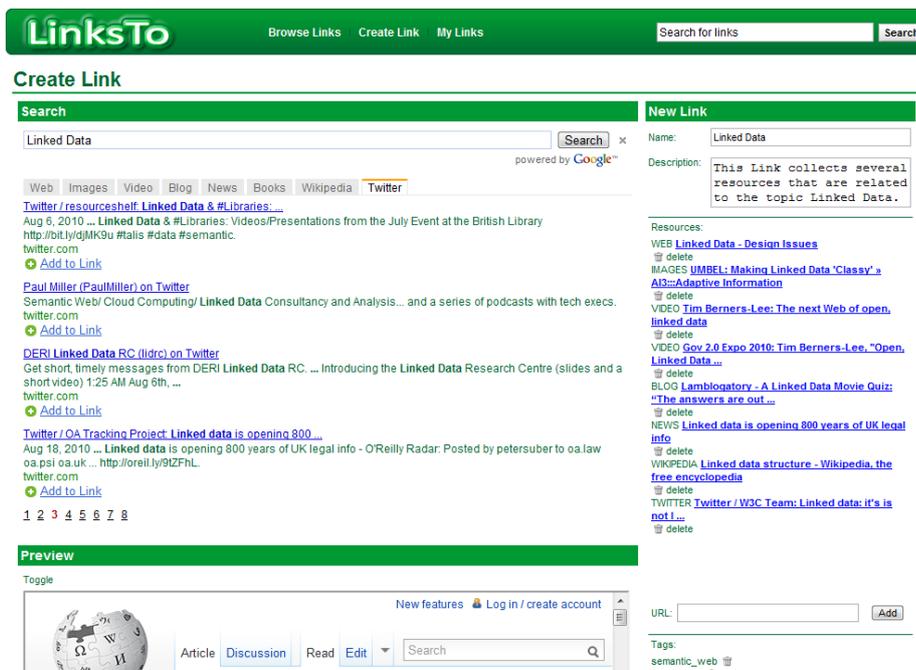


Fig. 1. A screenshot of LinksTo creation of a *Link*

3.3 LinksTo Architecture – A Technical Overview

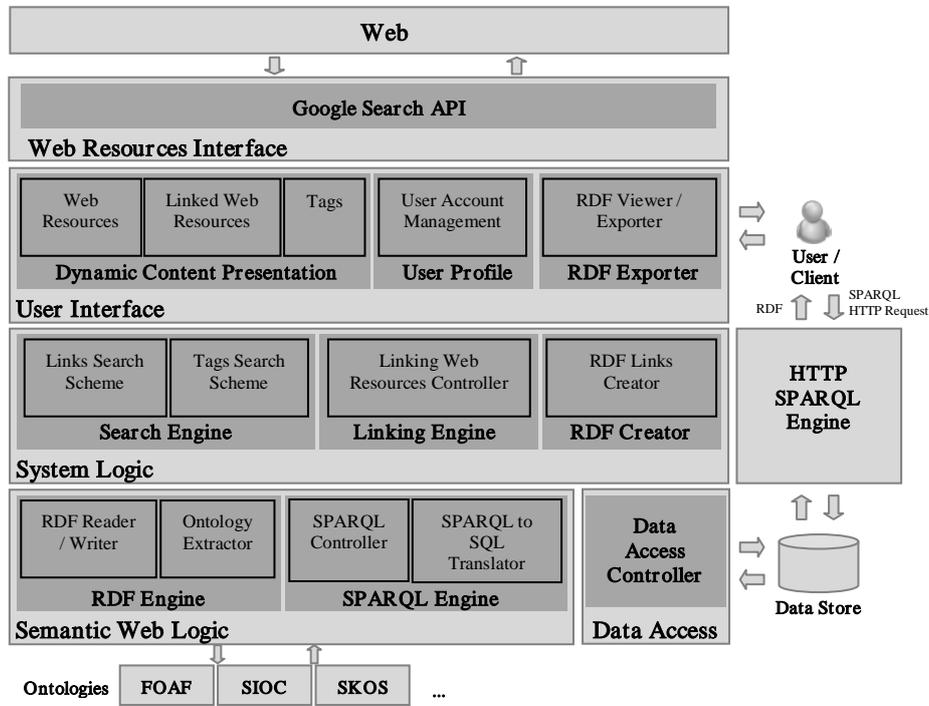


Fig. 2. LinksTo Technical Architecture

LinksTo technical architecture consists of various specific functions as illustrated in Fig. 2. These functions are briefly explained below.

Web Resources Interface. This module interacts with the Web resources located on the Web. This module contains the Google search API that retrieves search results of Web resources using the Google search engine. Moreover, this module also retrieves tags (if available) assigned to the resources from within social bookmarking websites such as delicious.

User Interface. This module provides a user interface for: 1) searching Web resources on the Web; 2) previewing selected Web resources from within the same application; 3) to create *Links* and tag such *Links*; 4) to submit search queries to search for *Links*; and 5) an interface to manage user account information. The User Interface uses the Prototype framework to perform AJAX calls to retrieve data.

System Logic. The system logic module controls most of LinksTo technical functionality. This module contains the *Links* searching engine that consists of the mechanism to query and retrieve the relevant *Links* which are the closest to the user's query terms. The searching engine makes use of a ranking procedure called FolkRank [4] to rank the tags assigned to the *Links* and the ones with higher ranks are displayed as top results. This ranking procedure was used since the results stated in [4] have proved to be adequate. Although the FolkRank ranking algorithm was used as a

preliminary ranking strategy in LinksTo, in the near future other ranking algorithms will be studied and the one which provides the optimum results will be used. Apart from the searching engine, this module also contains a controller that controls the creation and amendments of *Links*.

Semantic Web Logic. This module is responsible for parsing and writing RDF statements, and for querying RDF models. The Jena framework² is embedded in this module since it provides the functionalities to: 1) read and write RDF statements; 2) a SPARQL engine that queries RDF models and 3) methods for in-memory and persistent storage. Moreover a REST web service is created to interact with Jena that acts as an HTTP SPARQL engine that allows SPARQL queries to be requested over HTTP and sends the results back to the client application.

4. Conclusion

The LinksTo prototype system provides a Web2.0 interface utilising Semantic Web technologies that provides users with the functionality to link Web resources. This system also provides Semantic Web developers with data formatted in RDF that can be reused. Since LinksTo is still in its infancy, once the *Links* dataset increases, this dataset can be used to analyse user behaviour with respect to how users link Web resources. This analysis can contribute to interesting research as to how to retrieve information based on linked resources. Moreover, LinksTo system adds value to the Web community by promoting the use of Semantic Web technologies adhering to the linked data practices in order for the Web of Data to continue to evolve.

References

1. Berners-Lee, T. (2006). Linked Data - Design Issues. [Online] <http://www.w3.org/DesignIssues/LinkedData.html>
2. Bizer C., Cyganiak R., Heath T. 2007. How to publish Linked Data on the Web. [Online] <http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/>
3. Baeza-Yates R., Tiberi. 2007. Extracting Semantic Relations from Query Logs. In KDD'07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, p. 76-85.
4. Hotho A., Jäschke R., Schmitz C., Stumme G. 2006. Information Retrieval in Folksonomies Search and Ranking. In: The Semantic Web Research and Applications, p. 411-426.
5. Hotho A., Jäschke R., Schmitz C., Stumme G. 2006. BibSonomy: A Social Bookmark and Publication Sharing System. In: Proceedings of the Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on Conceptual Structures, p. 87-102.
6. Krause B, Jäschke R., Hotho A., Stumme G. 2008. Logsonomy – Social Information Retrieval with Logdata. In HT'08: Proceedings of the Nineteenth ACM Conference on Hypertext and Hypermedia, p. 157-166.
7. Abel F., Frank M., Henze N., Krause D., Plappert D., Siehndel P. 2008. GroupMe! – Where Semantic Web Meets Web2.0. In: The Semantic Web p. 871-878.

² <http://jena.sourceforge.net/>