

Weaving Social E-Learning Platforms Into the Web of Linked Data

Selver Softic

(Institute of Information Systems & Information Management
JOANNEUM RESEARCH Forschungsges. mbH, Graz, Austria
selver.softic@joanneum.at)

Behnam Taraghi

(Graz University of Technology, Graz, Austria
b.taraghi@tugraz.at)

Wolfgang Halb

(Institute of Information Systems & Information Management
JOANNEUM RESEARCH Forschungsges. mbH, Graz, Austria
wolfgang.halb@joanneum.at)

Abstract: In this paper we present an approach for interlinking and RDFising social e-Learning Web 2.0 platforms like ELGG based on semantic tagging and Linked Data principles. A special module called SID (Semantically Interlinked Data) was developed to allow existing tagged and published user generated content an easy entrance into the Web of Data and to enrich it semantically on the other hand. Our approach uses commonly known vocabularies (FOAF, SIOC, MOAT and Tag Ontology) for modelling and generation tasks along with DBPedia as reference dataset for interlinking.

Key Words: semantic web, interlinking, semantic tagging, social platforms, e-learning

Category: H.3, H.4

1 Introduction

User generated content and social platforms play a major role in today's Web. The work presented in this paper aims at tapping this information source and shows approaches for improving the user experience by semantic interlinking and content enrichment in the case of a social e-Learning platform. Content generated by individuals can be enhanced with more objective semantics and thus becomes more valuable for all users.

Most content in the form of text, images, or video is also accompanied with associated meta data such as information about the creator, categorisations, or tags which can be used for search. The use of tags also enables a more selective information retrieval and appealing information visualisations like tag clouds. However, known issues such as ambiguity and the individual diversity of tag choice can lead to unsatisfying results. The approach described in this paper

proposes to enhance the descriptions and tags with semantics as a possible solution for this problem. By linking tags to datasets like DBpedia the content descriptions are given a well-defined meaning. If the meaning of a tag cannot be automatically resolved by the implemented interlinking module, the user is provided with a list of different meanings to choose from. This option implicitly also enables re-tagging of automatically falsely assigned meanings of tags done in the automatic interlinking step before. Content creators can contribute more reliable information by adding semantics to their content descriptions which in turn could positively influence the contributor's overall trustworthiness and reputation.

Social e-Learning platforms like ELGG¹ combine the advantages of social networks, Web 2.0 principles and educational techniques. They represent very interactive platforms with semi-controlled (tag generation is partly driven by the system) content generation but are limited to the area of their community. The exchange of information between two installations of the same platform (e.g. between two collaborating universities) cannot be done easily as there is no standardised exchange format used and it would require the implementation of specific Web services. The use of RDF and Linked Data in particular could solve this problem and enable the exchange and integration of information across platform boundaries. As a first step a module was implemented that focuses on tag-based semantically interlinked content.

2 Related Work

Recent research on tagging and folksonomies in social environments [Gruber 2007, Gruber 2008] emphasizes the meaning of tags for the Semantic Web and outlines the need for describing them by ontologies. These ontologies contribute to sophisticated knowledge representation, enable knowledge exchange among users and applications, and expose the human knowledge to the machines [Halb 2008, Kim 2008b]. All these three aspects fit into the vision and needs of social and e-Learning platforms, especially those used for higher education [Ebner 2008, Ebner 2008a] where the knowledge transfer along with knowledge and resource exchange is based on intensive interaction between the actors of learning process.

A review of current tag related ontologies can be found in [Kim 2008a]. The most relevant ontologies in this context are the Tag Ontology², MOAT (Meaning Of A Tag) [Passant 2008a, Passant 2008b], and SCOT (Social Semantic Cloud of Tags) [Kim 2008] which also offer features that underline social relatedness, relation to content and creators of content described by tags and the possibility of

¹ <http://www.elgg.org/>

² <http://www.holygoat.co.uk/projects/tags/>

involving external resources into the representing model. The results of comparisons and evaluations made between different tag ontologies based on recommendations and experiences of experts and on heuristics [Kim 2008b] are suggesting the usage of a tripartite model for the tagging process [Mika 2005, Halpin 2006]. This model is represented by a triple (U,T,R) containing information about users (U) that participate in tagging, set of available tags (T), and related resources (R) that should be tagged.

Our research work also focuses on the social ambience of tagged content and exposing this information to the Web of Data. This notion provides the escape from the “walled garden” of closed social platforms [Rowe 2008] and enables attaching social values to tags. For this purpose the widely deployed FOAF (Friend Of A Friend) [Brickley 2000] ontology seems to be the missing gap in a twofold way: Firstly it makes the social information exposed to humans and machines and secondly it allows to link the tags to its owners using *foaf:maker* property and MOAT description [Kim 2008a].

Motivated by prior works in that area (such as [Fernandez 2007, Bojars 2008, Hausenblas 2007]) our approach of RDFising and exposing the personal blogs, that every such platform includes, is based on the SIOC vocabulary [Breslin 2005] which is well suited for this purpose. Furthermore, in this way an easy integration of existing vocabularies into the proposed combination (MOAT and FOAF) can be realised building a pragmatic composition.

3 Modelling Approach

3.1 Platform and Data Characteristics

The ELGG Platform is one of many open source e-Learning environments used by numerous educational institutions and organizations around the world. Based on Web 2.0 technologies and its modular structure users can establish digital identities, connect with other users, collaborate with them and discover new resources through their connections. The common characteristic of Web 2.0 applications is the central role of the users. They are the actual content creators who share information with other users in form of blogs or some other type of content. Tagging plays an important role in this area. Similar data are inter-linked to each other within the system through tags. The search functionality of the platform bases on this principle, so that searching for items results in retrieval of resources that are tagged with those items as keywords. The fact that tags do not have a defined meaning and are considered merely as literal strings is a disadvantage for the entire application. Semantically interlinking tags with resources which describe tags such as DBPedia could lead to more appropriate search results.

3.2 Interlinking Model

Our proposed interlinking model for tag based semantic enhancement of user contributed content in social e-Learning platforms and its interlinking into the Linking Open Data cloud follows the recommendations for the usage of tripartite models [Mika 2005] containing information about users, tags, and related resources [Kim 2008b]. As simplicity is one of our model design goals the MOAT and Tag ontologies appear to be appropriate. The MOAT ontology offers to define meanings and to differ between different meanings of a tag. Additionally the context and scope of a tag can be explicitly described.

Inspired by the idea of linked data [Berners-Lee 2007], the meaning of tags is aligned with linked datasets in order to enrich them and the content they belong to. In a first phase interlinks with DBpedia have been generated. Based on the initial results it is possible to let the user refine the definitions as an applied way of User Contributed Interlinking (UCI) as proposed in [Hausenblas 2008].

The interlinking model for social e-Learning platforms makes use of the three wide spread ontologies MOAT, SIOC, and FOAF and simple relations between them (cf. example for ELGG Community in [Fig. 1]). For relating the user generated content with tags the Tag Ontology and restricted tagging method are used.

The proposed approach was inspired by recently introduced research work like SCOT Ontology [Kim 2008, Kim 2008a] and prior research efforts [Breslin 2005, Bojars 2008, Passant 2008b], but differs from them in its simplicity. The complexity of the interlinking model is intentionally kept at a low level to allow an easy implementation and thus potential higher acceptance.

The weaving process occurs basically in two steps: In the first step the content is RDFised. This includes social content represented through users and relations between them based on FOAF, user generated content represented through blog posts based on SIOC, and tags as binding member between them based on a combination of MOAT and Tag Ontology.

As depicted in [Fig. 1] a tag's meaning described with MOAT is related to the user's profile via the *foaf:maker* property which is also used for connecting the SIOC representation of a user's blog post to the FOAF profile.

The relation between blog posts and tags is realised using the *Restricted-Tagging* class and *tag:taggedResource* property of the Tag Ontology combined with *moat:tagMeaning* property from MOAT. The example in [Listing 1] demonstrates this constellation. Social e-Learning platforms like ELGG allow embedding of other social relevant resources like videos, slides, images, files, bookmarks or RSS Feeds as well. Those resources are also tagged and can be interlinked in an analogous manner.

In the second step a preliminary interlinking of RDFised tags to the meanings in DBpedia is done which is described in the following section.

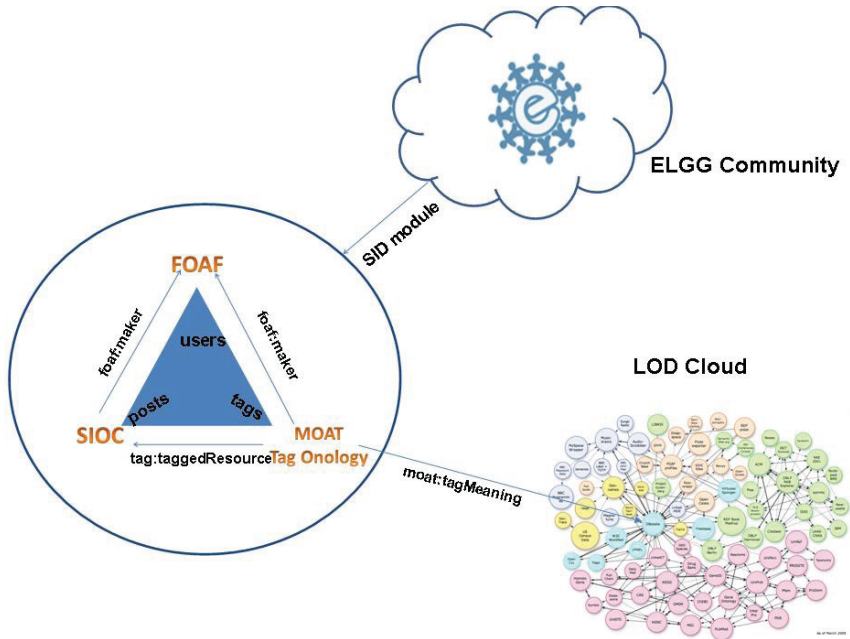


Figure 1: Model for tag-based interlinking of social content in ELGG.

4 Interlinking Module

The proposed interlinking model was applied in TUGLL (TU Graz Learn Land), an ELGG platform running at Graz University of Technology. It contains several thousands of users and tagged blog entries. ELGG is an open source PHP based social e-Learning network with a modular structure of functionality. Modules can be easily added as a plug-in to expand the functionality of the system. For the purpose of RDFising and interlinking we developed the SID (Semantically Interlinked Data) module containing classes for exporting FOAF profiles, SIOC posts and MOAT tag meanings. The module itself has an object oriented architecture which can also be used by other similar applications. The only modifications that need to be made for use in other applications concerns the queries for retrieving tags, user data and blog content from the corresponding platforms. The retrieved data must be prepared as input of SID classes. Once the input data is provided the classes can begin to parse the input data to create the triples in different RDF files using FOAF, SIOC and MOAT ontologies correspondingly. Finally the triples are stored in a RDF store (ARC2 RDF with MySQL³). A SPARQL endpoint is provided for potential further processing.

³ <http://arc.semsol.org/>

```

1 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
2 @prefix moat: <http://moat-project.org/ns> .
3 @prefix tag:
  <http://www.holygoat.co.uk/owl/redwood/0.1/tags/> .
4 @prefix rdf:
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
5
6 <http://tugll.tugraz.at/tag/post> a
  tag:RestrictedTagging;
7   foaf:maker <http://tugll.tugraz.at/96784/foaf/> ;
8   tag:tagName "post";
9   moat:tagMeaning <http://dbpedia.org/resource/Post>;
10  tag:associatedTag <http://tugll.tugraz.at/tag/post>;
11  tag:taggedResource
    <http://tugll.tugraz.at/96784/weblog/6965.html>.

```

Listing 1: Relating SIOC and FOAF with interlinked tags using RestrictedTag class.



Figure 2: Auto completed semantic tagging.

The trade-off made here is, that currently only one preliminary interlinked meaning per each found tag for all creators is assigned if the interlinking is done ex post where the user has no immediate possibility to interact and help refining the meaning. Although this is not the basic intention of MOAT vocabulary (that allows multiple meanings assignment for single tag) this approach provides a base for the semantic enrichment of social e-Learning platforms.

The semantic tagging mechanism provides an auto-complete function for disambiguation as depicted in [Fig. 2] that the user can optionally turn on. The feature makes use of the web service based DBpedia lookup API⁴.

⁴ <http://lookup.dbpedia.org/api/search.asmx>

5 Example Implementation

As preliminary result a simple user scenario represented through a recommender widget for a single resource (post, bookmark, slide, file, RSS ,video etc.) was implemented containing related resources tagged with same tag meaning or alternatively users having resources tagged with the same meaning. In [Listing 2] an exemplary query for the latter case used by widget to gather information is shown. All further relevant information can also be easily retrieved in this simple way.

```

1 SELECT DISTINCT ?pers WHERE {
2
3   ?reftag a tags:RestrictedTagging;
4           tags:taggedResource
5           <http://tugll.tugraz.at/medien07/weblog/778.html>;
6           moat:tagMeaning ?meaning_ref.
7
8   ?tag a   tags:RestrictedTagging;
9           foaf:maker ?pers;
10          tags:taggedResource ?res;
11          moat:tagMeaning ?meaning.
12
13  FILTER(?meaning = ?meaning_ref && ?res!=
14         <http://tugll.tugraz.at/medien07/weblog/778.html>)
15 }

```

Listing 2: Selecting related persons who tagged their resources with the same meaning.

As mentioned in section 4 the optional auto-complete function leads to a meaningful tagging in the whole system. Keywords are set more precisely in the sense of context than the case no auto-complete function is used. One of important advantages of RDFising the content is the interoperability. RDFising provides other learn management systems (LMS) with the whole public data (triples) as an open service. In a personal learning environment (PLE), for instance, it is possible now to access and consume the data generated or contributed through users (blogs, bookmarks, files, videos, etc.), build recommender widgets, and perform much more things by merely crawling the RDF triple store exposed using the SPARQL endpoint, even for users from external platforms, without any need to implement special web services for this purpose any more.

6 Conclusions

The presented approach for RDFising and interlinking a platform like ELGG shows several benefits: The problem of sharing information inside the community and to the outer world is diminished. The knowledge base of the community gets enhanced by reliable sources. Data present in RDF can be used for internal analysis or for integration of Semantic Web modules into the learning environment. Privacy issues are implicitly solved as users of the platform can generally choose the dissemination level (public or restricted to certain users or groups) of their content. Even though there are still open issues to be solved, with the proposed approach a large community finds its entrance to the Web of Data.

7 Future Work

Future work will focus on improvements of the data quality and user experience. This includes also evaluation of integration of semantics based tagging and re-tagging services like LODr [Passant 2008] and similar semantic applications which can contribute to knowledge expansion on the targeted social e-Learning platforms. Using the APIs for social relations resolution and presentation as existent e.g. in Facebook or based on Google Social Graphs will be considered. We also aim at stronger and more reliable interlinking with other linked datasets like e.g. RDF Book Mashup, Project Gutenberg, W3C WordNet, or GeoNames and between the social e-Learning platforms with focus on reliable educational content and social relevance. The benefits of open linked data enable e.g. that further related information existing on Flickr or in geographical database, can be retrieved for the given keyword and recommended to the user in form of additional information. It is also planned to make the SID module publicly available as plug-in for the ELGG platform and as independent RDFising and tag based interlinking tool for all resources containing tags.

References

- [Breslin 2005] Breslin, J., Harth, A., Bojars, U., Decker, S.: “Towards semantically-interlinked online communities”; Proc. 2nd European Semantic Web Conference (2005).
- [Berners-Lee 2007] Berners-Lee, T.: “Linked Data”; Technical report (2007).
- [Bojars 2008] Bojars, U., Passant, A., Cyganiak, R., Breslin, J.: “Weaving sioc into the web of linked data”; Proc. WWW 2008 Workshop: Linked Data on the Web (LDOW2008), Beijing, China, April 22, (2008).
- [Brickley 2000] Brickley, D. Miller, L.: “FOAF: Friend of a Friend”; Technical report (2000) <http://xmlns.com/foaf/0.1/>.
- [Ebner 2008] Ebner, M., Taraghi, B.: “A Blog Sphere for Higher Education”; Proc. 20th World Conference on Educational Multimedia, Hypermedia and Telecommunications (ED-Media) (2008).

- [Ebner 2008a] Ebner, M., Taraghi, B., Nagler, W.: "TUGLL Plug-ins: Special Needs for a University Wide Blogosphere"; International Conferences on Knowledge Management and New Media Technology (I-KNOW) (2008).
- [Fernandez 2007] Fernandez, S., Berrueta, D., Labra, J.: "Mailing Lists Meet The Semantic Web"; Proc. BIS 2007 Workshop on Social Aspects of the Web, Poznan, Poland (2007).
- [Gruber 2008] Gruber, T.: "Web Semantics: Science, Services and Agents on the World Wide Web"; February (2008), 4-13.
- [Gruber 2007] Gruber, T.: "Ontology of folksonomy: A mash-up of apples and oranges"; Intl Journal on Semantic Web & Information Systems, 3 (2007).
- [Halb 2008] Halb, W., Raimond, Y., Hausenblas, M.: "Building Linked Data For Both Humans and Machines"; Proc. WWW 2008 Workshop: Linked Data on the Web (LDOW2008), Beijing, China (2008).
- [Hausenblas 2008] Hausenblas, M., Halb, W.: "Interlinking of Resources with Semantics (Poster)"; Proc. 5th European Semantic Web Conference (2008).
- [Hausenblas 2007] Hausenblas, M., Rehatschek, H.: "mle: Enhancing the Exploration of Mailing List Archives Through Making Semantics Explicit"; Proc. Semantic Web Challenge 2007 at the 6th International Semantic Web Conference (ISWC07), Busan, South Korea (2007).
- [Kim 2008] Kim, H., Breslin, J.: "SCOT: Social Semantic Cloud of Tags" ; Technical report (2008) <http://scot-project.org/scot/ns>.
- [Kim 2008a] Kim, H.L., Scerri, S., Breslin, J., Decker, S., Kim, H.G.: "The state of the art in tag ontologies: A semantic model for tagging and folksonomies."; Proc. International Conference on Dublin Core and Metadata Applications, Berlin, Germany (2008).
- [Kim 2008b] Kim, H.L., Passant, A., Breslin, J., Scerri, S., Decker, S.: "Review and alignment of tag ontologies for semantically-linked data in collaborative tagging spaces"; Proc. ICSC 08: 2008 IEEE International Conference on Semantic Computing, Washington, DC, USA, IEEE Computer Society (2008), 315-322.
- [Mika 2005] Mika, P.: "Ontologies Are Us: A Unified Model of Social Networks and Semantics"; Lecture Notes in Computer Science (2005), 522-536.
- [Halpin 2006] Halpin, H., Robu, V., Shepard H.: "The Dynamics and Semantics of Collaborative Tagging"; Proceedings of the 1st Semantic Authoring and Annotation Workshop (SAAW'06), Vol-209.
- [Passant 2008] Passant, A.: "Lodr - a linking open data tagging system"; Proc. ISWC2008 Workshop on Social Data on the Web (SDoW2008) Karlsruhe, Germany, October 27, 2008.
- [Passant 2008a] Passant, A.: "MOAT: Meaning Of A Tag Ontology"; Technical report, DERI, Galway (2008) <http://moat-project.org/ns>.
- [Passant 2008b] Passant, A., Laublet, P.: "Meaning of a tag: A collaborative approach to bridge the gap between tagging and linked data"; Proc. WWW 2008 Workshop: Linked Data on the Web (LDOW2008), Beijing, China, April 22, (2008).
- [Rowe 2008] Rowe, M., Ciravegna, F.: "Getting to me - exporting semantic social network from facebook"; Proc. ISWC 2008 Workshop: Social Data On Web (SDOW2008), Karlsruhe, Germany, (2008).