The Ontology and Architecture for an Academic Social Network

Moharram Challenger

**Computer Engineering Department, Islamic Azad University – Shabestar Branch,
Shabestar, East Azerbaijan, Iran**

Abstract

In this paper, UniGrad, an ontology that extends FOAF (Friend Of A Friend) with graduate education and research concepts has been presented. Nowadays, having lots of information exist throughout the Web, e.g. for academicians and researchers, finding a specialist in an area as a query is somehow cumbersome. The proposed technique is using Semantic Web technology including ontologies such as FOAF and SIOC to provide possibility of defining a structure for information to take the advantage of desired queries. We also propose the design of a social network based on UniGrad and using of this ontology to integrate different data sources.

Keywords: *FOAF, SIOC, Graduate Education, Ontology, Social Network.*

1. Introduction

Today's web is designed to be useful for human; it does not aim to be understood by machines. Semantic Web is an extension of the current Web, in which data is represented in a standard format with metadata that allows integration and processing of different data sources automatically by computer programs [1].

Suppose that a user is going to apply for a PhD program in a university and has some basic criteria for his or her choices. The existing solution for this applicant is searching universities to find his or her research area and the academicians in this university should have worked and have publications in this area. Using current web, the applicant has to search manually and compare all universities' PhD programs, courses and research areas of the academicians individually to find a program with the required criteria. In other words, the person should find the relations between data. This is an inefficient and error-prone solution. The data on the Web is distributed and is not linked effectively. To make this process more efficient, the data must be linked. As a matter of fact, the aim of the Semantic Web is to enable people to share structured data on the Web as easily as they can share documents today. However, using metadata, machines can read the contents by looking at pictures and reading descriptions like a person.

The Semantic Web is not designed just as a new data model - it is specifically appropriate for the linking of data between many different models. It also helps adding information relating to different databases on the Web to allow sophisticated operations to be performed across them [2].

One of the new and useful semantic web standards is FOAF [3] which is a specific schema that is built on top of the RDF. It provides a mechanism for people to link with their own FOAF files and helps them to define their social networks. Another key point is that data on the Web must be structured. Since FOAF has become a widely accepting standard vocabulary for representing social networks, many social networking web sites use it; enable to produce FOAF profiles for their users [4]. In addition, like Google Social Graph API [5] many services allow developers to expose social relationships embedded in the Web sites. Getting single and standardized access mechanism to distributed data is also semantic web's ability.

The paper is organized as follows. Next section describes background information about FOAF, SIOC and their integration. Following section describes our methodology to create the UniGrad ontology. After that the architecture of the system designed for the UniGrad Social Network has been presented. Last section concludes and informs about related and future work.

2. Related Work

In this section related work is elaborated including FOAF, SIOC and their integration.

2.1 FOAF

FOAF is a machine-readable ontology describing people, their relations and activities with other people. Any person can have FOAF document and describe himself. People introduce their friends and relations to other people, and they build a social network without any centralized database (FOAF documents are distributed among people's web servers).

FOAF is a descriptive vocabulary based on RDF and OWL; therefore computers can use these profiles (for example they can traverse them to do a specific search)

due to their machine readability. This is possible will lead to "Social Semantic Information Spaces" where

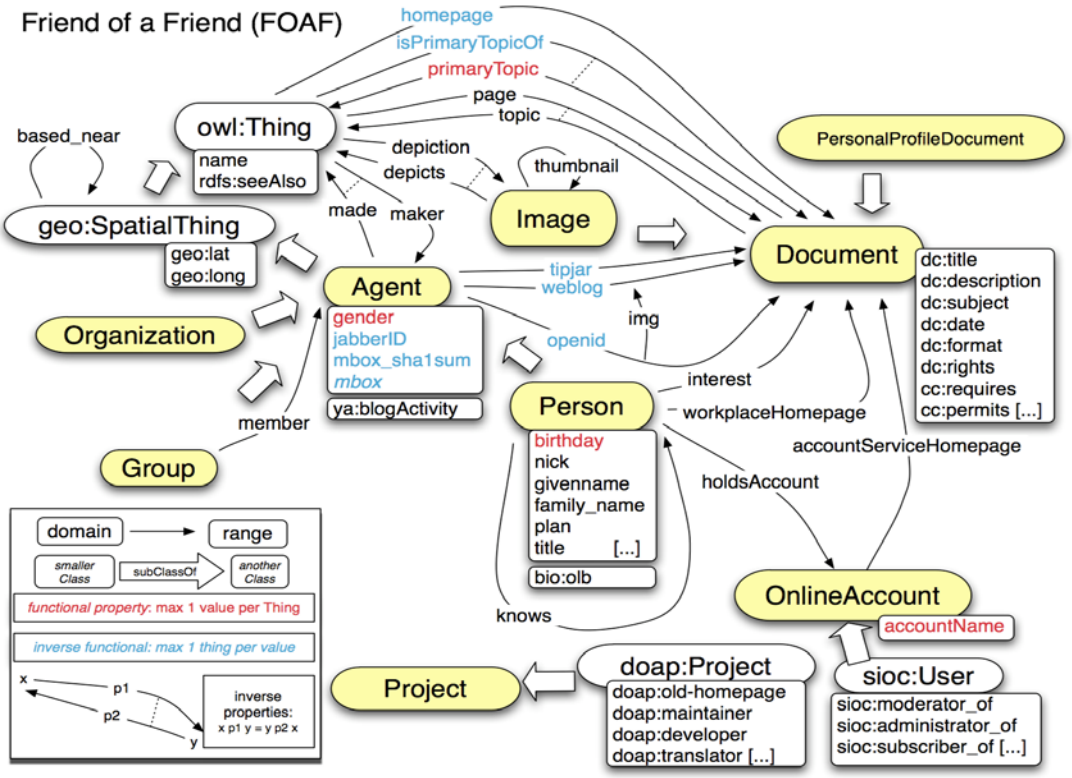

Fig. 1 FOAF Terms by Class and Property [6]

according to relations defined between people. Each person has unique id in FOAF, e.g. email, openID, URI, and so on [7]. In Fig. 1, FOAF terms are given in by classes and the properties. FOAF helps to use a simple set of modeling constructs for creating extensible, distributed information networks of people. Due to the distributed nature of RDF, a network of information can be distributed throughout the web. It also uses inference structure of RFDS to add completeness to their information structure. FOAF takes something of an evolutionary approach to information extension. FOAF takes very seriously the AAA slogan, to the point that the actual preferred parts of the representation will be determined to a large extent by its use. The free extension style of FOAF is performed using the graph overlay mechanism of RDF. FOAF effort is a standard effort which is supported by a committee. Meanwhile it is not a complete work yet.

2.2 SIOC

Semantic Web technologies provide standards and models to build a Web of Data, with unified models to represent interlinked data from different sources. Hence, combining Semantic Web technologies and social media paradigms information is socially created and maintained as well as being interlinked and machine-understandable, leading to new ways to discover information on the Web [8].

SIOC (Semantically-Interlinked Online Communities) [9] is Semantic Web technology which provides methods for interconnecting discussion methods such as blogs, forums and mailing list. It consists of the SIOC ontology; an open-standard machine readable format for expressing the information contained both explicitly in internet discussion methods. The SIOC vocabulary [10] is based on RDF and is defined using RDFS.

One of the aims when developing the SIOC ontology was to keep it as simple as possible, yet powerful, so that it could be easily deployed in existing applications [11]. Fig. 2 describes the main classes of the SIOC core ontology. While the SIOC core ontology provides a sufficient model to describe the activities of online communities, it is not fine-grained enough to allow a person to distinguish between, e.g., a blog post and an update on a microblogging service, as it only introduces a single sioc:Post class. In order to define this additional level of semantics, a SIOC Types module was designed, modeling more than 20 different types of content, such as sioct:MicroblogPost and sioct:WikiArticle.

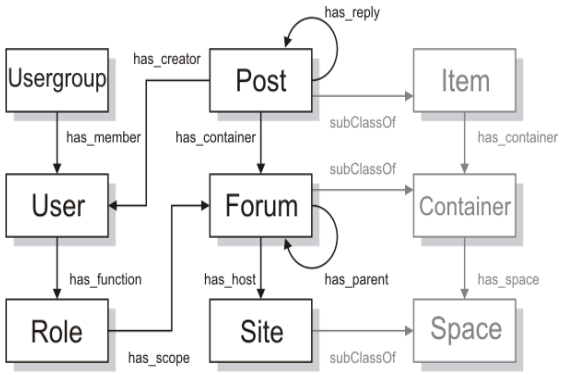

Fig. 2 The SIOC core ontology model [10]

2.3. FOAF and SIOC Integration

SIOC reuses the Dublin Core model to define various attributes of created content (such as the creation date of an item, using dcterms:created) and FOAF - Friend Of A Friend - to model personal identity and related attributes. Interlinking FOAF and SIOC also provides a model for identity federation on the Web. While one can have different user profiles on different websites (represented as instances of sioc:User), they all relate to the same physical person (foaf:Person). Using relations, all the SMCs (social media contributions) of a given physical person can be modeled in a single global graph of RDF data, enabling a model for interoperability and portability of social data between services.

3. UniGrad Ontology

We've extended FOAF in order to be able to represent new concepts defined in our ontology. Our extensions can be divided into two parts; new concepts defined in UniGrad namespace and an imported ontology, DBLP [12].

We defined new concepts within the "unigrad" namespace and connected them to the FOAF concepts. An example extension is given in Fig. 3. FOAF Person class is extended by two subclasses: Academician and Student; two main concepts of our graduate education and research ontology. They are connected through advisorOf/advisedBy relations, which are sub properties of the foaf:knows property. They are sub properties of foaf:knows property because if an academician is an advisor of a student, this means that they know each other. Another similar extension is implemented for the University concept. unigrad:university is defined as a sub class of foaf:organization.

To create a graduate education and research ontology, we need to define publications, with their types and connections to other concepts. FOAF has a concept called document, and it is used as a range of hasPublications relation. It doesn't have the properties we need, such as publication types or a coauthor relation. For this reason, we decided to use DBLP ontology [12]. It is created by Lehigh University by adapting UGA's SwetoDblp ontology [13]. It basically creates OWL descriptions of the corresponding DBLP concepts. DBLP web site [14] is a popular tool to trace the work of researchers in computer science. DBLP originally has an XML based API, which is not a perfect representation as discussed in [15]. Its data is accessible via XML files and its application is as a data source as described in the following section. In Fig. 4, DBLP's class hierarchy and part of its definitions is depicted.

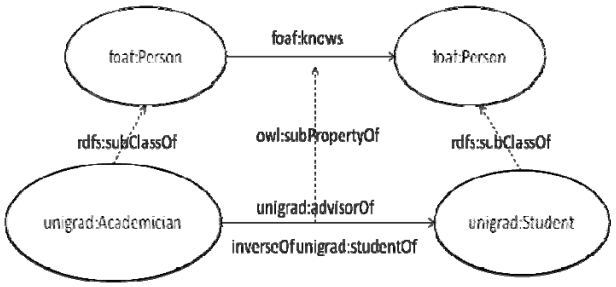

Fig. 3 Extending FOAF Person class with UniGrad classes Academician and Student

Our motivation while creating the UniGrad ontology is to create an ontology that can model a person's education, work experience, research interests, and publications. The purpose was to model it in a flexible way. For instance, since a person job changes in time, the new information should be added to the ontology while the information about the previous experience still remains. This is modeled with an experience concept, which can be academic or work experience. It includes start and end dates, organizations and research/work areas. Academic experience can have one of the academic titles; assistant professor, associate professor and professor. Another similar concept is created for education information. It is named degree and can be master or PhD. Similarly it has start/end dates and university as an organization. We also created a research area concept and instantiated it in the schema of the ontology as subclasses. So different research areas are part of our ontology schema. In Fig. 5, an overview of the resulting UniGrad Ontology is given. It has concepts from FOAF, SIOC, UniGrad and DBLP.

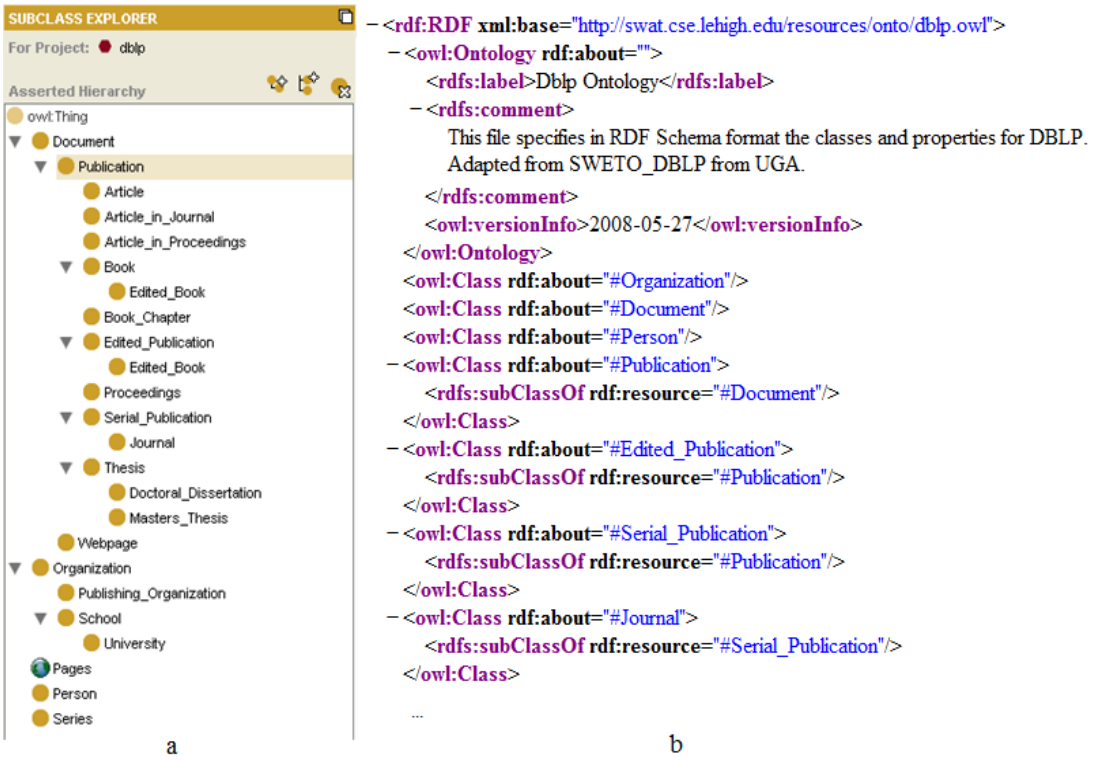

Fig. 4 a) Class Hierarchy of DBLP Ontology (using Protégé [8]) b) Part of DBLP Ontology definitions in OWL

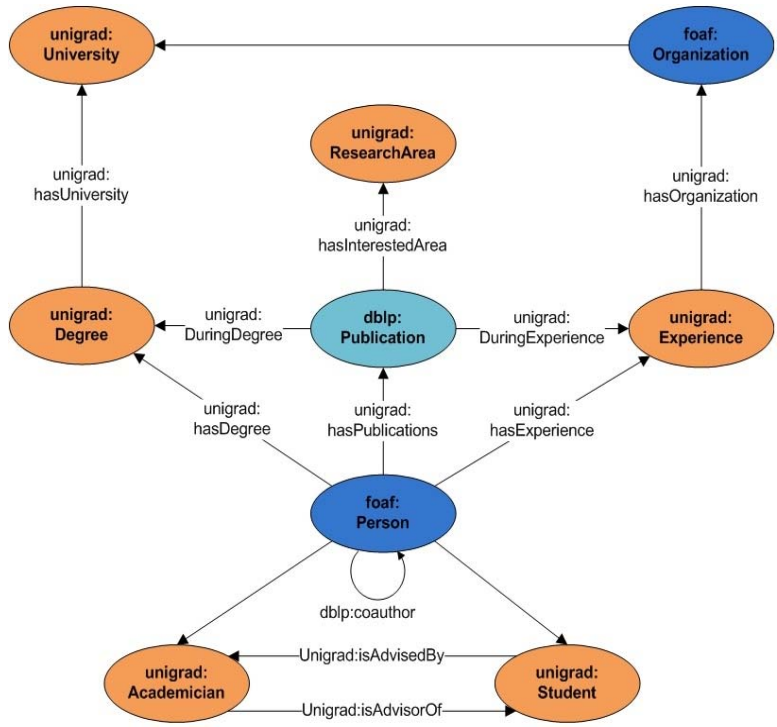

Fig. 5 Overview of UniGrad Ontology

4. Application: An Academic Social Network

In this section we describe the use of DBLP data in details. It should be noted that this is a typical scenario to use UniGrad ontology and it can be extended with many other data sources.

Fig. 6 shows an example of architecture defined on top of UniGrad ontology. The architecture has three different data sources plus user defined data. These data sources are (1) other social networks, via SIOC, (2) user profiles, via FOAF and (3) publication data, via DBLP.

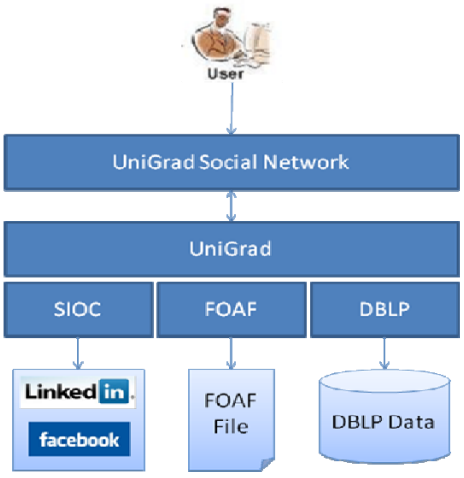

Fig. 6 The Application Architecture

According to FOAF files, we have a friend list for a person, so that it is clear whom they know. Nevertheless, we need to extract other relations, either from the information that the user gives or using Internet and other data sources. An example relation which is needed to be extracted explicitly is "coauthor" relation. The designed system uses DBLP data as a data source and thus it can have access to publications of a user. It is assumed that a user registers to the system with a FOAF file. Thus we have some information about the user, including the name and surname. We can query the DBLP data and retrieve publication information of the user with user name. More interesting part starts after getting this information. If we know a person's publications and the publication information containing a list of authors, we can infer the coauthors. This means some implicit relations of the person can be extracted. Other methods including web mining, to extract implicit relations is discussed in the future work section.

We have instantiated the ontology manually with the information of the popular semantic web researchers, to be able to execute an example query. Then we queried it to search for researchers with similar research interests. The SPARQL query and the results are given in Fig. 7. First, it retrieves research areas of the selected researcher, using unigrad:hasResearchInterest property. This property is defined as a sub property of foaf:topic_interest, hence builds a connection with FOAF interests. Then the query searches for the researchers with the research interests of first.

5. Conclusion

In this paper, UniGrad, an ontology that extends FOAF with graduate education and research concepts have been presented. The design of a social network based on UniGrad that shows how this ontology could be used to integrate different data sources has been described.

To the best of our knowledge, this is the first work intended to extend the FOAF with research ontology and design a social network using this ontology. However, there are similar other efforts extending FOAF in different fields, such as music [16][17] and film recommendations [18]. Other similar efforts include Polyphonet [19], a social network for researchers but it doesn't use FOAF or any other semantic technology as a basis.

Our system uses DBLP as an external data source and extracts relationships from this data. There are many other data on the web to integrate with our system and will increase with the increasing use of semantic web. Similar to the work in Polyphonet [19], web mining techniques can also be used to extract more data and relations about the users. Implementations of these techniques are possible future work of this work.

Fig. 7: A SPARQL Query and Results

Acknowledgments

The author would like to thank Rahele Eslampanah for her valuable comments and proofread.

References

[1] W3C Semantic Web Activity: http://www.w3.org/2001/sw/ (Last access: March 2012)
[2] What the Semantic Web can represent: http://www.w3.org/DesignIssues/RDFnot.html (Last access: March 2012)
[3] The Friend of a Friend (FOAF) project: http://www.foaf-project.org/ (Last access: March 2012)
[4] J. Goldbeck, M. Rothstein, "Linking social Networks on the web with FOAF : a semantic web case study", Proceedings of the 23rd national conference on Artificial intelligence (AAAI'08), 2008, Vol. 2.
[5] Google Social Graph API: http://code.google.com/intl/tr-TR/apis/socialgraph/ (Last access: March 2012)
[6] FOAF Specification: http://xmlns.com/foaf/spec/images/foafspec.jpg (Last access: March 2012)
[7] D. Allemang, J. Hendler, Semantic Web for the Working Ontologist, Effective Modeling in RDFS and OWL, Morgan Kauphman publication, 2008.
[8] J. G. Breslin, U. Bojārs, A. Passant, S. Fernández, and S. Decker, "SIOC: Content Exchange and Semantic Interoperability between Social Networks", W3C Workshop on the Future of Social Networking, Barcelona, January 2009, pp. 15-16.
[9] The SIOC (Semantically-Interlinked Online Communities) project: http://sioc-project.org/ (Last access: March 2012)
[10] U. Bojars, J.G. Breslin, D. Berrueta, D. Brickley, S. Decker, S. Fernández, C. Görn, A. Harth, T. Heath, K. Idehen, K. Kjernsmo, A. Miles, A. Passant, A. Polleres, L. Polo, M. Sintek, "SIOC Core Ontology Specification", W3C Member Submission, 12 June 2007.
[11] U. Bojārs, J. G. Breslin, V. Peristeras, G. Tummarello, S. Decker, "Interlinking the Social Web with Semantics", IEEE Intelligent Systems, Vol. 23, Issue 3, May/June 2008, pp. 29-40.
[12] The DBLP Ontology. http://swat.cse.lehigh.edu/resources/onto/dblp.owl (Last access: March 2012)
[13] B. Aleman-Meza, F. Hakimpour, I.B. Arpinar, A.P. Sheth, "SwetoDblp Ontology of Computer Science Publications", Journal of Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 5, Issue 3, 2007, pp. 151-155.
[14] The DBLP Computer Science Bibliography: http://www.informatik.uni-trier.de/~ley/db/ (Last access: March 2012)
[15] Protégé Ontology Editor: http://protege.stanford.edu/ (Last access: March 2012)
[16] O. Celma and X. Serra. "FOAFing the music: Bridging the semantic gap in music recommendation", Journal of Web Semantics, Vol. 6, No 4, 2008, pp. 250–256.
[17] A. Passant, , and Y. Raimond, "Combining Social Music and Semantic Web for Music-related Recommender Systems", In Proceedings of the First Social Data on the Web Workshop, Karlsruhe, Germany, October 27, 2008.
[18] J. Golbeck, and J. Hendler, "Filmtrust: Movie recommendations using trust in Web-based social networks", In Proceedings of the IEEE Consumer Communications and Networking Conference, 2006.
[19] Y. Matsuo, J. Mori, M. Hamasaki, K. Ishida, T. Nishimura, H. Takeda, K. Hasida, and M. Ishizuka, "POLYPHONET: An Advanced Social Network Extraction System", Journal of Web Semantics, Vol. 5, No. 4, 2007.

Moharram Challenger is a PhD candidate of Information Technology in the International Computer Institute, EGE University, Izmir, Turkey. He is also member of Computer Engineering faculty at IAU-Shabestar University in East-Azerbaijan, Iran. Moharram is also graduate student member of IEEE. He received his bachelor and master in computer engineering and software engineering respectively, from IAU University, Iran. His interesting area includes: Multi-agent systems, Domain specific (modeling) languages, distributed algorithms and Semantic web. He is also interested in using graph theory and its algorithms in distributed algorithms. In addition, Moharram has several conference and journal papers.