The SIOC Project: Semantically-Interlinked Online Communities

Uldis Bojārs and Alexandre Passant and John G. Breslin and Stefan Decker
Digital Enterprise Research Institute,
National University of Ireland, Galway, Ireland.
Email: [uldis.bojars,alexandre.passant,john.breslin,stefan.decker]@deri.org

***Abstract*—The SIOC project is aimed at expressing information about the structure and content of online community sites and at enabling interoperability on the Social Web using Semantic Web technologies. In this paper we briefly describe the SIOC project, introduce the SIOC Core ontology and its modules, and discuss some SIOC-based applications in terms of human and agent communication.**

I. Introduction

While new paradigms, tools and web services introduced by the Social Web (such as blogs, wikis, tagging practices) are now widely accepted in both public and scientific communities, these tools generally act as independent data silos; hence, interoperability between applications is a complex issue. The SIOC project[1] aims at solving this by providing a comprehensive data model (as well as related tools and applications) based on Semantic Web technologies [1] in order to represent online communities and their activities in an homogenous way.

The SIOC project consists of two main parts:

- The SIOC Ontology, composed of a Core ontology and different modules;
- A set of applications, for both producing and consuming SIOC data, constantly evolving based on the user needs and implementations.

II. The SIOC Ontology

The SIOC Ontology is composed of a Core ontology and a set of modules, focusing on the ease of integration of SIOC in existing applications by Web developers. With an emphasis on standard Semantic Web technologies since its beginning, the whole ontology has been designed using RDF(S)/OWL. A comprehensive overview of the SIOC ontology is provided in [2].

A. SIOC Core Ontology

Main classes and properties in the SIOC ontology[2] are shown in Figure 1. While relatively small and simple, this model is yet powerful enough to represent the content produced and exchanged within online communities. For instance, a `Forum` represents a space in which discussion happen (not necessarily a bulletin board, in spite of its name, but for example as a weblog), and contains different `Posts`, written by `Users`. The following example then describes that Alice has created a post in a particular forum (*i.e.* an area of discussion) and that Bob replied to it — as follows (using the N3 notation - prefixes omitted). In order to represent more abstract containers (such as a personal information space), more general `Container` and `Space` classes can be used.

```
:post a sioc:Post ;
  sioc:has_creator :alice ;
  sioc:has_container :forum ;
  sioc:has_reply :reply .
:forum a sioc:Forum .
:reply a sioc:Post ;
  sioc:has_creator :bob .
:alice a sioc:User .
:bob a sioc:User .
```

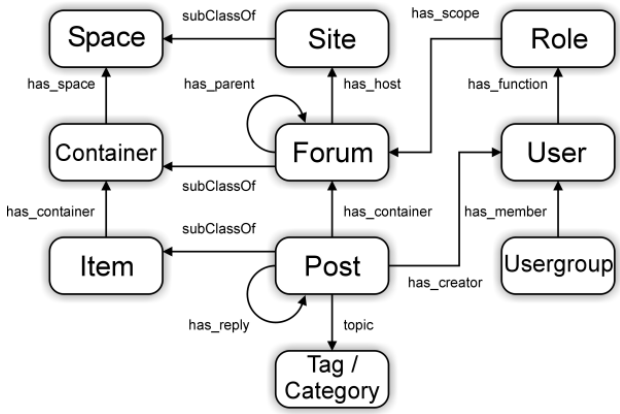

Fig. 1. Main classes and properties in the SIOC ontology.

Other properties and classes exist in the ontology, for instance properties to represent previous and next versions of an item, which can be used when representing Wikis.

It is worth noticing that each `sioc:User` is actually related to the `Agent` class (from the FOAF[3] — Friend Of A Friend [3] — ontology), and that `sioc:User` can be consequently associated with both software agents and human users. Hence, SIOC can be used to represent communication between bots on the Social Web (*e.g.* on IRC using the

[1] http://sioc-project.org
[2] http://rdfs.org/sioc/spec
[3] http://foaf-project.org

SIOC-IRClog project[4]) and then enable a machine-readable description of communities of agents interacting together on social websites. Moreover, another link between FOAF and SIOC is that the social networking aspect can be represented using FOAF (by `foaf:knows`), while SIOC can be used to represent aspects of relationships inside the community, *e.g.* a user following another (on microblogging services) can be expressed by the `sioc:follows` property.

B. SIOC Modules

Several SIOC modules have been defined to extend the available terms and to avoid making the SIOC Core Ontology too complex and unreadable. SIOC has 4 ontology modules: Access, Argumentation, Services and Types[5].

- the Access module defines simple classes and properties regarding the notions of Role and Permission to represent access rights and permissions in online communities websites;
- the Argument module defines classes and properties to represent simple argumentative discussions in online communities websites.
- the Services module defines classes and properties to reprsent Web services related to online communities (*e.g.* API endpoint and return format, etc.);
- the Types module defines advanced content-types to be used when defining user-generated content from online-communities. For instance, it includes classes such as `sioct:BlogPost`or `sioct:Wiki` that respectively subclass the `sioc:Post`and `sioc:Forum`classes from the Core Ontology;

In addition, one of the recent development of SIOC is a module defining alignments between SIOC and the SWAN — Semantic Web Applications in Neuromedicine — ontology [4][6], providing a complete model for fine-grained argumentative discussions in online scientific communities.

III. Status and Update of SIOC

Since the goal of SIOC is to provide interoperability between communities on the Social Web, one way to evaluate its success is to consider its uptake on the Web. To illustrate the amount of SIOC data on the Web, according to the PingTheSemanticWeb (PTSW) service[7] on June 2009 there were 132'475 URIs which contain data described using the SIOC ontology[8].

In addition to the uptake in terms of number of documents, SIOC is now widely accepted as a core ontology to describe Social Web communities using Semantic Web technologies, alongside with FOAF. Hence, the use of SIOC is suggested by the Yahoo! SearchMonkey developer documentation[9] (SIOC data being indexed by SearchMonkey to improve presentation of search results) and by various best practices documents describing data publishing on the Semantic Web such as [5].

Various SIOC-enabled applications have been created[10], forming a *food-chain* (Figure 2) that includes data exporters for Web 2.0 platforms and services (*e.g.* Drupal[11], WordPress, Flickr, etc.), tools for indexing of distributed data, and browser and visualization applications, allowing users to consume SIOC data from various sources in an integrated way.

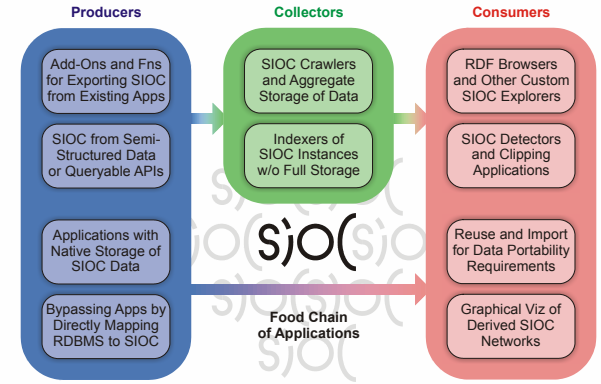

Fig. 2. The SIOC food-chain

IV. Conclusion

In this paper, we briefly introduced the SIOC project, its goals and means as well as uptake and services overview. In particular, in the context of multi-agent systems and online communities, an already active fields for several years if we consider for instance [6], the SIOC ontology could be used to express in a unified way information about activities of both humans and agents interacting in these communities.

Acknowledgements

The work presented in this paper has been funded in part by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Líon-2).

References

[1] T. Berners-Lee, J. A. Hendler, and O. Lassila, "The Semantic Web," *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001.

[2] U. Bojārs, J. G. Breslin, V. Peristeras, G. Tummarello, and S. Decker, "Interlinking the Social Web with Semantics," *IEEE Intelligent Systems*, vol. 23, no. 3, pp. 29–40, May/June 2008.

[3] D. Brickley and L. Miller, "FOAF Vocabulary Specification," FOAF Project, Namespace Document 2 Sept 2004, 2004, http://xmlns.com/foaf/0.1/. [Online]. Available: http://xmlns.com/foaf/0.1/

[4] P. Ciccarese, E. Wu, G. Wong, M. Ocana, J. Kinoshita, A. Ruttenberg, and T. Clark, "The SWAN biomedical discourse ontology," *Journal of Biomedical Informatics*, vol. 41, no. 5, pp. 739–751, 2008.

[5] C. Bizer, R. Cyganiak, and T. Heath, "How to Publish Linked Data on the Web," Tech. Rep., 2007, http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/.

[6] F. Gandon and R. Dieng-Kuntz, "Distributed Artificial Intelligence for Distributed Corporate Knowledge Management," in *CIA 2002 - 6th Cooperative Information Agents - Intelligent Agents for the Internet and Web September*. Springer LNAI 2446, pp. 202–217.

[4] http://irc.sioc-project.org/about.html

[5] Information about SIOC modules: http://rdfs.org/sioc/spec/#sec-modules

[6] http://rdfs.org/sioc/swan

[7] http://pingthesemanticweb.com

[8] The full amount of SIOC information on the Web is larger than described here as PTSW indexes only a part of available RDF data.

[9] http://developer.yahoo.com/searchmonkey/smguide/profile_vocab.html

[10] http://rdfs.org/sioc/applications/

[11] http://drupal.org/project/sioc