A SIOC Enabled Explorer of Shared Workspaces

Ke Ning, Vassilios Peristeras, Uldis Bojars, John Breslin
DERI, NUIG

Abstract: Shared Workspaces are getting more and more popular in helping people collaborate. However, it's difficult for user to search information across different Shared Workspaces since they are currently islands that are not interlinked. Based on semantic web technologies, SIOC provide ontology and tools to interlink online communities. By using SIOC to develop an explorer for Shared Workspaces, we make distributed and heterogeneous shared workspaces interlinked and allow user to perform advance search functions across them.

Introduction

A Shared Workspace or Collaborative Workspace is an inter-connected environment in which all the participants in dispersed locations can access and interact with each other just as inside a single entity [1]. Typically, people put objects like documents, notes, calendars, contact information etc. onto a Shared Workspace to allow other related participants to access regardless space and time differentials. Also, the sharable objects on Shared Workspaces are organized in a structural way to help people navigate and search efficiently. Nowadays, there has been lots of Shared Workspace supporting tools developed, among the popular ones include BSCW [2], Microsoft Sharepoint [3], etc.

At present, different Shared Workspaces are information islands, i.e., they can not talk to each other. However, as collaborations are more and more across project and organization boundaries, it's quite common that people need to combine information from different Shared Workspaces in their daily work. With isolated Shared workspaces, it's up to the user to build up and maintain the connections between different Shared Workspaces [4]. Especially, it's not easy to search information across shared workspaces, since people have to go into each related Shared workspace to perform the search function and integrate the results manually.

In this paper, we are trying to use SIOC [5] to solve this problem. SIOC has been developed for semantically interlinked online communities. It provides methods for interconnecting discussion methods such as blogs, forums and mailing lists to each other. It consists of the SIOC ontology, an open-standard machine readable format for expressing the information contained both explicitly and implicitly in internet discussion methods, of SIOC metadata producers for a number of popular blogging platforms and content management systems, and of storage and browsing / searching systems for leveraging this SIOC data. Since Shared Workspace can be seen as a kind of online communities, it's quite natural to think of reusing SIOC to interlink Shared Workspaces.

In the rest of this paper, we will describe a motivation scenario firstly, and then give a brief introduction of SIOC ontology and its supporting tools, and then present a SIOC based Explorer for Shared Workspaces that can exactly solve Tom's problem. Finally we will give conclusions and present the future work.

Motivation Scenario

Tom is working on two EU research project: Ecos and Incon. In each project, there is an internal shared workspace to help partners collaborate efficiently: Ecos' shared workspace is based on BSCW, while Incon's shared workspace is based on Microsoft Sharepoint. These two projects are all aiming at bringing new technologies for CWE (Collaborative Working Environment), so there are some similar interesting topics between them, also some of the their participants are from the same organization (e.g. Tom).

As the projects progress, more and more documents, calendar events, discussions, etc, have been created on both shared workspaces. Now it becomes not easy for Tom to search related information on the shared workspaces. For example, if he wants to find deliverables related to the topic of "web 2.0" on both projects, or he wants to find out calendar events that are related to him, he has to log on to both shared workspaces and perform a search action, and then combine the search results manually. It's not that bad if there are only two shared workspaces as in this case, but what if there are 5 or even more? Obviously it's not a trivial task.

SIOC Ontology and Tools

Semantically-Interlinked Online Communities or SIOC is a framework aimed at connecting online community sites and internet-based discussions [6]. Currently, online communities (boards, blogs, etc.) are like islands - they contain valuable information but are not well connected. SIOC allows us to interlink these sites, and enables the extraction of richer information from various discussion services.

The vision of SIOC is: To create an ontology that fully describes the content and structure of most online community sites; To create new connections between discussion channels and posts, and to allow users to browse discussion data in interesting ways using these connections; To overcome a "chicken-and-egg" problem with the Semantic Web (no applications without data, and no data without applications) by making it easy to generate and use SIOC data.

The core of SIOC is the ontology [7]. It is a vocabulary that contains concepts necessary to express information contained in online community sites. Figure 1 shows the main concepts in the SIOC ontology. To better fit specific description requirements of different sub domain of online community, SIOC modules are used. At present SIOC has two modules: Types and Services. In next section we will se that most of the Shared Workspace concepts can be mapped directly to SIOC concepts.
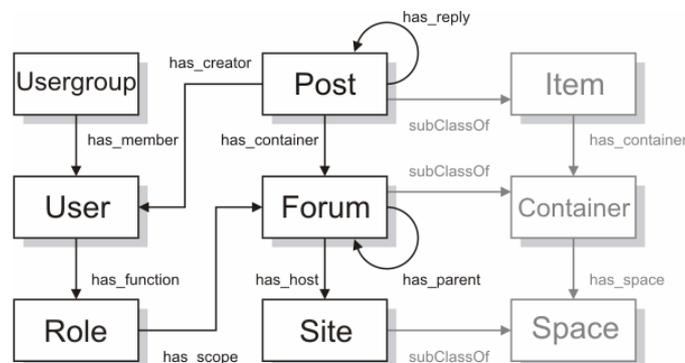

Figure 1: The main concepts in the SIOC ontology

Figure 2 describes a typical SIOC application framework. Firstly, Web-Based Communities and Legacy systems should be able to export SIOC data through SIOC Interfaces or Legacy Data Wrappers. Various tools, exporters and services have been created to expose SIOC data from existing online communities. Then there is a RDF Store to collect and store SIOC data from different sources. Basically, any RDF Store could be used here, e.g. Redland, Jena, and so on. Finally, a Query/User Interface is provided for user to navigate or query across the SIOC data. Interesting tools have also been developed here, e.g. SIOC Browser and SIOC Explorer based on ActiveRDF. For more information about SIOC Application, please refer to [8].
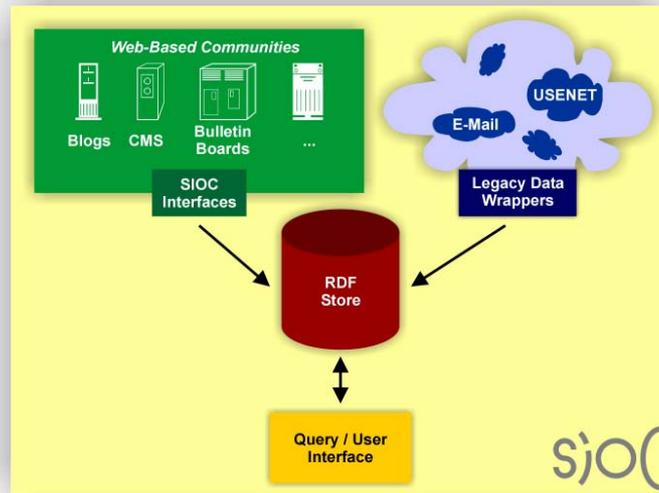

Figure 2: A Typical SIOC Application Framework

Shared Workspaces Explorer

To solve Tom's problem in the Motivation Scenario, we are trying to develop a Shared Workspaces Explorer by reusing and extending the current SIOC Explorer [9]. Figure 3 shows the framework of Shared Workspace Explorer.
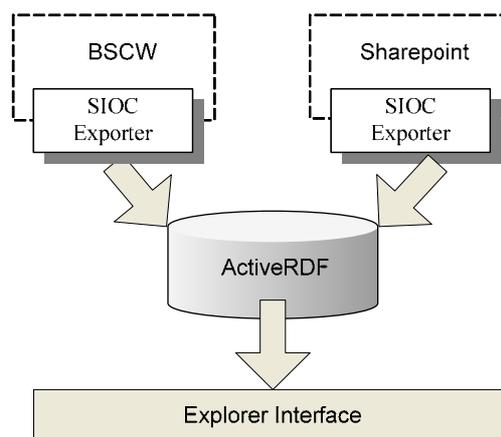

Figure 3: Framework of Shared Workspace Explorer

Before we can develp a SIOC Exporter for BSCW or Sharepoint, we need to map their sharable objects concepts to SIOC. Table 1 shows the Concept Mapping from BSCW to SIOC. Basically, SIOC can support all concepts in BSCW. However, to

more precisely describe Shared Workspace concepts, we also plan to extend SIOC Type Module with some new concepts (as showed in the SIOC(extended) column) in the future.

Table 1: Concept Mapping from BSCW to SIOC

BSCW	SIOC	SIOC(extended)
BSCW Site	sioc:Site	
Workspace	sioc:Forum	sioctypes:Workspace->Container
Document	foaf:Document	
Note	sioc:Post	
URL	annotea:Bookmark	
Folder	sioc:Forum	sioctypes:Folder->Container
Search folder	sioc:Forum	sioctypes:Folder->Container
Calendar	sioctypes:EventCalendar	
Discussion	sioctypes:ArgumentativeDiscussion	
Email message	sioctypes:MailMessage	
Document set	sioc:Forum	sioctypes:Folder->Container
Project	sioc:Forum	sioctypes:Project->Container
Flow folder	sioc:Forum	sioctypes:Folder->Container
Task	sioc:Forum	sioctypes:Task->Container
Community	sioc:Forum	
Poll	sioctypes:Poll	
Contact list	sioctypes:AddressBook	
Contact	sioc:Post	sioctypes:Contact->Item
User	sioc:User, foaf:Person	
Tag	sioctypes:Tag	

Explain of the table:
- Column "BSCW": Shared Workspace concepts in BSCW;
- Column "SIOC": SIOC concepts mapped from BSCW;
- Column "SIOC(extended)": If the current mapping is not ideal, a new SIOC concept will be proposed, which we think is more suitable to describe shared workspace;
- ->: subclass_of;
- Namespace:
  - sioc: "http://rdfs.org/sioc/ns#"
  - sioctypes: "http://rdfs.org/sioc/types#"
  - foaf: "http://xmlns.com/foaf/0.1/"
  - annotea: http://www.w3.org/2002/01/bookmark#

To develop a SIOC Exporter for BSCW, currently we use XML-RPC to connect to BSCW, and then use Jena to create the SIOC RDF model. In the future we might try to use Python to develop the Exporter within the BSCW source code, this way it will be much more efficient.

Regarding SIOC Exporter for Sharepoint, we are thinking two ways to develop it. The easier way is to use an HTML Wrapper to extract information directly from the HTML web pages, and then use a XSLT translator to transfer the original format to SIOC format. The more efficient way is to retrieve information from the Sharepoint database directly, which will need more efforts.

Nevertheless, we expect in the near future the Shared Workspace supporting tools could have their own SIOC exporter, just like most of the current major news site all support RSS feed.

We are now still working on developing of Shared Workspace Explorer. Currently we have been able to retrieve core SIOC data from BSCW, and to provide a basic Facet Browsing Interface (as shown in figure 4) [9].
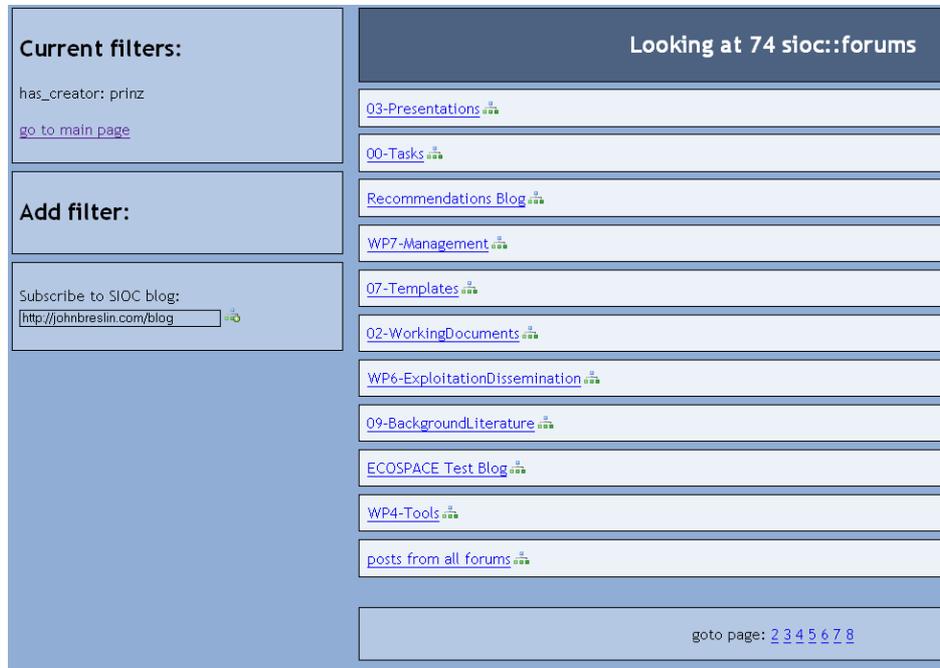

Figure 4: Facet Browsing Interface of Shared Workspace Explorer

Conclusion and future work

Our approach provides an extensible and scalable way of aggregating information from different shared workspaces. For making systems SIOC-enabled, just two steps are required: building a SIOC-exporter and mapping the concepts from the internal representation/model used by the system to the SIOC ontology. In this way, SIOC plays the role of a meta-language or a reference vocabulary for the overall CWE domain and provides the knowledge infrastructure upon which a whole family of tools could be built. Although the prototype is still under developing, it has show great potential to use SIOC to interconnect shared workspaces.

An exporter for Sharepoint is currently being developed and conceptual mapping of Sharepoint concepts to SIOC are proposed in a similar way as presented in table 1 for BSCW. Meanwhile, in order to better fit description of shared workspaces, we are continuing to refine the SIOC Type Module and provide a more dedicated explorer interface for Shared Workspaces. At present we did not address the issues of access control, security and confidentiality, we believe these are important questions needed to research further in the future.

Except SIOC Explorer, there are still a number of different architectures how SIOC can be used for Shared Workspaces. Potential architectures might include:
- Distributing queries to Shared Workspaces - the information is requested from Shared Workspaces at user query time. Queries are built and sent to Shared Workspaces (which need to have SIOC-aware query interface

created before that), which then return an answer information in RDF using the SIOC ontology.

- Importing SIOC into Shared Workspaces - one (or some) of the Shared Workspaces could import the SIOC information created by other Shared Workspaces. This way this one Shared Workspace acts as a master that can also incorporate content from those other, possibly smaller Shared Workspaces.

Reference

1.   Wikipedia, *Shared Workspace entry*. 2007, http://en.wikipedia.org/wiki/Shared_workspace.
2.   BSCW, *BSCW Home Page*. 2007, http://bscw.fit.fraunhofer.de/.
3.   Sharepoint, M., *Microsoft Sharepoint Products and Technologies*. 2007, http://www.microsoft.com/sharepoint/default.mspx.
4.   Hupfer, S., et al. *Reinventing team spaces for a collaborative development environment*. in *"Beyond Threaded Conversation" Workshop, CHI 2005*. 2005.
5.   SIOC, *SIOC Home Page*. 2007: http://sioc-project.org/.
6.   Breslin, J.G., et al. *Towards Semantically-Interlinked Online Communities*. in *Proceedings of the 2nd European Semantic Web Conference (ESWC '05)*. 2005. Heraklion, Greece: LNCS.
7.   Uldis Bojārs and J.G. Breslin, *SIOC Ontology Specification*. 2007, http://rdfs.org/sioc/spec/.
8.   Bojārs, U., J.G. Breslin, and A. Passant, *SIOC Application*. 2007, http://rdfs.org/sioc/applications/.
9.   Heitmann, B. and E. Oren, *Leveraging existing web frameworks for a SIOC explorer to browse online social communities*, in *ESWC Workshop on Scripting for the Semantic Web*. 2007.