SocioBiblog: Enabling Communication on Bibliography with Semantic Blogging

Aman Shakya* shakya_aman@grad.nii.ac.jp
Hideaki Takeda* takeda@nii.ac.jp
Vilas Wuwongse‡ vw@cs.ait.ac.th
Ikki Ohmukai* i2k@nii.ac.jp

*National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
‡Asian Institute of Technology, Klong Luang, Pathumthani 12120, Thailand

Abstract

Sharing of information about publications is very important in a research community. The paper describes a system which demonstrates how semantic blogging can be used for the purpose. It incorporates SWRC ontology into blogging for entering metadata about publications; facilitates commenting on publications and provides a decentralized aggregation mechanism to aggregate publications in the community. RSS aggregation has been extended to handle metadata in BuRST feeds. The system uses FOAF links of authors and friends to explore the social network of the research community and gather RSS/BuRST feeds.

Keywords

Blog, semantic web, bibliographic metadata, information sharing, RSS aggregation, SWRC, BuRST, FOAF

1. Introduction

Communities of researchers are good examples of online communities spread worldwide. Researchers need to share bibliographic information and comment about publications. As research activities are highly autonomous, we need a decentralized technology for this. Blogging is a popular mechanism for easy publishing on the web. However, currently blogs can only publish unstructured passages of text. The semantic web offers a decentralized platform to publish structured information. Semantic blogging is a technology that builds upon blogging and adds semantic structure to blog items [1]. SWRC (Semantic Web for Research Communities) [3] is an ontology for modeling entities of research communities and their relationships. The semantic web also provides popular technologies such as FOAF(Friend of a Friend) which help in binding our research community together.

2. Example Scenario

Fig. 1 illustrates an example scenario. Researcher 'A' publishes information about his publication. Another researcher 'B' has some comments. B's entry (comment on X) acts as an annotation to A's publication. At the same time, a trackback ping is also sent to A's blog. SocioBiblog aggregates posts by friends of the blog-owner by subscribing BuRST feeds from the blogs of people in the blogroll. BuRST(Bibliography Management using RSS Technology) [2] is a lightweight specification for publishing bibliographic information. The aggregated feeds can be searched. When a blog entry about any publication is opened, latest posts by the authors of the publication are also listed. In the example, while viewing X, other posts by the co-authors P and Q are listed.

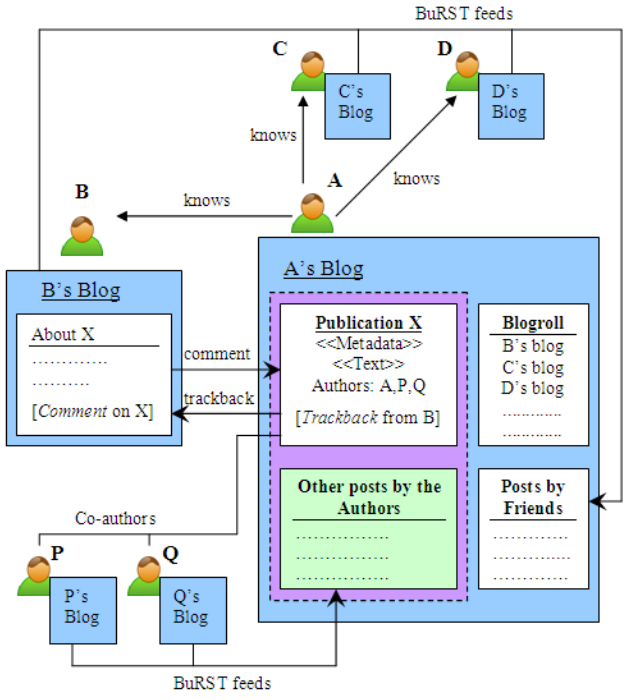

Fig. 1: *Example scenario*

3. Implementation

Fig. 2 shows the architecture of the system. The *publishing system* facilitates publishing blog entries and metadata about publications based on the SWRC ontology. The metadata is embedded in BuRST feeds. It also publishes FOAF profile of the blog-owner. The *aggregation system* aggregates publications from multiple blogs. RSS/BuRST feeds to be aggregated may be retrieved from FOAF profiles. The aggregated feeds can be search by metadata. SocioBiblog can easily be introduced in the current web to co-exist with existing systems. Standard metadata formats like SWRC, FOAF and BuRST have been adopted to ensure compatibility and interoperability.

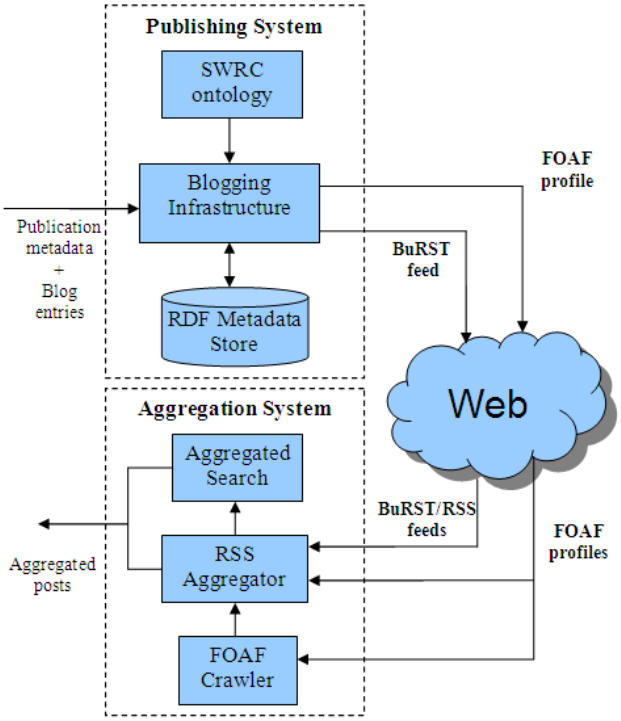

Fig. 2: *System architecture of SocioBiblog*

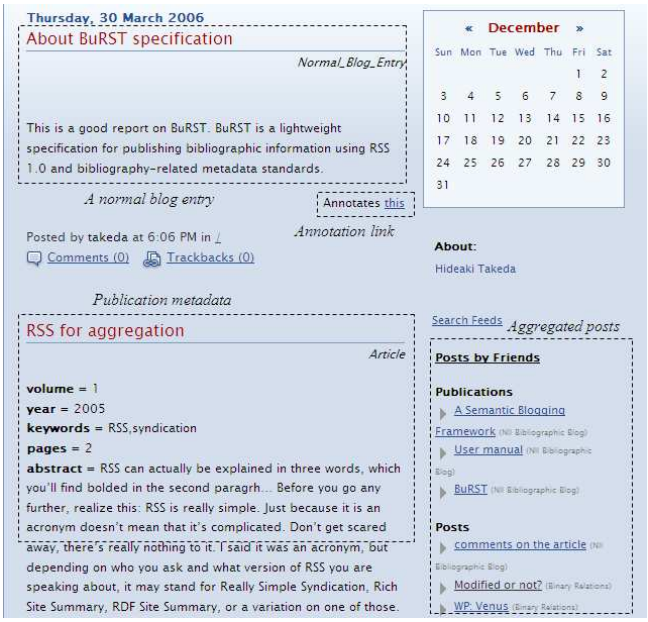

Fig. 3: *The blog interface*

3.1 Publishing

The semantic blog provides metadata entry forms for different SWRC publications. Fig. 3 shows a blog with both a normal blog entry and a publication with some metadata. Blojsom[1] has been used as the blogging platform. Metadata in RDF format is stored in a MySQL database using the Jena[2] framework.

Blogging has been made convenient by employing javascript bookmarklet which captures the title, URL, and trackback ping URL. The blog entry acts as a comment or annotation to the entry pointed by the URL. An "annotates" property has been introduced to model this forward link. A trackback ping is also sent to the annotated blog entry if desired.

A web-based interface to maintain the FOAF profile of the blog-owner has also been provided in the blog itself. The profile serves as a blogroll too. The information is maintained in a simple readable text file which can be easily understood and updated by the user.

3.2 Aggregation

The system aggregates BuRST/RSS feeds from multiple blogs. The latest publications and posts by friends of the blog-owner are aggregated in the blog itself as shown in Fig. 3. If feeds are in BuRST format, publications can be distinguished from other posts and SWRC metadata can be extracted. When a publication is opened, RSS feeds of the authors are only downloaded. Thus, when a user reads about a publication, he/she is also presented a list of other publications by the same authors.

The RSS aggregator makes subscriptions for RSS feeds from the friends list of the blog owner. A FOAF profile includes FOAF links and optionally RSS feed URLs. When a single publication entry is opened, RSS feeds of the authors are only aggregated. The Flock[3] RSS aggregator has been extended to handle BuRST feeds.

The Elmo scutter[4] has been used as a *FOAF crawler* to find out FOAF links of publication authors. The FOAF crawler traces rdfs:seeAlso elements to gather FOAF profiles into a database. The crawled FOAF database may be searched to find out FOAF links of authors. The system provides an interface to search the FOAF link of a person by entering his/her name.

The aggregated BuRST feeds in a blog can be searched and sorted by SWRC fields like title, author, type, etc. Search is done by filtering the aggregated RSS feeds on the basis of search criteria provided through a web-based form. The *aggregated search* does a metadata based search over multiple blogs in the community.

References

[1] S. Cayzer. Semantic blogging and decentralized knowledge management. *Communications of the ACM*, 47(12):48–52, December 2004.

[2] P. Mika. Bibliography management using RSS technology (BuRST). http://www.cs.vu.nl/ pmika/research/burst/BuRST.html, 2005.

[3] Y. Sure, S. Bloehdorn, J. H. Peter Haase, and D. Oberle. The SWRC Ontology semantic web for research communities. In *Proceedings of the 12th Portuguese Conference on Artificial Intelligence - Progress in Artificial Intelligence (EPIA 2005)*, pages 218–231. Springer, December 2005.

[1] http://blojsom.sourceforge.net
[2] http://www.hpl.hp.com/semweb/jena.htm
[3] http://flock.sourceforge.net
[4] http://www.openrdf.org/doc/elmo/users/index.html