

SocioBiblog: A Decentralized Platform for Sharing Bibliographic Information

Aman Shakya¹, Hideaki Takeda¹, Vilas Wuwongse², and Ikki Ohmukai¹,

¹ National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku,
Tokyo, 101-8430, Japan

{ shakya_aman, takeda, i2k }@nii.ac.jp

²Asian Institute of Technology, Klong Luang, Pathumthani
12120, Thailand
vw@cs.ait.ac.th

Abstract. Sharing of bibliographic information is very important in research communities. SocioBiblog is a semantic blogging system that provides a decentralized environment for this. The SWRC ontology has been used for adding metadata about publications in blogs. SocioBiblog aggregates publications from the social network neighborhood and co-authors of a researcher. RSS aggregation has been extended to handle embedded publication metadata in BuRST feeds. Interoperability with other systems has been maintained by adopting standard formats. The aggregated collections may be searched and filtered flexibly by metadata criteria. The results can be redistributed as new feeds. Thus, a decentralized ecosystem can be formed where each unit can publish, aggregate and redistribute information.

Keywords: Decentralized information sharing, semantic blog, bibliographic metadata, RSS aggregation, SWRC, BuRST.

1 Introduction

Communities of researchers are good examples of online communities. Researchers need to share information about their publications, bookmark interesting publications and write comments on them. Moreover, they always need to be up-to-date about latest publications. Research activities are highly autonomous and thus, we need a decentralized mechanism for such information sharing. We may expect to aggregate highly relevant resources from closely linked people in the social network. Pooling information from various relevant sources and filtering them based on metadata can help in creating useful tailored information channels. Blogging is a popular decentralized technology for easy publishing on the web. However, bibliographic information needs to be structured in a standard way and semantic blogging [1] provides this by combining the features of blogging and the Semantic Web. SocioBiblog is a prototype semantic blogging system enabling researchers to share bibliographic metadata, aggregate information about publications from different sources and filter and redistribute such aggregated information.

2 The SocioBiblog Decentralized Platform

SocioBiblog serves information sharing in the community with the following features.

- **Decentralized Information Sharing.** The system is fully decentralized and supports both publishing and aggregation of information. We propose personalized aggregation of bibliographic metadata on individual blogs.
- **Social Network based Aggregation.** Bibliographic metadata is aggregated from the social network neighborhood of the blog-owner, up to the depth of two levels. Further, posts by the co-authors of publications are also listed.
- **Interoperability.** Bibliographic information feeds can be aggregated from different systems. Moreover, metadata from different systems can be scraped and quoted in blog entries. Comments can be posted on the publications.
- **Integration and Filtering.** Information from different sources can be integrated using a common semantic standard. The collection can be filtered by metadata criteria and the results can be redistributed as a new feed.

SocioBiblog has been implemented as two sub-systems. The *publishing system* facilitates publishing blog entries with metadata about publications. The SWRC (Semantic Web for Research Communities) ontology [2] has been used for metadata. BibTeX scrapers extract bibliographic metadata from blogs and bibliographic sites. Metadata is exported in BuRST (Bibliography Management using RSS Technology) feeds[3]. The *aggregation system* aggregates feeds from multiple sources. RSS/BuRST feed URLs are retrieved from FOAF profiles. A search interface allows the user to filter aggregated publications by defining various metadata criteria. The aggregated and filtered BuRST feed thus obtained is exposed as a new BuRST feed.

3 Conclusion

SocioBiblog can act as an active unit which can both publish and aggregate contents to facilitate information flow in the community forming a decentralized ecosystem. The network of such units allows flexible handling of information channels and offers the potential of evolution. The Semantic Web provides the base for sharing structured information and interoperating with each other and existing systems.

References

1. Cayzer, S.: Semantic Blogging and Decentralized Knowledge Management. Communications of the ACM, Vol. 47 (2004) 48-52
2. Sure, Y., Bloehdorn, S., Haase, P., Hartmann, J., Oberle, D.: The SWRC Ontology - Semantic Web for Research Communities. In: Bento, C., Cardoso, A., Dias, G. (eds.): Proceedings of the 12th Portuguese Conference on Artificial Intelligence, Vol. 3803. Springer, Covilha Portugal (2005) 218 – 231
3. Mika, P.: Bibliography Management using RSS Technology (BuRST) (2005) <http://www.cs.vu.nl/~pmika/research/burst/BuRST.html>

SocioBiblog - Demo Description

1 Publication and Commenting

SocioBiblog¹ provides metadata entry forms for different SWRC publications. BibTeX snippets can also be imported directly to fill the forms. Publication metadata entry is exported in SWRC, BibTeX formats and BuRST feeds. Blog entries can quote publication metadata from other blogs and bibliographic sites and comment on them. It supports bidirectional commenting using trackback and annotation.

Blog this bookmarklet. Commenting has been made convenient by employing javascript bookmarklet. The bookmarklet captures the title, URL, trackback ping URL of the blog entry being annotated and any highlighted text.

BibTeX Scraping. *Blog this* also scrapes out BibTeX information if available. SocioBiblog currently provides scrapers for SocioBiblog instances, the ACM digital library and a generic BibTeX scraper which works for any web page that contains BibTeX snippet (applicable for Citeseer, DBLP, BibSonomy, CiteULike, etc).

Metadata Search. Bibliographic metadata in a blog can be searched by specifying various metadata values. It also searches into the metadata quoted from other sources.

Blogroll and FOAF Profile. The blogroll is also exported as a FOAF profile.

2 BuRST/RSS Aggregation

Social Network based Aggregation. The system aggregates BuRST/RSS feeds from the social network neighborhood of the blog-owner, up to two levels deep – directly linked friends and friends of the friends. The path to the linked friends is also shown.

Feeds from other bibliographic systems and repositories can also be added to the blogroll. For example, SocioBiblog can aggregate BuRST feeds from BibSonomy.

The latest aggregated publications and posts are displayed alongside in the blog. The system separates publications and non-publications. Further, when a publication is opened, BuRST/RSS feeds of the co-authors are downloaded and shown alongside.

FOAF Crawler. A crawled FOAF database can be searched for FOAF links.

Aggregated search, filtering and mixing. The BuRST feeds aggregated from multiple sources can be searched and sorted by SWRC fields like title, author, type, etc. The system also searches the metadata quoted in comments on publications. Commented publications are distinguished from original publications. The aggregated and filtered results obtained are exported as new BuRST feed. The user can subscribe to this BuRST feed in his/her RSS reader and receive the latest desired updates. The feed can further be combined with other sources to form new channels.

¹ An online demo is available at <http://dutar.ex.nii.ac.jp/>