

# A Web Mashup for Social Libraries

Domenico Gendarmi, Filippo Lanubile

University of Bari

Dipartimento di Informatica

Via E. Orabona, 4 - 70125 Bari

+390805442286

{gendarmi, lanubile}@di.uniba.it

## ABSTRACT

User-generated content on the Social Web is often locked within information silos. Inadequate APIs or, worst, the lack of APIs obstruct reuse and prevent the opportunity to integrate similar content from different communities. In this paper we present a Web mashup which combines information from different social libraries. Aggregated information, including both classic book metadata and user-generated content, is represented as linked data in RDF to allow machine computation and foster reuse among different applications.

## Categories and Subject Descriptors

H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces – Web-based interaction.

## General Terms

Design, Human Factors.

## Keywords

social libraries, mashup, linked data, social web, semantic web.

## 1. INTRODUCTION

The rise of the Social Web, also known as Web 2.0 [8], created new challenges and opportunities for user interface integration. In such a context, Web mashup has emerged as a dynamic approach for composing content and functionalities originating from disparate web sources [10].

Nowadays, plenty of user-generated content is available on the Social Web. However, these forms of collaborative contributions are restricted to one single application and do not let data flow from one point on the Web to another. Furthermore, most applications' creators on the Social Web are quite reluctant in providing programmatic access to user generated content which is hosted within their web sites. As a consequence, often no APIs are made available. Although some programming interfaces are sometimes provided, third-party applications are only allowed to retrieve information concerning one user at time, thus hampering the reuse of community-generated content and the integration of similar content from different communities.

Social libraries, also referred to as "Libraries 2.0" [7], can be regarded as an innovative form of library services which replace traditional ones. Social libraries attempt to harness users in the design and implementation of the application by encouraging feedback and participation. As part of the Social Web, these

applications are also affected by inadequate APIs or, worst, the lack of APIs. We thus paradoxically face a situation where one book can be tagged and reviewed in Amazon, LibraryThing, and Anobii, but there is no way to know that people refers to the same paperback.

In this paper, we illustrate a Web mashup, named Colibrary, which combines classic book metadata (e.g., title, authors, publisher) and user-generated information (e.g., tags and reviews) from different social libraries. The main goal of the mashup is to combine such data from different communities of book readers, as they can be shared rather than locked within 'information silos'. Shared data are structured as linked data, thus represented through semantic web technologies, in a manner that allows machine computation and fosters reuse among different applications.

## 2. COLIBRARY MASHUP

Colibrary initially originated as a Web API to exploit the opportunity to reuse content in the digital library domain and provide other applications with different types of information about books. Since its inception, the basic requirement was the possibility to unambiguously distinguish one single paperback. A straightforward solution was the adoption of the International Standard Book Number (ISBN) as a unique identifier for referring to a specific book (Figure 1).

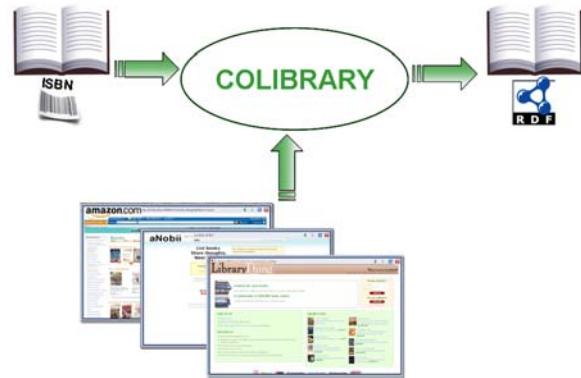


Figure 1. Colibrary Overview

Colibrary then evolved as a server-side mashup [6]. Once a valid ISBN is given as input to Colibrary, data about the specific paperback are retrieved from three large web sites where different community of readers can organize their virtual bookshelves by sharing keywords and reviews. In particular, Colibrary collects the following bibliographic data for a book: title, author(s), editor(s), publication date, number of pages and the cover, whether available through the Amazon Web API. Then, Colibrary supplements such information with social data directly generated from end users of popular applications in the library domain. Specifically, Colibrary retrieves all the tags and reviews assigned

to a specific book by users from Amazon, Anobii and LibraryThing. Among the three providers, only Amazon provides a suitable API to retrieve bibliographic information. Social data concerning both tags and reviews are retrieved through scraping services from any of the content providers.

The integration logic resides on a web server and is executed as a RESTful web service. Given as input the identifier of a specific resource (and other optional parameters), the service returns linked RDF containing bibliographic information and social data coming from different social libraries. The service has been implemented as a set of PHP scripts which (1) handle different type of requests from any client, (2) invoke content providers, and (3) filter the results and finally create the RDF output according to the Linked Data guidelines [4].

The client of the Colibrary Web API has been developed as a simple web-based interface which allows a user to launch a query and obtain an HTML representation of the linked data published by the service. The client has been implemented by means of PHP scripts which invoke the methods provided by the Colibrary web service and present results into the browser.

The main motivation behind using RDF in Colibrary stands in the opportunity to enhance the Social Web experience by promoting the vision of the Web of Data [9]. This vision can be regarded as a simplification of the Semantic Web, as the aspects of automated reasoning are de-emphasized and research focuses on linking data across boundaries. Linked Data relies on mainstream web technologies, like HTTP, URI and RDF to realize the four simple principles, firstly introduced by Tim Berners-Lee [2].

Within the Colibrary web service, once the data has been gathered from the three content providers, an RDF description about a book is generated. Such description reuses existing vocabularies, including Dublin Core for bibliographic data, Holygoat Tag ontology for tags, RDF Review Vocabulary for reviews or comments and FOAF for reviews contributors. New vocabulary terms were introduced within the Colibrary default namespace when no existing term could be reused. Adopting such widespread vocabularies makes Colibrary's RDF data interoperable with data from other sources.

All the query methods provided by the service are available through classic HTTP GET requests. Currently, just the ISBN is a mandatory parameter; if no other parameters are added to the query, the service returns all available data about a specific book. On the other hand, optional parameters can be added to the URI for retrieving only small portions of the full RDF book description. For example, we can retrieve only bibliographic data of the same book, or just social data (tags and/or reviews) through several HTTP GET requests to different URIs.

### **3. CURRENT AND FUTURE RESEARCH**

While many mashups pay more attention to aspects related to presentation, as the main concern is to provide existing information through a rich interactive application, Colibrary mashup is also able to publish integrated data in semantic web formats in order to enable both human and machine consumption. By encoding the output of the mashup as linked data in RDF, Colibrary provides different URIs for all the retrieved information concerning the book. Given as input an ISBN we have one URI to obtain the RDF description of just bibliographic data, one URI to retrieve just social data, and analogously other different URIs for RDF descriptions of tags, reviews or both.

Providing linked RDF data enables any third party to make reference to such information in other RDF statements, opening the way for links between Colibrary and other data sets. Incoming links could be set between URIs provided by Colibrary and other RDF descriptions about books created by other services such as the RDF Book Mashup [3], Revyu [5] and DBpedia [1]. For example, an application like RDF Book Mashup could augment their data sets with social information retrieved using the Colibrary Service, while Revyu could supplement reviews about things identified as books with comments on the same paperback originated from other communities.

As future work, we will enhance current Colibrary RDF descriptions by automatically generating outgoing links to other data sources that exploit Semantic Web technologies to represent information concerning books. Thus, by publishing and linking these data in RDF we might help to populate the Web with links between different data sets, supporting thus the vision of a Web of Linked Data.

### **4. ACKNOWLEDGMENTS**

We would like to thank all the students at the University of Bari, which contributed to the development of the current version of Colibrary.

### **5. REFERENCES**

- [1] Auer, S., Bizer, C., Lehmann, J., Kobilarov, G., Cyganiak, R. and Ives, Z. 2007. DBpedia: A Nucleus for a Web of Open Data. In Proceeding of the 6th International Semantic Web Conference, Springer LNCS 4825, 722-735.
- [2] Berners-Lee, T. 2006. Linked data. W3C Design Issues. <http://www.w3.org/DesignIssues/LinkedData.html>
- [3] Bizer, C., Cyganiak, R., & Gauss, T. 2007. The RDF Book Mashup: From Web APIs to a Web of Data. Workshop on Scripting for the Semantic Web. CEUR Workshop Proceedings, Vol. 248, ISSN 1613-0073.
- [4] Bizer, C., Cyganiak, R. & Heath, T. 2007. How to Publish Linked Data on the Web. <http://www4.wiwiiss.fuberlin.de/bizer/pub/LinkedDataTutorial/>
- [5] Heath, T. & Motta, E. 2008. Ease of interaction plus ease of integration: Combining Web2.0 and the Semantic Web in a reviewing site. Web Semantics 6, 1, 76-83.
- [6] Liu, X., Hui, y., Sun, W., Liang, H. 2007. Towards Service Composition Based on Mashup. IEEE Congress on Services (Services 2007), 332-339.
- [7] Maness, J. M. 2006. Library 2.0 Theory: Web 2.0 and Its Implications for Libraries. Webology, 3, 2.
- [8] Murugesan, S. 2007. Understanding Web 2.0. IT Professional, 9, 4, 34-41.
- [9] Shadbolt, N., Berners-Lee, T., & Hall, W. 2006. The Semantic Web Revisited. IEEE Intelligent Systems, 21, 3, 96-101.
- [10] Yu, J., Benatallah, B., Casati, F., and Daniel, F. 2008. Understanding Mashup Development. IEEE Internet Computing 12, 5, 44-52.