DataConf: Enriching Conference Publications with a Mobile Mashup Application

Lionel Médini
Université de Lyon, CNRS
Université Lyon 1, LIRIS, UMR5205
F-69622, France
+33 4 72 43 16 36
lionel.medini -at- liris.cnrs.fr

Florian Bâcle
Université de Lyon
Université Lyon 1
F-69622, France
florian.bacle
-at- etu.univ-lyon1.fr

Nguyen Hoang Duy Tan
Université de Lyon
Université Lyon 1
F-69622, France
hoang_duy_tan.nguyen
-at- yahoo.com

ABSTRACT
This paper describes a mobile Web application that allows browsing conference publications, their authors, authors' organizations, and even authors' other publications or publications related to the same keywords. It queries a main SPARQL endpoint that serves the conference metadata set, as well as other endpoints to enrich and explore data. It provides extra functions, such as flashing a publication QR code from the Web browser, accessing external resources about the publications, and it can be linked to external Web services. This application exploits the Linked Data paradigm and performs client-side reasoning. It follows recent W3C technical advances and as a mashup, requires few server resources. It can easily be deployed for any conference with available metadata on the Web.

Categories and Subject Descriptors
E.1.2 [**DATA**]: Data structures – *Distributed data structures.*

Keywords
mobile Web, mobile reasoning, mashup, Linked Data, publication browsing.

1. INTRODUCTION
The WWW'2012 conference that was held in Lyon was an occasion for the local Web community to initiate several innovating projects around conference material and Web technologies. We designed a mobile Web application to augment the poster track by representing posters as physical objects on the Web. This app initially targeted conference attendees discovering the 100 posters exposed at this track and allowed them to navigate among poster metadata (title, authors, abstract, keywords…), as well as authors and other publications metadata, using their smartphones and tablets. Each poster was associated with a QR-Code encoding the poster URI; flashing a code led to the corresponding poster view. Posters can also be searched from the application homepage. As the whole conference publication metadata set was available, we extended our application to all the conference publications.

As it targeted a WWW series conference, technical choices were oriented towards a full Web application applying recent/emerging Web technologies and standards. In particular, HTML5 APIs encouraged us designing the application as client-sided as possible. It also takes advantage of Linked Data by enriching publication metadata from several sources. We thus explored the feasibility, using currently available browsers on mobile devices, of: (i) dynamically constructing complex SPARQL queries and sending them to cross-domain endpoints, (ii) representing and allowing browsing among these metadata using a textual and a graphical interface, (iii) locally building, classifying and querying an ontology and (iv) capturing and processing images in JavaScript using the device built-in camera.

In this paper, we present a generic, configurable version of this application. It can easily be deployed for any conference that has its publication metadata available on the Web. Moreover, the WWW'2012 experience showed that a conference schedule was a dynamic artifact: session rooms frequently changed and paper presentations even took place in different session that the ones initially planned. Printing this schedule necessarily led to an obsolete version of this artifact. In order to provide an accurate version of the time schedule, the herein presented version of our app can be coupled with a Web service that exposes the conference schedule, allowing attendees to view the time and location of a paper presentation while browsing its metadata on their smartphones. The application homepage can be accessed at http://dataconf.liris.cnrs.fr/ and the sources are downloadable at https://github.com/ucbl/DataConf.

This paper is organized as follows: we first present the two main functions related to metadata querying and processing, followed by the application other functions. We then conclude and present evolution perspectives of this app.

2. Linked Metadata browsing
A publication homepage shows its metadata, based on the data received from the endpoint serving the conference metadata set. Semantic Web Conference (SWC) and Semantic Web for Research Communities (SWRC) ontology models. Metadata are enriched so that authors and keywords are navigable. Clicking on an author provides information such as name, most probable homepage, affiliation organization, co-authors, publications in the desired track/conference and other publications. Among these data, the organization, co-authors and publications items are navigable. The "organization" page displays the name, homepage and other authors referred in DBLP for this organization, and the "Other publications" page displays DBLP-specific metadata about the publications (title, DOI URI, year, publication type and name of the conference / journal). The co-authors list of an author is a typical Linked Data usage example of the available data: it is generated after parsing the other publications list and gathering each all authors of each publication. Another feature of our application enables users to browse the DBLP database from an

WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

author or publication view. Such a feature seemed interesting for evaluating the interest conference attendees would find in our interface for browsing large publications databases. To extract and enrich metadata, we use 3 SPARQL endpoints:

SWDF[1] contains (among other conferences) the metadata about the WWW'2012 conference in the SWRC and SWC formats. We rely on this endpoint for accessing publication metadata. SWDF is accessed using the CORS [1] recommendation working draft.

DBLP / L3S[2] provides a wider dataset, used to enrich SWDF paper metadata, such as other publications of an author; but it lacks publications keywords. As L3S is not CORS enabled, this server is accessed using JSONP [2].

DuckDuckGo! RDF data access[3] is a search engine that can be queried in SPARQL. We also use it for metadata enrichment. As we do not want to retrieve all data satisfying a query but get the most probable value (e.g. get an author's or organization homepage), we use its "I'm feeling ducky" feature, with the risk to end up on another page. DuckDuckGo! is accessed using JSONP.

3. Local ontology enrichment and reasoning

Our application also allows browsing by keywords. As DBLP does not provide information about publications keywords, navigable keywords list is only accessible for publications that have an entry in SWDF. For each of those publications, our application dynamically constructs an ontology composed of the publication, its keywords, its authors, their other publications and the keywords of these publications.

Once the data gathered, we launch a classification process using the OWLReasoner [3] JS inference engine in order to be able to launch SPARQL queries on keywords (e.g. suggesting other publications that match the same keywords, more generic ones, or publications of the same authors that match the user's keywords). For this, we rely on the taxonomy of the conference keywords (extracted from the call for papers pages) and the locally stored user's keywords in this taxonomy.

4. Other features

Publication and author search: forms allow accessing a publication or author's description page. In each form, automatic suggestions are provided by sending asynchronous SPARQL requests to the metadata server. Some requests require recursively querying transitive properties (e.g. swc:isSubEventOf). This is only done for endpoints that support SPARQL 1.1. In order to avoid overflowing SPARQL 1.0 servers with chained requests, for the latter, we use other relations existing in the ontology.

View metadata as graphs: users can choose to view and navigate in tree-based graphical representations of the publications, authors or keywords. The graphical representation is built using the JavaScript Infovis toolkit[4].

Flashing a QR-code inside the Web application: in browsers that support the getUserMedia API or the capture input type, users can take a picture of the QR representing a paper URI and send it to the server. It is then decoded using the ZXing library and the user is redirected to the paper homepage. Of course one can also flash the QR codes using a native application to retrieve the same publication homepage.

External services: we also introduced the possibility of linking the application to extra services regarding the conference publications. For instance, we worked with the company who designed the WWW'2012 schedule database, so that our app can add to the displayed metadata of a paper, the time and location of the session where it will be presented.

Extra information access: URIs of extra material (detailed PDF, video presentation…) can be stored in our database and displayed in the homepage of the corresponding paper. For WWW'2012 posters, the 2-page PDF papers published in the proceedings are accessible by clicking on the posters ids.

5. Conclusion

Our mobile Web application proposes various features among which querying and browsing conference publications. It has been designed w.r.t. recent Web standards and technologies and so that as much computation as possible is done on client side: we use cross-domain requests to access and enrich the metadata from three different SPARQL endpoints; during the navigation, the application constructs a local ontology and uses it for suggesting publications to the user. Server-side dynamic content is only used to store extra data about the publications, such as the keyword ontology. Additional features provided by our app on different browsers may vary depending on the technologies they embed, but the core functions (i.e. metadata navigation) is available on all browsers.

Since the last edition of the WWW Conference, we refactored the application. It is now both generic and modular. Using an XML configuration file, a conference chair can now specify several URIs, referencing the conference metadata set, the different tracks exposed to the application users and other data. We tested our app with several datasets related to different conferences available on the SWDF website. During the demo, we will show how to write a configuration file to tailor our application to a conference once its metadata set has been made available on the web. An instance will probably be deployed to query the WWW'2013 dataset when it is available.

6. REFERENCES

[1] Cross-Origin Resource Sharing. W3C Candidate Recommendation, 29 January 2013 : http://www.w3.org/TR/cors/

[2] Ippolito, B., 2005. Remote JSON – JSONP. Available at: http://bob.ippoli.to/archives/2005/12/05/remote-json-jsonp/

[3] OWLReasoner JavaScript Inference engine. Available at: http://code.google.com/p/owlreasoner/

[1] http://data.semanticweb.org/

[2] http://dblp.l3s.de/d2r/

[3] http://duckduckgo.com/1/c/RDF_data_access

[4] http://thejit.org/