

A Distributed Semantic Microblogging Platform

Alexandre Passant¹, Tuukka Hastrup², Uldis Bojārs², John Breslin²

¹ LaLIC, Université Paris-Sorbonne,
28 rue Serpente, 75006 Paris, France
`firstname.lastname@paris4.sorbonne.fr`

² DERI, National University Of Ireland,
Galway, Ireland
`firstname.lastname@deri.org`

Abstract. The application showcases the ideas of a distributed, Semantic-Web enabled microblogging architecture, providing a way to leverage this new Web 2.0 practice to the Semantic Web.

Key words: Microblogging, SIOC, Data Portability, Linked Data Web

Microblogging is one of the recent social phenomena of Web 2.0 but unlike blogs or wikis has not yet been leveraged to the Semantic Web. To achieve this goal, we designed a semantically-enabled distributed architecture for semantic microblogging, which relies on an open world of publishing clients and aggregation servers that exchange data modelled in RDF.

When users write microblog posts within their clients, RDF files are created on the client webservers, describing the posts using FOAF [3] and SIOC [2], and pushed live to a number of aggregation servers. Thus, the user really owns his data and can reuse it locally for other purposes, either browsing or merging with other RDF data, while aggregation servers are mainly dedicated to providing a browsing interface for shared communities. To model updates, we extended the SIOC types module [1] with a `MicroblogPost` class, as well as `Microblog` to model the service itself.

Thanks to the use of existing libraries, the code of both the client and the server is really light¹. The client uses the SIOC PHP API² to create the RDF files from an HTML form submission, and is only 57 lines of code. This file is pushed to some aggregation servers (chosen from the list of servers stored in the client configuration file) using CURL. Regarding the server, we rely on ARC2³ which provides a lightweight environment for developing RDF-based applications in PHP. The server uses the SPARUL `LOAD` instruction to store received updates in the server backend store, and a single SPARQL query to render a view of public updates. To make the interface fancier, we use Exhibit [4] to display a faceted view of these latest updates. These facets include date and author but also some user-defined data. Indeed, the server features a preprocessor

¹ <http://code.google.com/p/smob/>

² <http://wiki.sioc-project.org/index.php/PHPEXportAPI>

³ <http://arc.semsol.org>

that allows users to use some *semantic hashtags* in their updates. The current implementation includes a GeoNames⁴ mapping, allowing users to use tags like `#geo:paris_france` to retrieve the URI of the related resource, thus providing a way to leverage location-based microblogging to the Linked Data Web. Consequently, this mapping permits the use of the geographical rendering part of Exhibit, as shown on Fig. 1 Other simple topics can be extracted with a similar processor and can also be linked to DBPedia with a given prefix.

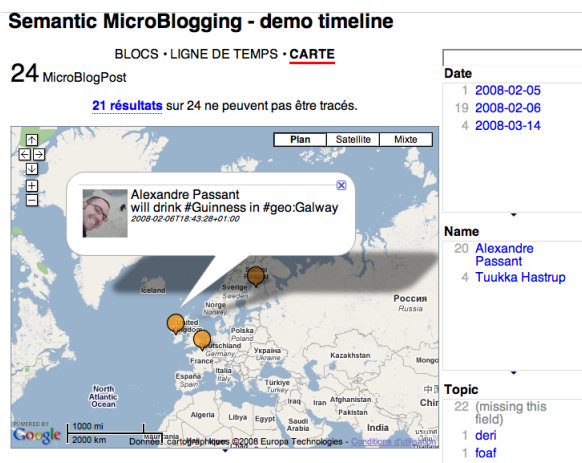


Fig. 1. Geographical faceted browsing of updates with Exhibit

This material is based upon work supported by Science Foundation Ireland under grant number SFI/02/CE1/I131.

References

1. Uldis Bojārs, John Breslin, Aidan Finn, and Stefan Decker. Using the Semantic Web for Linking and Reusing Data Across Web 2.0 Communities. *The Journal of Web Semantics, Special Issue on the Semantic Web and Web 2.0 (Forthcoming)*, 2008.
2. John G. Breslin, Andreas Harth, Uldis Bojars, and Stefan Decker. Towards Semantically-Interlinked Online Communities. In *Proceedings of the Second European Semantic Web Conference, ESWC 2005, May 29–June 1, 2005*, Heraklion, Crete, Greece, 2005.
3. Dan Brickley and Libby Miller. FOAF Vocabulary Specification. Namespace Document 2 Sept 2004, FOAF Project, 2004. <http://xmlns.com/foaf/0.1/>.
4. David Huynh, David Karger, and Rob Miller. Exhibit: Lightweight structured data publishing. In *16th International World Wide Web Conference*, Banff, Alberta, Canada, 2007. ACM.

⁴ <http://geonames.org>