

Open, Distributed and Semantic Microblogging with SMOB*

Alexandre Passant¹, John G. Breslin^{1,2}, and Stefan Decker¹

¹ Digital Enterprise Research Institute, National University of Ireland, Galway
`firstname.lastname@deri.org`

² School of Engineering and Informatics, National University of Ireland, Galway
`john.breslin@nuigalway.ie`

Abstract. This demo paper introduces SMOB, an open, distributed and semantic microblogging system using Semantic Web technologies (RDF(S)/OWL and SPARQL) and Linked Data principles. We present its ontology stack and related annotations, its distributed architecture, and its interlinking capabilities with the Linking Open Data cloud.

Keywords: Social Web, Semantic Web, Linked Data, Microblogging, Distributed Architectures.

1 Introduction

As many Web 2.0 services, microblogging¹ applications suffer from various limits. On the one hand, their close-world architecture strengthens the Web 2.0 data silo issues and makes microblog posts difficultly interoperable with other applications. On their other hand, their lack of machine-readable metadata and semantics entails that microblog posts cannot be fully exploited for advanced querying and reuse. In this demo paper, we present SMOB, an open-source framework for open, distributed and semantic microblogging that relies on Semantic Web technologies and Linked Data principles [1] to solve the aforementioned issues. It provides means to enable machine-readable description of microblog posts, and defines an open architecture where anyone can setup his own service, keeping control over his own data. To achieve this goal, SMOB relies on: (1) an ontology stack to represent microblogs (and their posts) combined with RDFa annotations to represent such data, (2) an open architecture based on distributed hubs that communicate and synchronise together using SPARQL/Update and its related HTTP protocol, and (3) interlinking components, so that microblog posts can be linked to existing resources from the Semantic Web, and especially from the Linking Open Data Cloud².

* The work presented in this paper has been funded in part by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion 2).

¹ Microblogging consists in sharing short (generally under 140 characters) status update notifications. It notably became popular *via* Twitter — <http://twitter.com>.

² <http://richard.cyganiak.de/2007/10/lod/>

2 SMOB — Semantic MicroBlogging

2.1 The SMOB Ontologies Stack

In order to semantically-enhance microblogging services and microblog posts, there is a need for: (i) ontologies to represent users, their features (such as names, homepages, etc.) and their related social networking acquaintances; and (ii) ontologies to represent microblogging posts (and microblogs), including particular features such as hashtags (`#tag` patterns included in microblog posts to emphasise particular words), replies, and some contextual information (geographical context, presence status, etc.). Moreover, there is a need to link microblog posts to existing resources from the Semantic Web, in order to represent topics discussed in these posts without any ambiguity.

Regarding the first aspect, we naturally relied on FOAF — Friend of a Friend [3] — as it provides a simple way to define people, their attributes and their social acquaintances. Furthermore, SMOB users can reuse their existing FOAF profiles so that existing information about themselves is automatically linked from their posts. Then, to describe microblog posts and microblogs, we relied on and extended SIOC — Semantically-Interlinked Online Communities [2]. We also used OPO — Online Presence Ontology [5] — to describe users' presence information, such as geolocation. Finally, we relied on MOAT — Meaning Of A Tag [4] — to represent links between hashtags and Semantic Web resources.

Combined together, these ontologies form a complete stack for semantic microblogging, depicted in Fig. 1, and each microblog post generated with SMOB is provided in RDFa using these ontologies.

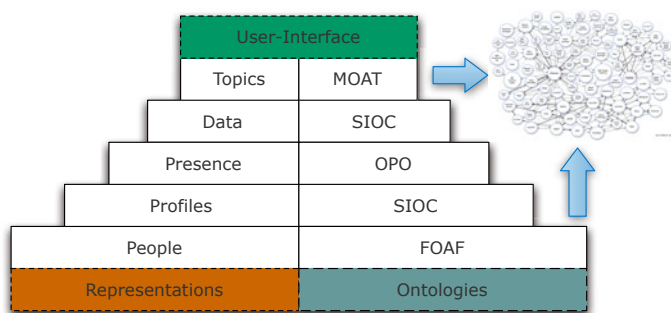


Fig. 1. The SMOB ontologies stack

2.2 A Distributed Architecture

The SMOB architecture is based on distributed hubs that act as microblogging clients and communicate each other to exchange microblog posts and following/followers notifications. That way, there is no centralised server but rather a set of hubs that contains microblog data and that can be easily replicated and extended, also letting users control and own their status updates. Hubs communicate each others and exchange content via HTTP using SPARQL/Update

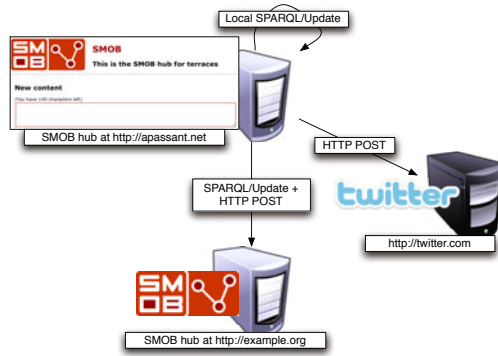


Fig. 2. Communication between SMOB hubs using SPARQL/Update

(the Update part of SPARQL, currently being standardised in the W3C³) and its LOAD clause (Fig. 2).

When a new microblog post is created, it is immediately stored in the user’s hub and sent to the hubs of his followers using the aforementioned principles, so that they immediately receive and store it in their hub. SMOB also enables cross-posting to Twitter, and each hub provides its own SPARQL endpoint so that the data it contains can be easily queried and mashed-up with other data.

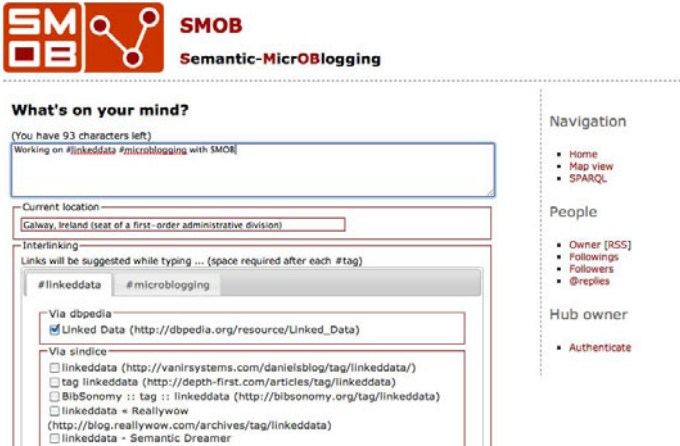


Fig. 3. The SMOB publishing interface and its interlinking components

2.3 Integrating Microblogging in the Linking Open Data Cloud

Finally, SMOB provides features for integrating microblog posts into the Linking Open Data cloud, in addition to the aforementioned reuse of FOAF profiles. On the one hand, it features a set of wrappers that automatically suggest URIs of

³ <http://www.w3.org/TR/sparql11-update/>

existing resources for each hashtag used in microblog updates, relying on services such as Sindice⁴ or DBpedia⁵. Once these suggestions are validated by the author, the mappings are modelled with MOAT and exposed in the post with RDFa. In addition, new wrappers can be deployed for integration with other sources, such as corporate knowledge bases. On the other hand, SMOB provides auto-completion features for geolocation information, relying on GeoNames⁶ data. Fig. 3 depicts its user-interface, where these two interlinking components can be observed.

Using such interlinking, new features can be enabled, such as real-time geolocation or topic-based discovery of microblog data using SPARQL queries.

3 Conclusion

In this demo paper, we gave a short overview of SMOB, a system for Open, Distributed and Semantic Microblogging. SMOB is available at <http://smob.me> under the terms of the GNU/GPL license, and can be setup on any LAMP — Linux, Apache, MySQL, PHP — environment.

References

1. Berners-Lee, T.: Linked Data. Design Issues for the World Wide Web, World Wide Web Consortium (2006), <http://www.w3.org/DesignIssues/LinkedData.html>
2. Breslin, J.G., Harth, A., Bojars, U., Decker, S.: Towards Semantically-Interlinked Online Communities. In: Gómez-Pérez, A., Euzenat, J. (eds.) ESWC 2005. LNCS, vol. 3532, pp. 500–514. Springer, Heidelberg (2005)
3. Brickley, D., Miller, L.: FOAF Vocabulary Specification. Namespace Document, FOAF Project (September 2, 2004), <http://xmlns.com/foaf/0.1/>
4. Passant, A., Laublet, P., Breslin, J.G., Decker, S.: A URI is Worth a Thousand Tags: From Tagging to Linked Data with MOAT. *International Journal on Semantic Web and Information Systems (IJSWIS)* 5(3), 71–94 (2009)
5. Stankovic, M.: Modeling Online Presence. In: *Proceedings of the First Social Data on the Web Workshop*. CEUR Workshop Proceedings, vol. 405. CEUR-WS.org (2008)

⁴ <http://sindice.com>

⁵ <http://dbpedia.org>

⁶ <http://geonames.org>