Using Social Data as Context for Making Cultural Heritage Recommendations: An Ontology based Approach

Salma Noor, Kirk Martinez,

School of Electronics and Computer Science
University of Southampton,
Southampton, UK, SO17 1BJ
{sn07r, km}@ecs.soton.ac.uk

Abstract. This research explores the potential of utilizing social-web data as a source of contextual information for searching and information retrieval tasks. While using a semantic and ontological approach to do so, it works towards a support system for providing adaptive and personalized recommendation of Cultural Heritage Resources.

Keywords: Ontology, Context, Semantic Web, Web 2.0, user interest modeling, tagging, Recommender Systems, Cultural Heritage, linked data.

1 Research Challenges and Motivation

Motivated by the need of a generalized context model for recommendations we aim to develop a semantic support system which can assist in providing brief and precise, high valued, context dependant, recommendations for the semantic web. This brings forward following research challenges. (1)Discovering new user information, (2)Unobtrusive information gathering, (3)Basic concept location and (4)Removing Vocabulary gap, especially in domains like the cultural heritage. We intend to answer these challenges in the light of ongoing research in the fields of Social and Semantic Web.

2 Proposed Methodology and Contributions

The architecture proposed below intends to capture user interest from different social web identities owned by a person. The idea is to utilize this information as context for making recommendations. This model is then utilized in recommending cultural heritage resources that might be of interest for the user. The use of cultural heritage as an interest domain here, aids our research due to the existence of semantically marked-up datasets from previous projects (e.g. eCHASE [2]). This provides a strong base for testing our proposed ideas. Following are the main modules in our proposed architecture.

- *Identifying a user's profile across social networks:* The first module of the system identifies user's different social identities across the web. This will help in deciding where to extract user's data from. The "otherme" method from Google's Social Graph API[1] will be utilized to achieve this alongside other techniques.
- *Data Extraction Module:* Describes a set of tag extraction techniques mostly utilizing public APIs provided by the sites and some scripts.
- *Data Filtering Module:* Specifies a set of filters for cleaning the tag clouds and making it usable for the next step.
- *Concept mapping and Ontology mapping:* Module Takes the set of filtered tags and equips them with semantics by categorization and ontology mapping. Global ontologies, YAGO [1] and Dbpedia [4] are used for conceptualization of tags. While later these concept-tags are mapped to a domain specific ontology The CIDOC Conceptual Reference Model (CRM) [3], to show how the concept-tags can be used in specific domains.
- *User interest profile:* This module takes as an input a set of 'concept-tags' and applies a concept expansion algorithm on them to make them more suitable for recommendations.
- *Recommender system, CH repository and Open linked data:* The final portion is a recommender system residing upon a cultural heritage repository containing data from V&A museum and the National Gallery London, and a query system for the open linked data on web.

3 Possible Contributions

This work strives to contribute towards

- Building fairly complex contexts using strong semantics supported on known ontologies (rather than keywords), utilizing social web as a context source. This generalized context model will help systems keep track of user's browsing/searching contexts and hence aid recommendations.
- Provide a mechanism of avoiding the cold-start problem, which is a very common and major problem in most of the search and recommender systems.

References

1. Suchanek, F. M., G. Kasneci & G.Weikum.: YAGO: A core of semantic knowledge. In Proc. of WWW (2007).
2. Addis, M. J., Hafeez, S., Prideaux, D., Lowe, R., Lewis, P. H., Martinez, K. and Sinclair, P. A.S.: The eCHASE System for Cross-border Use of European Multimedia Cultural Heritage Content in Education and Publishing. In: AXMEDIS 2006, ICSRiM –(Leeds, UK, 12-15 December 2006)
3. Nick Crofts, Martin Doerr, Tony Gill, Stephen Stead, Matthew.: Definition of the CIDOC Conceptual Reference Model, (October 2006).
4. Auer, S.; Bizer, C.; Lehmann, J.; Kobilarov, G.; Cyganiak, R.; Ives, Z.: DBpedia: A Nucleus for a Web of Open Data.6th International Semantic Web Conference (2007).

---

[1] http://code.google.com/apis/socialgraph/docs/otherme.html