

# User Profiling for Semantic Browsing in Medical Digital Libraries

Patty Kostkova, Gayo Diallo, and Gawesh Jawaheer

City ealth Research Centre, City University, Northampton Square, London,  
EC1V 0HB, UK

patty@soi.city.ac.uk,  
{Gayo.Diallo.1,Gawesh.Jawaheer.1}@city.ac.uk

**Abstract.** Semantic Browsing provides contextualized dynamically generated Web content customizing the knowledge to better meet user expectations. The real-world medical digital library, the National electronic Library of Infection (NeLI, [www.neli.org.uk](http://www.neli.org.uk)), enriched with an infection domain ontology enables new semantic services to be developed qualitatively. In this paper, we will address the use of group profiling to customize semantic browsing by integrating distributed knowledge sources. The service is evaluated by web server logs analysis, dynamically enhancing the profiles and by qualitative feedback from real users of the NeLI portal.

## 1 Introduction and Background

The Semantic Web leverages the knowledge integration on the Web to new levels. Despite the efforts put into the technical and research issues, there are few applications actually deploying and evaluating semantic web with real users. Semantic web can only deliver if it is driven by user needs, context or profiles to seamlessly integrate the knowledge on the web to really provide desirable content.

This is in particular relevant in the medical domain. The Internet enabled patients and healthcare professionals to access vast amount of available information but this often results in the inability to find what is needed and when it is needed [1]. Medical sites need to support a profile-based semantic search and a contextualized browsing to integrate knowledge from other medical portals needed by particular medical users or patients.

Context and customisation are some of the key factors for accurate, effective relevant information access in Internet digital libraries and in general – in the Semantic Web. Allan et al. [2] define contextual retrieval as a general framework combining search technologies and knowledge about a query and so called “user context” into a single framework in order to provide the most appropriate results for users’ information needs. The context of the user may include his/her level of expertise and domains of interest. An user profile is a record of user specific data that define users interests, his/her level of expertise, and his/her context.

## 2 Semantics Profiling in NeLI

### 2.1 Overview

The work presented in this paper demonstrates using profiles to customise user access in the UK based National electronic Library in Infection (NeLI) and to semantically integrate NeLI with other medical portals. Since 2000, NeLI has provided a single access point to the best available evidence around all aspects of infection and is currently being used by more than 15 000 real-world medical professionals a month. Portal users range from members of the public, General Practitioners (GPs), nurses, consultants communicable disease control (CCDCs) to senior Primary Care Trust executives. NeLI has been gathering user and group profiles providing contextual information essential for support of customised services. The development of the new version of NeLI portal takes into consideration the emerging Semantic Web where ontologies are one of the essential components [3]. This is particularly the case in the (bio)medical domain where substantial efforts have been made to develop standards, medical terminologies and coding systems (SNOMED, MeSH and UMLS, which integrates more than 100 most relevant vocabulary sources in medicine<sup>1</sup>), thereby, providing knowledge bases for encoding medical evidence. As there is no standard ontology meeting our specific need, the NeLI ontology has been developed representing the infection medical domain with several hundred concepts.

Resources in the library are indexed using a NeLI ontology that has been created from a pruned sub-tree of the MeSH<sup>2</sup> (Medical Subject Headings) vocabulary and a customised classification for infection control and public health. The development of the NeLI ontology has been carried out by 3 NeLI content managers, in a close collaboration with the project Advisory Board members and infection experts in the UK.

Profiling for recommending research papers based on ontologies was investigated by Middleton [4]. Dynamic profiling applied to information retrieval on the web has been widely applied [5]. However, a very common application of profiles is the vision of customised content relevant to the particular profile [6].

In this section, we will look into the spectrum of NeLI users and discuss how they could be categorised in terms of profiles. NeLI users come from different professional backgrounds and specialties which determine their medical interests, information needs and type of questions they are asking on the portal. In addition to their professions, they have particular specialities and are likely to have particular interests in treatment or investigation of a particular disease or a group of diseases. Due to the size limitation of this demo paper, personal user profiles are outside the remit of this publication and can be found in [7]. Based on personal profiles, which have been inserted manually by NeLI users, we have an understanding of their professional backgrounds and personal interests. Users have provided this information to us when subscribing to project updates at conferences or online. We use the personal profiles as a base for semi-automated development of professional group profiles.

---

<sup>1</sup> <http://umlsks.nlm.nih.gov/>

<sup>2</sup> [www.nlm.nih.gov/mesh/2002/index.html](http://www.nlm.nih.gov/mesh/2002/index.html)

## 2.2 Professional Group Profiles

The professional group profile is defined as a tuple:  $N=\{p,s,e,o,t\}$  where  $p \in P$ ,  $s \in P$ ,  $e \in O$ ,  $o \in O$  and  $t \in T$ , where  $O$  is the set of NeLI ontology concepts,  $P$  is the list of NeLI recognised professions and  $T$  is a list of NeLI-recommended external knowledge web portals in a form of a URL (called Targets). The list  $P$  of NeLI recognised professions include: clinical scientists, nurse, consultant, environmental health officer, general practitioner, lecturer, microbiologist. An example of a user is illustrated in Table 1.

**Table 1.** Example of a speciality group user profile

Profession	Nurse
Speciality	Infection Control
Expertise Area	TB
Other Area of Interest	Hand washing Antibiotic Resistance National Policy
Targets	Department of Health WHO

Currently, we are looking into a dynamic web server log-based generation of the profiles (the Other Area of Interest and the Targets fields being dynamically populated by frequently searched keywords from the NeLI ontology). The aim is to dynamically enhance the profiles by evaluation of the search keywords and navigation terms from the web server logs. Users in a group may have different interests, but also sufficiently similar interests for the development of group default preferences. The need for a combination of professional, speciality and topic comes from different questions asked by different users. For example, all users search for TB (tuberculosis) but the information they actually need varies: a clinician will ask about latest TB guidelines provided by the Department of Health (target), an infection control nurse about isolating patients with TB, a public health office is more concerned about high-risk TB populations and outreach to them, while a GP might need to check the latest diagnosis and treatment recommendations for TB. A case scenario around TB is shown in Table 2. Based on the user group (profession), different information is to be provided (targets – portals with further knowledge).

**Table 2.** User profiles information needs example based on a TB scenario

Professional group	Targets
GPs	BNF (British National Formulary) NeLI treatment pages
clinicians	PubMed Clinical Evidence
nurse	Department of Health
public	Wikipedia HPA Public leaflets Nathnac

### 3 Semantic Browsing and Semantic Knowledge Integration

Semantic browsing provides users with dynamically selected concepts or links from an ontology – enriched by the profile-based customization this selects and integrates web portals by working as a “semantic recommender” system. NeLI semantic browsing is being developed in the context of the SeaLife project [8] enabling users to semantically browse the Web by highlighting ontology concepts and providing dynamic access to Web servers or knowledge portals semantically related to the Web content retrieved (targets).

To facilitate this, NeLI has integrated the Conceptual Open Hypermedia Service (COHSE) system [9] developed at University of Manchester. COHSE automatically inserts hyperlinks on web pages by recognizing terms contained in background knowledge, based on an uploaded ontology, and presents the user with search web services linking to relevant targets, see Table 2. For any term from the ontology, resources are provided for broader, narrower and related terms based on ontological taxonomical and non-taxonomical relationships. Selecting the relevant set of targets for a particular profile improves the contextualization of the search and recommender function.

#### Demonstrator

Unlike many other demos serving just as a proof of concept, NeLI is a real-world Internet medical library with over 15 000 unique users a month. In the demonstration, we will show the ontology-based selection of targets based on the user group profiles. A medical scenario semantically integrating the following targets based on ontology-based relationships will be shown for two profiles: public and nurses.

A “public” user visits NeLI to search for tropical diseases. A list of search results based on taxonomical relationships in the NeLI ontology is displayed containing (among others) malaria resources. NeLI ontology terms on the page are highlighted by the NeLI-COHSE system and ontological relationships to malaria are shown in a pop-up box. Among non-taxonomical relationship “*is caused*” is shown giving *Plasmodium*. When selected, the targets for public profile show the Natnac web site <http://www.nathnac.org> giving travelers information on malaria for public.

A “nurse” searches for healthcare associated infection. Results contain documents on MRSA, TB etc., provided by the NeLI ontology. When a TB concept is highlighted a pop-up box providing the ontology relationships shows a non-taxonomical “transmission mode” giving *airborne*. The HPA ([www.hpa.org.uk](http://www.hpa.org.uk)) and Department of Health ([www.dh.gov.uk](http://www.dh.gov.uk)) targets, provided for a nurse profile, are searched to give policy resources on management of healthcare acquired TB – in particular, respiratory isolation and respiratory protection.

These two case scenarios demonstrate the semantic search, use of taxonomical and non-taxonomical relationships in the NeLI ontology and the new concept of selection of targets for different professional profiles and the construction of searches on those external sites to provide contextualized information. These are new features enhancing the COHSE system developed by the NeLI team to support profile-based customization. A screen shot can be found in Figure 1.

The visitors will see a unique working example of domain ontology-driven medical portal in use by practising clinicians and a semantically integrated knowledge from other portals based on professional group profiles.

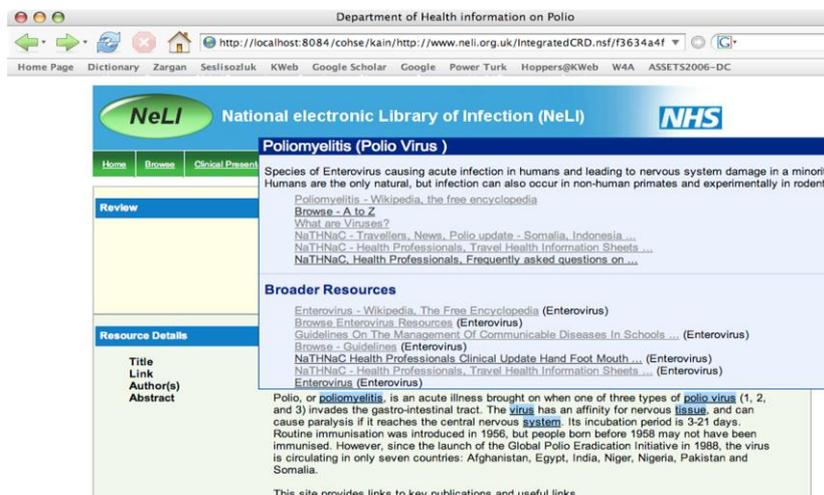


Fig. 1. NeLI-COHSSE Semantic Browsing and Recommender Case Scenario

**Acknowledgments.** We acknowledge the EC for providing funding for the SeaLife project (IST-2006-027269). We thank the University of Manchester for providing the COHSE System used in this research.

## References

1. Muir Gray, J.A., de Lustignan, S.: National electronic Library for Health (NeLI). *BMJ* 319, 1476–1479 (1999)
2. Allan, J., et al.: Challenges in Information Retrieval and Language Modelling: Report of a workshop held at the centre for intelligent information retrieval, University of Massachusetts Amherst, September 2002 (2003). *ACM SIGIR Forum* 37(1), 31–47
3. Berners-Lee, T., Hendler, Lassila, O.: The Semantic Web. *Scientific Am.*, 34–43 (May 2001)
4. Middleton, S.E., Shadbolt, N.R., De Roure, D.C.: Ontological User Profiling in Recommender Systems. *ACM Trans on Information Systems (TOIS)* 22(1), 54–88 (2004)
5. Danilowicz, C., Indyka-Piasecka, A.: Dynamic User Profiles Based on Boolean Formulas. In: Orchard, B., Yang, C., Ali, M. (eds.) *IEA/AIE 2004. LNCS (LNAI)*, vol. 3029, pp. 779–787. Springer, Heidelberg (2004)
6. Wærn, A.: User Involvement in Automatic Filtering: An Experimental Study. *User Modeling and User-Adapted Interaction* 14(2-3) (June 2004)
7. Kostkova, P., Diallo, G., Jawaheer, G.: Application of User Profiling on Ontology Module Extraction for Medical portals. In: *The Proceedings of the MedSemWed 2007 Workshop*, held in conjunction with the MedInfo 2007 Conference, Brisbane, Australia (August 2007)
8. Schroeder, M., Burger, A., Kostkova, P., Stevens, R., Habermann, B., Dieng-Kuntz, R.: From a Service-based eScience Infrastructure to a Semantic Web for the Life Sciences: The Sealife Project. In: *Workshop on Network Tools and Applications in Biology, NETTAB 2006*, Italy (2006)
9. Bechhofer, S., Yesilada, Y., Horan, B.: COHSE: Knowledge-Driven Hyperlinks. In: *The Semantic Web Challenge at the International Semantic Web Conference (ISWC 2006)* (2006)