

The Epidemic Marketplace Platform: towards semantic characterization of epidemiological resources using biomedical ontologies

Francisco M. Couto^{1*}, João D. Ferreira¹, João Zamite¹, Carlos Santos¹, Tiago Posse¹, Paulo Graça¹, Dulce Domingos¹ and Mário J. Silva²

¹Lasige, Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa, Portugal

²IST, INESC-ID, Technical University of Lisbon, Portugal

ABSTRACT

This software demonstration paper presents the Epidemic Marketplace, a platform for integrating and sharing epidemiological data. It uses its own semantic metadata model for characterizing the resources submitted, which integrates a network of epidemiology related ontologies. The Epidemic Marketplace platform is available at <http://www.epimarketplace.net/>.

1 INTRODUCTION

The Epidemic Marketplace is a platform for integrating and sharing epidemiological data, which aims at supporting transparent, seamless access to distributed, heterogeneous and redundant resources (Silva *et al.*, 2010). The development of this platform is one of the goals of EPIWORK, a multidisciplinary European research project funded by a Future and Emerging Technologies (FET) Proactive initiative, aimed at developing the appropriate framework of tools and knowledge needed for the design of epidemic forecast infrastructures.

The Epidemic Marketplace uses its own semantic metadata model for characterizing the submitted resources (Lopes *et al.*, 2010). To improve the system interoperability, the semantic metadata model is internally mapped to the Dublin Core Metadata Element Set, a standard way to describe web-based information resources. However, the integration of available ontologies in the semantic metadata model will have a crucial role on enhancing data integration and the consistency and coherency of their semantic characterization. Thus, we proposed a Network of Epidemiology Related Ontologies, dubbed NERO, that establishes a collection of ontologies to be integrated in the Epidemic Marketplace's metadata model (Ferreira *et al.*, 2012). Ontology matching techniques will be used to align shared concepts between ontologies (Cruz *et al.*, 2011). The network will provide a source of concepts to characterize epidemiological resources.

The usage of ontologies enables the restriction of the semantic characterization of every resource to standard and unambiguous terms, making the information retrieval and analysis of the stored resources more effective. Additionally, it also facilitates the semantic characterization process, for example by providing term suggestions according to the user profile and their previous resource submissions.

2 USER INTERFACE

The platform provides a web user interface to upload, characterize, browse, search, download, request and comment the epidemic resources and their metadata. The resources are disseminated under access policies that can be personalized.

The interface was developed using Drupal, a highly customizable free software package for the organization, management and distribution of any content¹. The resources are stored in a repository based on Fedora Commons, a general-purpose, open-source digital object repository system².

Figure 1 shows an example of the resources returned by the platform when searching for flu in 2010 and by ordering them by date (oldest first). This is just an example of a generic search, but more specialized filters based on ontological terms may be performed.

The resources and their metadata are also directly accessible through the set of RESTful web services that are invoked by the user interface to store and manage the resources. These web services are currently being used by computational applications that require automatic retrieval, search, upload, update or removal of resources stored in the Marketplace platform. For example, these web services are currently being used to store simulation data from the GleamViz simulator³, tweets retrieved by the MEDCollector⁴, and monitorization data gathered by the Influenzanet system⁵. The web services also enable computational applications to manage group-based access control policies. The detailed specification of these web services is presented on the platform's web site.⁶

3 CONCLUSION

This software demonstration paper presents the Epidemic Marketplace, a platform that enables users to semantically characterize their resources in a standard way by exploiting the knowledge encoded in a comprehensive network of epidemiologically relevant ontologies. This model not only enhances the characterization process, but also provides the community with an important asset by its own, and improves the integration, retrieval and analysis of resources.

¹ <http://www.drupal.org/>

² <http://www.fedora-commons.org/>

³ <http://www.gleamviz.org/simulator/>

⁴ <http://vimeo.com/14734417>

⁵ <http://www.influenzanet.eu/>

⁶ http://www.epimarketplace.net/developers_corner/web_services

*To whom correspondence should be addressed: fcouto@di.fc.ul.pt

The screenshot displays the Epidemic Marketplace interface. At the top, there is a navigation bar with the logo 'EM Epidemic Marketplace' and a search bar. Below the navigation bar, there are filter options for Title, Uploader, Subject, Type, and Date. The Date filter is set to '2010 (42)'. There are also 'Order by' options for Title, Uploader, Subject, Type, and Date, with 'Oldest First' selected. A red banner indicates the search results: 'Resources Filtered by: Date in 2010, All matching "flu", Ordered by: Date (Oldest First)'. Below the banner, it says 'Showing 12 resources, from a total of 42 results'. The resources are displayed in a grid of 12 cards, each with a title, author, date, description, subject, and type. Each card also has a 'See more' link. The resources include: 'Flu Survey', 'Influenzanet', 'Influweb', 'UN World population prospect (2008)', 'Google Flu Trends', 'Twitter Dataset H1N1 + Portugal 4-6-2009', 'Twitter Dataset H1N1 + Italy 29-5-2009', 'Twitter Dataset H1N1 + holland 4-6-2009', 'Twitter Dataset H1N1 + Spain 29-5-2009', 'Data collector - H1N1 - all locations', 'Arrival Times Global', and 'cumWHO worldwide noMexico'. At the bottom of the grid, there are navigation links: '1 2 3 4 next last »'. At the very bottom of the page, there are logos for 'EPIWORK' and the European Union.

Fig. 1. Epidemic Marketplace screenshot from the current user interface.

The goal is to provide a unified and integrated platform for the management of epidemic resources, which will be semantically characterized by extensively exploiting the knowledge encoded in available ontologies. This platform will eventually serve as the basis for a distributed information reference for epidemic modelers, which will further aid the communication within the community of epidemiologists and the sharing of their knowledge.

ACKNOWLEDGEMENTS

The authors want to thank the European Commission for the financial support of the EPIWORK project under the Seventh Framework Programme (Grant #231807) and the Portuguese Fundação para a Ciência e Tecnologia through the financial support of the SOMER project (PTDC/EIA-EIA/119119/2010), the PhD grant SFRH/BD/69345/2010, through the PIDDAC Program funds

(INESC-ID multi annual funding) and through the LASIGE multi-annual support.

REFERENCES

- Cruz, I., Stroe, C., Pesquita, C., Couto, F., and Cross, V. (2011). Biomedical ontology matching using the AgreementMaker system. In *Software Demonstration at the International Conference on Biomedical Ontologies (ICBO)*.
- Ferreira, J., Pesquita, C., Silva, M., and Couto, F. (2012). Bringing epidemiology into the semantic web. In *International Conference on Biomedical Ontologies (ICBO)*.
- Lopes, L., Silva, F., Couto, F., Zamite, J., Ferreira, H., Sousa, C., and Silva, M. (2010). Epidemic marketplace: An information management system for epidemiological data. In *1st International Conference on Information Technology in Bio and Medical Informatics (ITABM - DEXA)*.
- Silva, M., Silva, F., Lopes, L., and Couto, F. (2010). Building a digital library for epidemic modelling. In *ICDL 2010 - The International Conference on Digital Libraries*.