

Chapter VIII

SWARMS:

A Platform for Domain Knowledge Management and Applications

Jie Tang

Tsinghua University, Beijing, China

Bangyong Liang

NEC Labs, China

Juanzi Li

Tsinghua University, Beijing, China

ABSTRACT

This chapter describes the architecture and the main features of SWARMS, a platform for domain knowledge management. The platform aims at providing services for 1) efficiently storing and accessing the ontological information; 2) visualizing the networking structure in the ontological data; 3) searching and mining the semantic data. One advantage of the system is that it provides a suite of components for not only supporting efficient semantic data storage but also searching and mining the semantics. Another advantage is that the system supports visualization in the process of search and mining, which would greatly help a normal user to understand the knowledge inside the ontological data. SWARMS can be easily customized to adapt to different domains. The system has been applied to several domains, such as News, Software, and Social Network. In this chapter, we will also present the performance evaluations of the system.

INTRODUCTION

With the rapid growing interest in Semantic Web, more and more ‘semantic’ data becomes available.

Semantic search can assist users and applications to fetch relevant knowledge in the domain (Berendt *et al.*, 2003). Semantic mining is used to reveal the hidden facts in order to empower the

'semantics' in ontological data. For example, to answer the question "get all persons involved in the project of web spider", data semantic search may be sufficient to give the answer by using SPARQL (Prud'hommeaux and Seaborne, 2007). However, for the question "get all the persons who are experts on data mining in the project of web spider", the simple search based method cannot be sufficient. Instead, mining might be a better solution to answer this question. Unlike mining on the unstructured Web, the metadata on the Semantic Web might be complicated, the traditional search and mining methods cannot be directly adapted to this scenario. In addition, as the semantic data is often represented by a complex metadata and how to make it easy for a normal user to understand is also a challenging issue. Consequently, a comprehensive investigation including semantic data storage, indexing, search, and mining is required. Furthermore, a practical platform for supporting the management is also necessary. This is exactly the problem addressed in this chapter.

There have been many researching and industrial efforts on semantic data management in recent years. However, most of the works focus on specific problems. For example, Swoogle (Ding and Finin, 2005) focuses on semantic search (searching for ontologies that contain the input query terms); Flink (Mika, 2005) provides a graphical view for the researcher social network. Different from the existing works, we aim at integrating the semantic search and mining in a general platform in order to provide a comprehensive tool for semantic data management; we also try to visualize the processes of search and mining so as to facilitate the normal user to easily browse and understand the complicated ontological data. For several features in the system, e.g., semantic data caching, expert finding, and association search, we propose new approaches trying to overcome the drawbacks that exist in the conventional methods. For some other features, e.g., storage, searching, and visualization, we utilize the state-

of-the-art methods. This is because, these issues have been intensively investigated previously and the conventional methods can result in good performances.

SWARMS (Semantic Web Aided Rich Mining System) is a general platform for semantic content management in Semantic Web. SWARMS provides various functions for domain semantic data management. SWARMS can be easily adapted to different domains by importing the ontological data of that domain (Liang, Tang, Li, and Wang, 2005b; Liang, Tang, Li, and Wang, 2006b)

The rest of the paper is organized as follows: Section 2 gives the whole architecture of SWARMS. Section 3 describes the knowledge access layer. Section 4 describes the search functions provided by SWARMS. Section 5 presents the association search algorithm in SWARMS. In Section 6, we introduce an application of the SWARMS platform to the researcher community domain. Finally, before concluding the chapter we give related work and future research directions.

ARCHITECTURE OF SWARMS

We start by giving a general overview of SWARMS. SWARMS includes (1) an indexing module that creates indexes for both semantic data and text-based data, (2) a knowledge access layer which provides a common API for accessing the knowledge data, which also includes a cache module to fast the search process, (2) a knowledge search engine that can search both semantic data and the text-based data, (4) an association search module that searches for the connections between instances, (5) a mining module that includes several basic mining algorithms including propagation and clustering, and (6) a visualization module that supports visualizing the process of search, association search, and mining.

Figure 1 shows the SWARMS architecture. Domain knowledge base stores the semantic data

29 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/chapter/swarms-platform-domain-knowledge-management/30390?camid=4v1

This title is available in InfoSci-Books, InfoSci-Multimedia Technologies, InfoSci-Knowledge Management, Semantic Web and Web Services, Business-Technology-Solution, InfoSci-Computer Science and Information Technology, Science, Engineering, and Information Technology, InfoSci-Select, InfoSci-Select. Recommend this product to your librarian:

www.igi-global.com/e-resources/library-recommendation/?id=1

Related Content

Ontologies with Semantic Web/Grid in Data Integration for OLAP

Tapio Niemi, Santtu Toivonen, Marko Niinimaki and Jyrki Nummenmaa (2007). *International Journal on Semantic Web and Information Systems* (pp. 25-49).

www.igi-global.com/article/ontologies-semantic-web-grid-data/2841?camid=4v1a

Finding Healthcare Issues with Search Engine Queries and Social Network Data

M. Ikram Ullah Lali, Raza UI Mustafa, Kashif Saleem, M. Saqib Nawaz, Tehseen Zia and Basit Shahzad (2017). *International Journal on Semantic Web and Information Systems* (pp. 48-62).

www.igi-global.com/article/finding-healthcare-issues-with-search-engine-queries-and-social-network-data/172422?camid=4v1a

Entity-Centric Semantic Interoperability

Paolo Bouquet, Heiko Stoermer, Wojcech Barczynski and Stefano Bocconi (2010). *Cases on Semantic Interoperability for Information Systems Integration: Practices and Applications* (pp. 1-21).

www.igi-global.com/chapter/entity-centric-semantic-interoperability/38036?camid=4v1a

Extraction and Prediction of Biomedical Database Identifier Using Neural Networks towards Data Network Construction

Hendrik Mehlhorn, Matthias Lange, Uwe Scholz and Falk Schreiber (2013). *Teaching Cases Collection* (pp. 58-83).

www.igi-global.com/chapter/extraction-prediction-biomedical-database-identifier/77200?camid=4v1a