

Chapter 21

OSIRIS: Ontology-Based System for Semantic Information Retrieval and Indexation Dedicated to Community and Open Web Spaces

Francky Trichet

University of Nantes: Team Knowledge and Decision (KOD), France

Xavier Aimé

University of Nantes: Team Knowledge and Decision (KOD), France

Christophe Thovex

University of Nantes: Team Knowledge and Decision (KOD), France

ABSTRACT

OSIRIS (Ontology-based Systems for Semantic Information Retrieval and Indexation dedicated to community and open web Spaces) is a platform dedicated to the development of community web spaces which aim at facilitating both semantic annotating process and searching process of multimedia resources. Based on the use of both heavyweight ontologies and thesauri, OSIRIS allows the end-user (1) to describe the semantic content of its resources by using an intuitive natural-language based model of annotation which is founded on the triple (Subject, Verb, Object), and (2) to formally represent these annotations by using Conceptual Graphs. Each resource can be described by adopting multiple points of view, which are usually provided by different end-users. These different points of view can be defined by using multiple ontologies which can be related to connected (or not-connected) domains. Developed from the integration of Semantic Web technologies and Web 2.0 technologies, OSIRIS aims at facilitating the deployment of semantic, collaborative, community and open web spaces. The use of OSIRIS is illustrated in the context of a project dedicated to the preservation of French popular and cultural heritage.

DOI: 10.4018/978-1-61520-883-8.ch021

INTRODUCTION

Currently, the collective and interactive dimension of Web 2.0 coupled with the lightness of its tools facilitates the rise of many platforms dedicated to the sharing of multimedia resources such as Flickr (<http://www.flickr.com/>) for the images or YouTube (<http://www.youtube.com>) for the videos. However, the success of these platforms (in terms of number of listed resources and number of federated users) must be moderated in comparison with the poverty of the approach used for Information Retrieval (IR). Indeed, the search engines integrated in such systems are only based on the use of tags which are usually defined manually by the end-users of the communities (*i.e.* the social tagging which leads to the creation of folksonomies). In addition to the traditional limits of IR systems based on keywords, in particular the poverty of semantic description provided by a set of tags and consequently the impossibility of implementing a semantic search engine, these systems suffer from a lack of openness because the tags provided by the end-users remain useful and efficient only inside the platforms; they cannot be exported when the resources are duplicated from a platform to another.

OSIRIS (*Ontology-based Systems for Semantic Information Retrieval and Indexation dedicated to community and open web Spaces*) is a platform dedicated to the development of community web spaces which aim at facilitating both semantic annotating process and searching process of multimedia resources. Such a community space corresponds to an Internet-mediated social and semantic environment in the sense that the resources which are shared are not only tagged by the users (which thus construct a folksonomy in a collaborative way) but they are also formally described by using one (or several) ontolog(ies) shared by all the members of the community. The result is an immediate and rewarding gain in the user's capacity to semantically describe and find related content.

Based on the use of heavyweight ontologies (Furst & Trichet, 2006a) coupled with thesauri¹, OSIRIS allows the end-users to semantically describe the content of a resource (for instance, this photography of Doisneau represents “*A woman who kisses a man in a famous French place located in Paris*”) and then to formally represent this content by using Conceptual Graphs (Sowa, 1984). Each resource can be described according to multiple points of view (*i.e.* representation of several contents) which can also be defined according to multiple ontologies, which can cover connected domain or not. Thus, during the annotating process, OSIRIS allows managing several ontologies which are used jointly and in a transparent way during the searching process, thanks to the possibility of defining equivalence links between concepts and/or relations of two ontologies. Moreover, OSIRIS is based on heavyweight ontologies, (*i.e.* ontologies which in addition to including the concepts and relations [structured within hierarchies based on the relation of Specialisation/Generalisation] characterizing the considered domain, also include the axioms [rules and constraints] that govern this domain). This confers to OSIRIS the possibility to automatically enrich the annotations (manually associated to a resource) by applying the axioms which generally correspond to inferential knowledge of the domain.

From a technical point of view, OSIRIS is based on the integration of technologies currently developed in Web 2.0 and Semantic Web areas. In its current version, OSIRIS allows implementing Semantic Web Spaces dedicated to the sharing of images, videos, music files and office documents, respectively in JPEG, MP3, OpenOffice and Office 2007 format. The choice of these formats is justified by the fact that it is possible to store the semantic annotations (represented in terms of conceptual graphs) within the files via the use of standards such as IPTC (<http://www.iptc.org>) for JPEG, ID3 (<http://www.id3.org>) for MP3 and ODF (<http://www.oasis-open.org/>) for OpenOffice and Office 2007. These standards make it possible first

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/osiris-ontology-based-system-semantic/45055

Related Content

Reply Speed as Nonverbal Cue in Text Messaging with a Read Receipt Display Function: Effects of Messaging Dependency on Times until Negative Emotions Occur While Waiting for a Reply

Shogo Kato, Yuuki Kato and Yasuyuki Ozawa (2020). *International Journal of Technology and Human Interaction* (pp. 36-53).

www.irma-international.org/article/reply-speed-as-nonverbal-cue-in-text-messaging-with-a-read-receipt-display-function/239530

Security Benefits of Little Data From the Socio-Technical Perspective

Peter Imrie and Peter M. Bednar (2018). *International Journal of Systems and Society* (pp. 45-53).

www.irma-international.org/article/security-benefits-of-little-data-from-the-socio-technical-perspective/210592

Parochial School Teachers Instructional Use of the Interactive Whiteboard

Jillian R. Powers (2018). *Handbook of Research on Human Development in the Digital Age* (pp. 109-134).

www.irma-international.org/chapter/parochial-school-teachers-instructional-use-of-the-interactive-whiteboard/186213

Electronic Aggression among Adolescents: An Old House with a New Facade (or Even a Number of Houses)

Jacek Pyzalski (2011). *Youth Culture and Net Culture: Online Social Practices* (pp. 278-295).

www.irma-international.org/chapter/electronic-aggression-among-adolescents/50705

Machine Learning in Fake News Detection and Social Innovation: Navigating Truth in the Digital Age

Vitthal B. Kamble, Nilesh J. Uke, Dhiraj Ghanshyam Karwatkar, Rahul Dilip Dhongade and Priyanka Kasare (2025). *Exploring Psychology, Social Innovation and Advanced Applications of Machine Learning* (pp. 87-108).

www.irma-international.org/chapter/machine-learning-in-fake-news-detection-and-social-innovation/371660