

GoNTogle: A Tool for Semantic Annotation and Search

Giorgos Giannopoulos^{1,2}, Nikos Bikakis^{1,2}, Theodore Dalamagas², and Timos Sellis^{1,2}

¹ KDBSL Lab, School of ECE, Nat. Tech. Univ. of Athens, Greece
giann@dblab.ntua.gr, bikakis@dblab.ntua.gr

² Institute for the Management of Information Systems, "Athena" Research Center, Greece
dalamag@imis.athena-innovation.gr,
timos@imis.athena-innovation.gr

Abstract. This paper presents GoNTogle, a tool which provides advanced document annotation and search facilities. GoNTogle allows users to annotate several document formats, using ontology concepts. It also produces automatic annotation suggestions based on textual similarity and previous document annotations. Finally, GoNTogle combines keyword and semantic-based search, offering advanced ontology query facilities.

Keywords: Document Annotation, Ontology, Semantic Search.

1 Introduction

Semantic annotation and search tools are at the core of Semantic Web Technology. Annotations involve tagging of data with concepts (i.e., ontology classes) so that data becomes meaningful. Annotating data can help in providing better search facilities, since it helps users to search for information not only based on the traditional keyword-based search, but also using well-defined general concepts that describe the domain of their information need.

A great number of approaches on semantic annotation have been proposed in the literature [2]. Most of them are focused on annotating web resources such as html pages or plain text [3,4,5,6]. As far as popular document formats are concerned, there are approaches that differ in the annotation and search facilities they offer [7,8].

In this paper we present GoNTogle. GoNTogle supports manual and automatic annotation of several types of documents (doc, pdf, rtf, txt, odt, sxw) using ontology classes, in a fully collaborative environment. It also provides searching facilities beyond the traditional keyword-based search, using a flexible combination of keyword and semantic-based search. In contrast with other works, our aim was to implement an easy-to-use document annotation and search tool, that would fully support (a) viewing and annotating popular document types while maintaining their initial format, (b) sharing those annotations and (c) searching for documents combining keyword and semantic-based search.

The key features of our tool are the following:

- It allows users to open and view widely used document formats such as *.doc* and *.pdf*, maintaining their original format.