Ontology-Based Legal Information Retrieval to Improve the Information Access in e-Government

Asunción Gómez-Pérez
Facultad de Informática. Universidad Politécnica de Madrid
Campus Montegancedo, s/n 28860
Boadilla del Monte. Madrid. Spain
34-913367439
asun@fi.upm.es

Fernando Ortiz-Rodríguez
Facultad de Informática. Universidad Politécnica de Madrid
Campus Montegancedo, s/n 28860
Boadilla del Monte. Madrid. Spain
34-913363670
fortiz@fi.upm.es

Boris Villazón-Terrazas
Facultad de Informática. Universidad Politécnica de Madrid
Campus Montegancedo, s/n 28860
Boadilla del Monte. Madrid. Spain
34-913363670
bvillazon@fi.upm.es

ABSTRACT
In this paper, we present EgoIR, an approach for retrieving legal information based on ontologies; this approach has been developed with Legal Ontologies to be deployed within the e-government context.

Categories and Subject Descriptors
H.3.3 Information Storage and Retrieval: Information Search and Retrieval – query formulation, retrieval models, search process.

General Terms
Design, Experimentation

Keywords
Ontology, Information Retrieval

1. INTRODUCTION
For more than two decades, the AI and Law community has been very active and productive. In the early 80’s, research was focused on logic programming. Other approach adopted was the case-based reasoning. Knowledge Engineering was also of interest for the research community and the field most applied since it allowed developing and using the legal ontologies that underlie the growth of the Semantic Web.

The e-Gov has been strengthened with all these previous studies carried out by the research community and now its main concern is data representation and information management. By its nature, the e-Gov is supported by the legal domain.

Our contribution consists of an ontology based approach for legal information retrieval that we called EgoIR. This system has as a main goal to retrieve e-Gov documentation. EgoIR deals with Real-estate transaction documents, and gives an opportunity to the citizens, business and governments to integrate and recover documents. For this purpose EgoIR provides facilities for managing, searching and sharing e-Gov documentation.

2. EgoIR
EgoIR is an Ontology-Based Legal Information Retrieval System. This system is the result of integrating Ontological Workbench WebODE\(^1\), and a text search engine library, Lucene\(^2\). In this section we describe the system architecture and the Legal Ontologies.

2.1 Architecture
The system integration of the EgoIR is built and composed by the Search Client, the Search Server and the Ontology Server modules, which are described in the next subsections. Figure 1 shows the general architecture of the system.

![EgoIR System Architecture](image)

**Figure 1. EgoIR System Architecture.**

2.1.1 Ontology Server
This module defines how the knowledge is structured in the application domain. This module includes the Legal Ontologies within WebODE\(^2\).

Within the Legal Ontologies, concept instances are associated with documents. Every time that a new concept instance is added the Ontology Server communicates with the Search Server to index its corresponding document.

2.1.2 Search Client
This module incorporates two sub-modules: a Query Builder and a Document Viewer.

Query Builder connects to the Ontology Server, in order to access Legal Ontologies, browse them and obtain concepts to build the query by using a graphical interface. This module sends the query to the Search Server.

---

\(^1\) http://webode.dia.fi.upm.es/
\(^2\) http://lucene.apache.org/
Another work reported in [4] is the EULEGIS (European User Views to Legislative Information in Structured Form), whose main goal is to provide a consistent user interface for legal IR generated in different legal systems and at different legislative levels. This system focuses on user interfaces.

4. CONCLUSIONS AND FUTURE WORK
In this paper we present our first approach to an ontology-based legal IR, which aims to retrieve government documents in a timely and accurate way. This is an approach of an entirely new wave of legal knowledge systems. At this time we can mention that the utility of ontologies within an IR is twofold: On the one hand, as a social impact, ontologies are a good way to guide user to the legal terms, thus avoiding him/her to make mistakes at the query construction; and on the other hand, mostly technical, ontologies are a key to the development the Semantic Web and improving interoperability on the legal applications.

Finally, in the near future we will improve the performance of EgoIR and we will focus on further enhancement of the ontology-based retrieval mechanism by means of Natural Language Processing (NLP) techniques for a user friendlier environment; on the automatic semantic annotation of the documents to improve the search process; and on security issues by providing a summary of the retrieved documents.

5. ACKNOWLEDGMENTS
We would like to thank our project partners for their helpful comments. Specially thanks to Rosario Plaza and Ángel López-Cima. This work is carried out within the ongoing Reimdoc Project which is supported by the Spanish Technology and Science Ministry (Project FIT-340100-2004-22).

This project is partially funded by a scholarship granted through the Program of Teaching Staff Improvement (PROMEP) at the Tamaulipas University.

6. REFERENCES