Legal Ontologies for the Spanish e-Government

Asunción Gómez-Pérez, Fernando Ortiz-Rodríguez, Boris Villazón-Terrazas

Facultad de Informática. Universidad Politécnica de Madrid
Campus Montegancedo, s/n 28860 Boadilla del Monte. Madrid. Spain
{asun, fortiz, bvillazon}@fi.upm.es

Abstract. The Electronic Government is a new field of applications for the semantic web where ontologies are becoming an important research technology. The e-Government faces considerable challenges to achieve interoperability given the semantic differences of interpretation, complexity and width of scope. In this paper we show the results obtained in an ongoing project commissioned by the Spanish government that seeks strategies for e-Government to reduce the problems encountered when delivering services to citizens. Here we present an e-Government ontology model; within this model a set of legal ontologies are devoted to represent the Real-estate transaction domain used to illustrate this paper; some examples of use of these legal ontologies are given.

1 Introduction and Motivation

Electronic Government (e-Gov) is an important application field [2] for the transformations that governments and public administrations will have to undergo in the next decades. Therefore to transform the e-Gov into the e-Governance, the e-Gov research needs to be based on a robust theory, on modelling approaches, and on planning. In this scenario, it is crucial to manage the legal knowledge for improving the systems applications in different ways.

For over 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, and all the efforts were centered on legislation and on legal reasoning. Another approach adopted was the case-based reasoning, which was not as formal as logic was, aimed at finding similarities in legal cases and allowed retrieving relevant cases for the judges. Knowledge engineering was also of interest for the research community and the field most applied; this area allowed developing and using the legal ontologies that underlie the growing of the Semantic Web.

The Semantic Web was proposed by Tim Berners-Lee [7] as a new field of research, and according to World Wide Web Consortium[1] (W3C) the Semantic Web is defined as "an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. It is based on the idea of having data on the Web defined and linked such that it can be

[1] http://www.w3.org/2001/sw

used for more effective discovery, automation, integration, and reuse across various applications".

The Semantic Web at e-Gov is new; it features knowledge representation, knowledge engineering, database design, information systems, database integration, natural language understanding, information retrieval and semantic portals, among others. The Semantic Web is considered to be the infrastructure upon which all intelligent e-Gov applications will be built in the near future.

Within the objectives of the Semantic Web the ontologies play an important role. "Ontology" is a word taken from Philosophy where it is used as a systematic explanation of "existence". In the field of the Artificial Intelligence, Neches [11] defined an ontology for the first time in the following way: "Ontology defines the basic terms and the relations that include the vocabulary of a specific area, in addition to the rules to combine terms and relations to define extensions to the vocabulary". It is possible to say that this definition serves us as a kind of guide to construct ontologies. According to Neche´s definition, an ontology does not include only the terms that explicitly are defined in it, but also those that can be inferred using rules. Gruber defines the ontology as: "An explicit specification of a conceptualization" [5, 6].

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. In Spain, legal ontologies for e-Gov applications have been scarce and to reverse this we pursue as the first goal of our paper. The second is to build ontologies that help reduce some important semantic problems presented when providing e-Gov services [3].

This research is based on the needs stated in a Spanish Project that seeks strategies for e-Gov and aims to provide knowledge conceptualizations given by legal experts that help improve information retrieval of legal sources in general.

This paper is organized as follows: section 2 deals with the related work done, section 3 shows the Reimdoc Project; section 4 describes the legal ontologies built and at section 5 their application use. Finally the conclusions are given.

2 Related Work

Nowadays the joint efforts put in by different research communities have made possible the birth of the semantic e-Gov. Since e-Gov ontologies are still in their initial state, only a few works carried out in this field are known; thus, in this section we provide a brief state of those works performed in IA, in the law field and in the Semantic Web. The sum up of all these efforts can produce robust ontologies for the e-Gov domain.

2.1 Law and e-Gov within the Semantic Web

Currently, the Semantic Web is a new area of research and applications within the law and e-Gov domains and is a promise for the next generation web; this new area

will transform the current web, which is now used mainly to communicate with people but not with machines; and this capability of communicating with machines is one of the main objectives of the Semantic Web. If the web were equipped with more meaning, every citizen would extract answers in a new, easy and simple way, and this action could be carried out by web powered semantics, what would enable citizens and businesses to obtain better information from the government. Web powered semantics could help the e-Gov in two ways: first, by allowing the government to delegate more intelligent tasks to computers and second, by solving daily problems with logic deductions and reasoning. But at present, the web is merely a common framework that allows data to be shared and reused.

Currently the legal and e-Gov Semantic Web applications are still in an experimental phase, but their potential impact over social, economical and political issues is extremely significant.

The main goals of e-Gov are to develop user-friendly and efficient services for the public and the business community; semantic interoperability is seen as an important issue to solve within this domain. Some of the works aimed at covering the semantic e-Gov domain are the following: the DIP project[2], the Reimdoc project[3], The IFIP Working Group 8.5[4], the Ontogov project[5], the Egov project[6], and the WEBOCRAT project[7].

2.2 Ontologies: Domain Considerations

The e-Gov scenario is a promising application field for the ontologies underlying the legal engineered knowledge. Many ontologies have been built in the legal domain but not all of them are available or are modelled just for a specific domain. The research efforts made in the legal domain by the AI community have contributed to the making of ontologies such as the following: LLD [9], NORMA [12, 13], FOL [14], FBO [8, 16] and LRI-Core Legal Ontology [1].

The emergence of legal ontologies as part of the Semantic Web initiative has provided a new opportunity for the research community and has brought about a solution to retrieve legal documents within the e-Gov domain. We can mention some of the efforts carried out by IA community on building e-Gov ontologies:

- The Government R&D[8] describes organizations and individuals participating in a government R&D program.
- The Government type[9] describes government concepts used in the CIA World Fact Book 2002.
- The E-Government Ontology[10] describes a seamless UK taxonomy.

---

[2] http://dip.semanticweb.org
[3] http://reimdoc.atosorigin.es
[4] http://falcon.ifs.uni-linz.ac.at/research/ifip85.html#aim
[5] http://www.ontogov.com/
[6] http://www.egov-project.org
[7] http://www.webocrat.org/
[8] http://www.daml.org/projects/integration/projects-20010811
[9] http://reliant.teknowledge.com/DAML/Government.owl

3 Reimdoc Project

We use a reference model to focus on and build a common understanding of the problem stated; Figure 1 shows the different actors within the e-Gov.

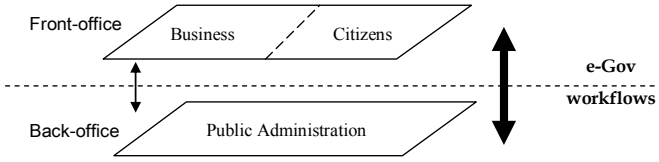

Fig 1. The e-Government Reference Model

In the Back-office, the main actor is the Public Administration; it has many processes inside which should work properly to provide efficient services. The dynamics of the Public Administration provides a huge amount of information to be processed and these data should be managed in a transparent and efficient way.

Within the Public Administration many processes are carried out and these must work properly to provide efficient services; since the Public Administration functions in a decentralized way and the dynamics of this field generates a huge amount of information to be processed, it is necessary to manage this vast amount of information in a transparent and efficient way. Therefore, the implementation of e-Gov ontologies and applications is crucial.

A particular case is being developed in the Reimdoc[11] Project. This project aims to develop tools that allow the legal document to be modelled in electronic support and its semantic retrieval to facilitate the government-citizen document transaction. The domain selected is related to the Real-estate transaction market with the sufficient juridical guarantees.

This project will permit verifying the Real-estate process gathered in digital support. These processes consist of procedures that happen in three areas: the Property Title, the Tributary Administration of the Autonomous Communities and the Justice Administration. In Spain these procedures are meticulously regulated in a coherent form by the context, which is marked by the legal knowledgeable community.

Reimdoc Project is currently developing an application based on the proposed Legal Ontologies described in section 4: EgoIR, an Information Retrieval system. The EgoIR will have as users: final users, which require consulting juridical documentation; agencies, which need to know the current legislation and lawyers, which have to consult concrete aspects. EgoIR is java-based system that offers an ontology-based approach to Information Retrieval. EgoIR has as a main goal to retrieve e-Gov documentation. The system 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.

---

[10] http://dip.semanticweb.org/documents/D9.3e-Governmentontology.doc
[11] http://reimdoc.atosorigin.es/

EgoIR offers an ontology browsing capability. EgoIR uses the ontologies described in section 4. These ontologies are stored in WebODE [4] (workbench for ontological engineering). Also EgoIR allows the query construction from the ontology concepts; the query obtained is composed by a set of concepts extracted from the ontologies. EgoIR connects to WebODE, throughout WebODE's ODE service, in order to obtain ontology concepts. Finally EgoIR employs Lucene[12] (search engine library) in order to retrieve the documents that match the given query.

4 Legal Ontologies

The Legal Ontologies described in this section were built to represent the Real-estate transactions in the Spanish Government domain. These Legal Ontologies were developed acquiring knowledge from academic and private sector experts and built with the methodology METHONTOLOGY [4] and the workbench WebODE [4].

The Legal Ontologies presented are part of an EGO Ontology Model (Figure 2), this model aims to represent a part of the legal processes carried out within the government.

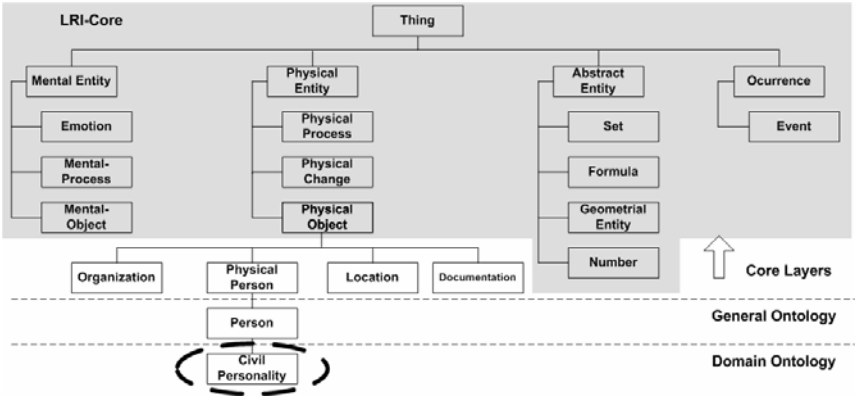

Fig 2. Excerpt of the EGO Ontology Model

The EGO Ontology Model is at present time under development. This model reuses parts of the first two layers of LRI-Core model. The EGO Ontology Model is one of the first efforts not intended for legal domain but for e-Gov domain instead. The e-Gov domain needs take account of law, regulations, citizen services, administrative processes, best-practices and different languages.

4.1 Real-estate Transaction Ontologies Roles

In [10, 15] are identified five main roles for ontologies: organize and structure information; reasoning and problem solving; semantic indexing and search; semantics

[12] http://lucene.apache.org/

integration and interoperation; and understanding the domain. Before building the Real-estate Transaction Ontologies we find useful to settle on the proper role(s) which the ontology will play.

The Real-estate Transaction Ontologies (Figure 3) will play three of the five roles aforementioned: one of the roles will be organizing and structuring information in the e-Gov domain, mainly defining terms used. The second role will be the reasoning and problem solving, basically represents the knowledge of the domain so that an automated reasoner can represent problems and generate solutions for these problems that imply the use of an inference engine to achieve specific goals. The third role will be the semantic indexing and searching, which in this case the ontology will represent the contents of documents, enabling semantic search for content.

Figure 3 shows the relationships between the Real-estate Transaction Ontologies aforementioned (each ontology is represented by a triangle). The aim of this figure is to show all the ad-hoc relations between the Real-estate Transaction Ontologies.

Eleven ontologies have been developed for Reimdoc Project: *person*, *civil personality*, *organization*, *location*, *tax*, *contract model*, *jurisprudence*, *Real-estate transaction verifications*, *Real-estate*, *legislation*, and *Real-estate transaction*. They play individually specific goals and models knowledge used in Reimdoc Project. In the following we describe the relationships between the main ontologies.

The Civil Personality Ontology has as main concept the `civil person`, which is spitted into two subclasses: `natural person` (representing citizens), `juridical person` (representing enterprises, public administrations, etc.). The ad-hoc relations specified for each concept are those whose domain is the concept. For example, the concept `civil person` has six binary relations: 'has data from juridical person', 'has residence', 'is buyer', 'is seller', 'realizes' and 'has data from Natural Person'.

The Real-estate Transaction Ontology has as main concept the `Real-estate transaction`, which is split into two subclasses: `buy` (representing the action of buying), `sell` (representing the action of selling.). The ad-hoc relations specified for each concept are those whose domain is the concept. For example, the concept `Real-estate transaction` has eight binary relations: 'is bought', 'is sold', 'based on' (tax, legislation, jurisprudence), 'acquires ', 'verifies' and 'uses'.

The Location Ontology has as main concept the `location`, which is split into three subclasses: `geographic division`, `town` and `country`. The ad-hoc relations specified for each concept are those whose domain is the concept. For example, the concept `location` has two binary relations: 'is residence' and 'is associated'.

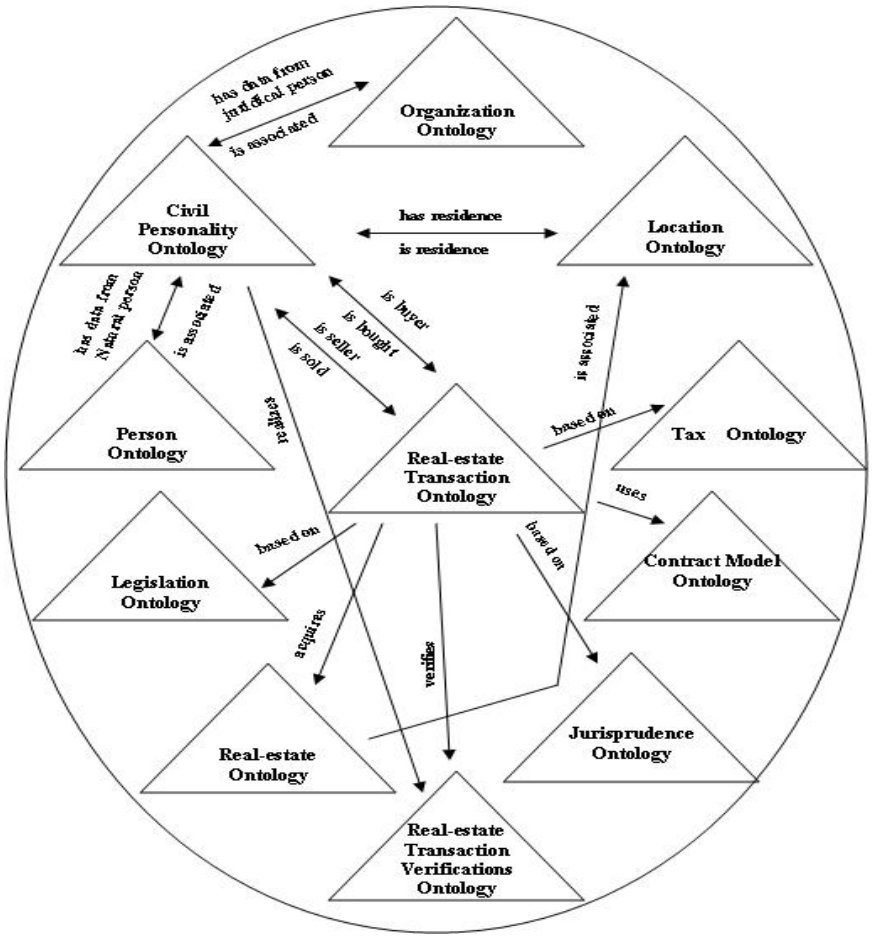

Figure 3. Main ad-hoc relationships for the Real-estate Transaction Ontologies

The Person Ontology has as main concept the `person`. The ad-hoc relations specified for each concept are those whose domain is the concept. For example, the concept `person` has one binary relation: 'is associated'.

The Organization Ontology has as main concept the `organization`. The ad-hoc relations specified for each concept are those whose domain is the concept. For example, the concept `organization` has one binary relation: 'is associated'.

The Real-estate Ontology has as main concept the `Real-estate`. The ad-hoc relations specified for each concept are those whose domain is the concept. For example, the concept `Real-estate` has one binary relation: 'is associated'.

4.2 Main Ontology Modelling Components

METHONTOLOGY [4] proposes to conceptualize ontologies with a set of tabular and graphical intermediate representations. Such intermediate representations allow modeling the components described in this section.

Concepts are taken in a broad sense. For instance, in the legal domain, concepts are: `Civil Personality, Natural Person, Juridical Person`, etc. Concepts in the ontology are usually organized in taxonomies through which inheritance mechanisms can be applied. For instance, we can represent a taxonomy of legal entities (which distinguishes persons and organizations), where a `Real-estate Contract` is a subclass of a `Contract`, etc.

Relations represent a type of association between concepts of the domain. If the relation links two concepts, for example, `has civil personality`, which links `Natural Person` to `Civil Personality`, it is called binary relation. An important binary relation is *Subclass-Of*, which is used for building the class taxonomy, as shown above. Each binary relation may have an inverse relation that links the concepts in the opposite direction.

Instances are used to represent elements or individuals in an ontology. An example of instance of the concept `Contract` is `Contract of merchanting Real estate`. Relations can be also instantiated. For example, we can express that Contract of merchanting real estate has a location in Madrid as follows: `has location(Contract of merchanting real estate, Madrid)`, using a first order logic notation.

Constants are numeric values that do not change during much time. For example: `legal age`.

Attributes describe properties of instances and of concepts. We can distinguish two types of attributes: instance and class attributes. ***Instance attributes*** describe concept instances, where they take their values. These attributes are defined in a concept and inherited by its subconcepts and instances. For example, the `date` of a `Contract` is proper to each instance. ***Class attributes*** describe concepts and take their values in the concept where they are defined. Class attributes are neither inherited by the subclasses nor by the instances. An example is the attribute `First Name` as a part of `Natural Person`. Ontology development tools usually provide predefined domain-independent class attributes for all the concepts, such as the concept documentation, synonyms, acronyms, etc. Besides, other user-defined domain dependent class attributes can be usually created.

Formal axioms are logical expressions that are always true and are normally used to specify constraints in the ontology. An example of axiom is a Natural Person has legal capacity at the age of sixteen if he/she gets married.

Rules are generally used to infer knowledge in the ontology, such as attribute values, relation instances, etc. An example of rule is a Natural Person could be a part of the Juridical Person.

Finally, we show the Real-estate Transaction Ontologies statistics: number of concepts is 58, number of relations is 20, number of attributes is 59 and number of constants is 2.

5 Conclusions and Outlook

We have presented a set of legal ontologies for Real-estate transaction within the Spanish government domain as a part of the EGO Ontology model, which in turn is part of an ongoing project aiming, on the one hand, at supporting semantic applications to retrieve legal documents and, on the other hand, at delivering services from public administration (within the government) to citizens. These legal ontologies are built following the methodology METHONTOLOGY and the workbench WebODE and they are application independent.

The e-Gov domain does still have many needs: knowledge, for instance, has not been modeled at all. These needs represent real challenges for researchers. One problem to be solved in the near future is that of knowledge acquisition by legal experts. Finally, we must add that the legal domain is very complex and evolving and this complexity provides a situation different than that provided by domains such as physics or mathematics, and this fact will bring about the deployment of future e-Gov ontologies.

In our future work, will be focused on further enhancement and evaluation of the Real-estate Transaction Ontologies; we will work towards the reasoning capabilities of these ontologies; we will continue integrating the legal knowledge captured on the EGO Ontology Model and we will compare the model with other ontology models. Also, we will evaluate the EgoIR application in order to improve its performance.

Acknowledgments

We would like to thank our project partners for their helpful comments. Specially thanks to Fernando Galindo, Ana Ibarrola and Ferran Jornet. This work is done 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.

References

1. Breuker, J.: Constructing a Legal Core Ontology: LRI-Core.University of Amsterdam, Leibniz Center for Law, Amsterdam, the Netherlands (2004)
2. Curtain, G., Sommer, M., Vis-Sommer, V.(editors): The World of E-Government. The Haworth Press. (2003)
3. Electronic Government: First International Conference, EGOV 2002, Aix-en-Provence, France, September 2-5,. Proceedings (Lecture Notes in Computer Science) by Roland Traunmüller (Editor), Klaus Lenk (Editor) (2002)
4. Gómez-Pérez, A.; Fernández-López, M.; Corcho, O. Ontological Engineering. Springer Verlag. 2003
5. Gruber, T.R.: A Translation Approach to Portable Oontologies. Knowledge Acquisition. 5(2) (1993) 199-220
6. Gruber, T.R.: Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Presented at the Padua workshop on Formal Ontology, to appear in an edited collection by Nicola Guarino (1993)
7. Hendler, J., Berners-Lee, T., and Miller, E. Integrating Applications on the Se-mantic Web, (2002) http://www.w3.org/2002/07/swint.html
8. Kralingen, R.W.: Frame-based Conceptual Models of Statute Law, Computer/Law Se-ries. No.16, Kluwer Law International. The Hague, the Netherlands (1995)
9. McCarty, L.T.: A Language for Legal Discourse, I. Basic Features, Proceedings of the Second International Conference on Artificial Intelligence and Law, Vancouver, Canada (1989) 180-189
10. Melz, E., Valente, A.: Modelling the Tax Code. The Second International Workshop on Regulatory Ontologies (2004)
11. Neches, R., Fikes, R.E., Finin, T., Gruber, T.R., Senator, T., Swartout, W.R.: Enabling Technology for Knowledge Sharing. AI Magazine. 12(3) (1991) 36-56.
12. Stamper, R.K.: LEGOL: Modelling Legal Rules by Computer. Computer Science and Law. Bryan Niblett. Cambridge University Press. Cambridge, United Kingdom (1980)
13. Stamper, R.K.: The Role of Semantics in Legal Expert Systems and Legal Reasoning. Ratio Juris, Vol. 4, No. 2 (1991) 219-244.
14. Valente, A.: Legal Knowledge Engineering; A Modelling Approach. University of Amsterdam, The Netherlands, IOS Press, Amsterdam, The Netherlands (1995)
15. Valente, A.: Types and Roles of Legal Ontologies. V.R. Benjamins et al. (Eds.): Law and the Semantic Web, (2005) 65-76
16. Visser, P.R.S.: Knowledge Specification for Multiple Legal Tasks; A Case Study of the Interaction Problem in the Legal Domain, Computer / Law Series. No. 17, Kluwer Law International. The Hague, The Netherlands (1995)