Knowledge Representation of Intelligent Public Services through a Semantic Model

Carlos Roberto Brys[1], José F. Aldana-Montes[2]

[1] Departamento de Informática, Facultad de Ciencias Económicas,
Universidad Nacional de Misiones,
Av. Fernando Llamosas Km 7,5. Campus UNaM.
N3304. Posadas. Misiones. Argentina
brys@fce.unam.edu.ar
[2] Departamento de Lenguajes y Ciencias de la Computación, Universidad de Málaga,
Boulevard Louis Pasteur 35. Campus de Teatinos.
29071 Málaga, España
jfam@lcc.uma.es

Abstract. Today citizens make intensive use of mobile communication technology, and they demand to public services providers for complex and sophisticated information. To meet these demands, the governments' services agencies must orchestrate a lot of information from various sources and formats, and deliver them in the data terminals that people commonly use: computers, net-books, tablets and smart-phones.
To overcome these problems, we propose a deductible model for conceptual representation of the organizational units of the State and his services, based on ontologies designed under the Linked Open Data principles. This model allows automatic extraction of information through machines, that support governmental decision-making processes and giving to citizens a comprehensive access to find and make formalities through intelligent agents.

Keywords: Web of Data, Public Administration, Semantic Web, Linked Data, Government Data, Automatic Data Extraction.

1. Introduction

Nowadays, information systems have to feed on a variety of data sources in real time, so that the process of decision-making is the most efficient and successful. The information is distributed over a wide geographical area, provided by many different sources and in different formats. To get the best possible information that supports the decision-making process, governments should find, extract, and process data using automatic machine data extraction, and present information in data terminals that users often use.

To ensure the quality and integrity of the information that governments need to find, filter and process, it is necessary the use of standards that define in detail the data entities and their relationships with others, to enrich their meaning. To make this possible, administrative agencies and the information needed for decision-making, can be represented using ontologies at different levels, and these can be linked to heterogeneous sources through Linked Data. Then the data resources can be incrementally obtained by following the Uniform Resource Identifiers of resources via the HTTP protocol [6].

Recent work in Artificial Intelligence is exploring the use of formal ontologies as a way to specify content-specific agreements for the sharing and reuse of knowledge between software entities. The use of ontologies and Linked Open Data in strategic areas of the State such as education, health and security will substantially improve the decision-making process. This will impact and transform the way the State manages its data, and will offer society itself to propose and solve problems accessing, processing, mixing and relating the Linked Open Data to create innovative solutions. [4].

Knowledge is information about a domain that can be used to solve problems in that domain and to solve these problems this knowledge must be represented in the computer. This representation must be rich to enough express the knowledge to solve the problem closest to the problem; should be compact, natural and easy to maintain. [10].

Ontologies are important for the artificial intelligence and cognitive computing. They are a description of declarative knowledge in the form of classes and the relation between them. In the literature are defined as, "a formal and explicit way to define a conceptualization about knowledge sharing" [3]. The underlying idea behind these technologies is to make computers capable of understanding the data with little or no human intervention.

An ontology allows machines and information systems to understand concepts and domains through a common vocabulary that represents the relationships between them. The structure of ontologies represents knowledge about the world and specific domain knowledge and allows systems to reason as the human mind does through the interplay of data elements within a system of categories.

The advantage of expressing the State organizational structure as an ontology, is that it can build an information model that allows data exploration in terms of the elements that represent the associations between the objects, the properties of elements and formally describe the semantics of the classes and properties used in dependency relationship, temporal and spatial; thus enabling machines to carry out automated reasoning, semantic and conceptual search, and provide services to decision support systems. The Semantic Web is a way to allow the knowledge distributed on the World Wide Web is interpretable by machines.

The paper is organized as follow. Section 2 introduces the theorethical framework, defines the State and their organization, defines what is Linked Data, geospatial data and their application. Section 3 outlines the semantic model that we propose, the selected language for implement it and how we represent the State political organization. Section 4 presents an sample of use case and how to extract information through the SPARQL language. Finally in Section 5 we present the conclusions and future work.

2. Theoretical Framework

We will use the term “State” as a definition of a political concept that refers to a form of social organization and political sovereignty, formed by a group of institutions. These institutions are structured functionally into administrative units, which are the basic elements of organizational structures. Generally, the organization of a State can be distinguished by: Functions, Institutions and Authorities.

2.1. Linked Government Open Data (LGOD) and Spatial Data

Tim Berners Lee [1] outlined a set of rules for publishing data on the Web in a way that all published data becomes part of a single global data space. These have become known as the “Linked Data principles”. Later, in 2010 he suggested a 5 Star Linked Open Data Rating Scheme specially for government data owners, to deploy Linked Open Data:

Linked Data is a standard model for data interchange on the web. This term is used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF. It should not be assumed that they are only datasets. It is a common understanding today that Open Data is simply “data on the web”, whereas Linked Data is a “web of data”. Linked Data is part of the Semantic Web movement which seeks to standardize data formats and create meaningful data and data relationships on the web. Linked Government Open Data is a useful tool to retrieve knowledge from diverse domains such as ontologies, governmental information, geospatial data, publications, and so on. [7].

According to the Geographical Information Platform of the Swiss Confederation [2], 60% to 80% of all decisions affecting citizens are linked to geographical information. Geographical information is becoming even more important in all aspects of life on issues such as transport, energy, protection of environment, agriculture and forestry, development planning, land regulations, information technology and telecommunications, education and culture, insurance, health provision, national defense, internal security, civil defense and disaster prevention, utility supplies and waste disposal. In this sense, GeoNames provides locality datasets that serve as a hub for other datasets that have a geographic component.

GeoNames is a geographical database containing open licensed linked data of about 10 million locations and geographical names. These datasets are linked to the OpenStreetMap project data. OpenStreetMap is a collaborative project to create a free and editable map, that provides information on over 350 million geospatial features and describes some 25 million km. of routes and streets. Whenever possible, the places GeoNames are linked to corresponding locations on DBpedia, ensuring that there are a number of interrelated data on geographic locations.

LinkedGeoData adds a spatial dimension to the Web of Data. It permits lifting OpenStreetMap data into the Semantic Web infrastructure, and makes it available as an RDF knowledge base according to the Linked Data principles. This simplifies information integration and aggregation tasks that require comprehensive background knowledge related to spatial features [12].

3. A Semantic Model for Public Administration

The administrative units of government are not a stable collection of instances. Their diversity and their evolutionary nature, need a model that eases and simplifies management use. To obtain quality, speed, opportunity and confidence in public services, the customers (citizens and businesses) should be able to find through intelligent agents the services they need and receive all the information clearly and concisely. The tool they will use will be a semantic searcher that will find this information collected and described in a semantic model.

From the perspective of an ontological model, we can represent public administration as a large virtual organization. Evolving capabilities of IT, the semantic web and the artificial intelligence allows the development of virtual organizations that exploit the capabilities of these new technologies. [9].

We follow the strategy for data integration proposed by Salmen [11], looking for the semantic improvement. The strategy can be applied incrementally; creates minimal barriers to the incorporation of new data in the semantic system; preserves existing data in its original form, and covers the entire range of data sources, types, models and their modalities.

In order to specify the ontology to represent semantically the State, his organizational structure and the inherent knowlegde we start from the three-layer model proposed by Guarino [5] and reviewed by Lacasta-Miguel [8], and we extend it by adding a new layer of Linked Open Data ontologies.

The extended ontology model have:

- A high-level ontology which defines the concepts of the State, the Powers, the legal framework, the basic concepts, types of data and general relations, independent of context.
- Other external ontologies and their linkage relationships through the Linked Open Data which enrich the instanced indivual's data and gives the geospatial information.
- And various application ontologies of formalities and services, and their instances.

For express the ontology in a machine readable language we use the Ontology Web Language (OWL) through which we add classes, attributes, relations and geolocation data. We choose OWL because is the proposed standard for Web ontologies. Formal semantics and reasoning support is provided through the mapping of OWL on logics. OWL builds upon RDF and RDF Schema and has the same kind of RDF's XML syntax, and extends RDF/RDFS by providing additional vocabulary along with a formal semantics. Also allow to describe the semantics of knowledge in a machine-readable way. The RDF has emerged as the main data model for representing information about the resources available in the Web. For querying the RDF-based knowledge bases we use the SPARQL language. [13].

In order to model a general ontology compatible with the political organization of different countries, we will use the basic jurisdictional domains: “State”, “Division”, “Organization” and “Sub-organizations” as defined in ISO/IEC 15944-5:2008. “State” is an entity with its own legal nature which may be national in scope (country), or sub-national (province). The “Divisions” are different partitions of higher level organizations (ministries) which are turned into smaller entities: under secretariats, general directions, directions, and departments which are the basic division entities. [8].

A domain ontology which defines the organizational structure and describes in detail the specific administrative units, their hierarchies, dependences and relations.

This relationship between the State, government and its divisions can be represented in Description Logic with a *has* property, the subset, existential quantification, and union symbols as follows:

State ⊆ ∃ hasPowers.(ExecutivePower ∪ LegislativePower ∪ JudiciaryPower)
ExecutivePower ⊆ ∃ hasDivision.(Governor ∪ (Ministry$_i$ ∪...∪ Ministry)
Ministry ⊆ Undersecretariats ⊆ General Directions ⊆ Directions ⊆ Departments

We are not completely aligning ontologies with existing ones because for our model it is very difficult or unnecessary to pursue the creation of a single global ontology, but instead we are building the ontology of e-government from the perspective of citizen's services and Linked Data. In this sense we want to answer the questions of competence populating the ontology with mapping to other ontologies and data sets. Then we want to align only some classes and instances that address the questions of competence. For that, we only consider one-to-one mappings between single entities and instances in different ontologies and we do not consider inverse functional properties.

4. Application Scenario

In a sample of use case, we consider a service where a citizen requires to update his ID card. The citizen must go to the nearest documentation center to begin the formality. For that, the citizen provides his location and the formality ID: (Posadas, Nuevo_DNI). As shown in Fig. 1, we graph an example of the interlinked semantic instances of "Nuevo_DNI" and "Posadas".

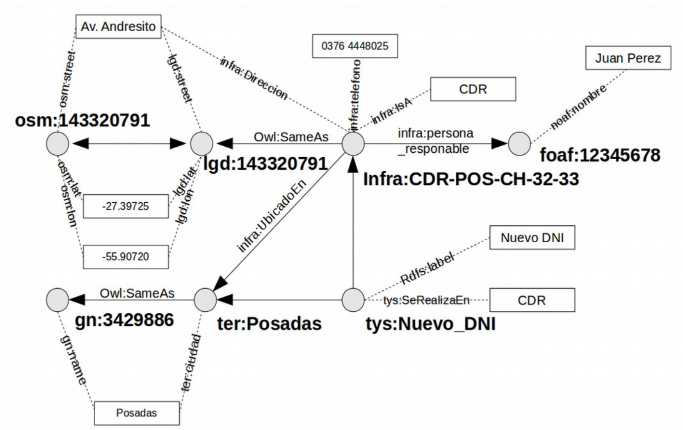

Fig 1. A semantic network of interlinked instances allowing inheritance

Definitions: V is a set of vertices that are the labels of data sets having the linked SameAs instances, E ⊆ V×V is a set of sameAs edges, and I is a set of URIs of the interlinked SameAs Instances.

GDNI = (V, E, I), where:

V = {S, T, O, L, I, G, F},

E = {(S,T), (S,I), (I,F), (I,L), (L,O), (T,G)},

I= {tys:Nuevo_DNI, ter:Posadas, osm:143320791, lgd:143320791, infra:CDR-POS-CH-32-33, gn:3429886, foaf:12345678}.

S, T, O, L, I , G anf F represents the labels of data sets Formalities and Services, Territory, OSM, LinkedGeoData, Infrastructure, GeoNames and Personnel respectively.

Through the SPARQL language we query the information explicity represented within the ontology. The implicit triples are the ones that may be inferred from the rules that characterize each interpretation. The query obtains the list of documentation centers in Posadas where doing the formality, informs who is the responsible person, attention hours, costs, and so one.

```
PREFIX ter: <http://www.datos.misiones.gov.ar/ontologias/gobierno/ar.gp.N.territorio.owl#>
PREFIX tys: <http://www.datos.misiones.gov.ar/ontologias/gobierno/ejecutivo/ar.gp.N.pe.tys.owl#>
PREFIX infra: <http://www.datos.misiones.gov.ar/ontologias/gobierno/ejecutivo/ar.gp.N.pe.infraestruc-
tura.owl#>
SELECT *
WHERE {
          ?Tramite tys:identificador "Nuevo_DNI" .
          ?Tramite tys:SeRealizaEn ?TipoLugar .
          ?Lugar rdf:type ?TipoLugar .
          ?Lugar infra:localidad ter:Posadas .
          ?Lugar infra:domicilio ?domicilio .
          ?Lugar infra:telefono ?telefono .
}
```

(SPARQL code for query the ontology whith user requirements)

5. Conclusions and Future Work

We have proposed an original approach that shifts the way in which the existing literature represents the knowledge inherent of public administrations. With this model, the information of public services are those that find the citizen and bring it wherever they are, through intelligent agents. By applying the proposed experimental model, we demonstrated that an ontology of geolocated public services could be used to integrate different government services distributed geographically that are relevant to citizens.

We currently have available tools to use SPARQL queries from native applications for smartphones. Applications such as *MIT App Inventor,* an open source web application which allows that SPARQL and RDF can be used as part of a mobile application.

With this tool we aim to develop mobile applications that use intelligent agents to extract the knowledge from the ontology and bring the public service to citizens.

References

1. Berners-Lee T. (2006). Linked Data. http://www.w3.org/DesignIssues/LinkedData.html
2. GPSC The Geoportal of the Swiss Confederation. (2011) http://www.geo.admin.ch/internet/geoportal/en/home.html
3. Gruber, T. (1993). A translation approach to portable ontology specifications. Knowledge Acquisition 199–220.
4. Gruber, T. (1995) Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal of Human-Computer Studies. Volume 43, Issues 5–6, Pp 907–928. DOI:10.1006/ijhc.1995.1081.
5. Guarino N. (1998). Formal Ontologies and Information Systems. Proceedings of FOIS'98, 3-15. Trento, Italy
6. Heath T., Bizer C., (2011). Linked Data, Evolving the Web into a Global Data Space, Synthesis Lectures on the Semantic Web: Theory and Technology, ISBN: 9781608454303
7. Jain, P., Hitzler, P., Sheth, A. P., Verma, K., & Yeh, P. Z. (2010). Ontology Alignment for Linked Open Data. (P. F. Patel-Schneider, Y. Pan, P. Hitzler, P. Mika, L. Zhang, J. Z. Pan, I. Horrocks, et al., Eds.)Information Retrieval, 6496(November), 402-417. Springer.
8. Lacasta-Miguel, J., López-Pellicer, F.J., Floristán-Jusué, J., Nogueras-Iso, J., Zarazaga-Soria, F.J.(2006): Unidades administrativas, una perspectiva ontológica. Avances en las Infraestructuras de Datos Espaciales. Treballs dínformática i tecnologia. Castelló de la Plana: Universidad Jaime I de Castellón. (85-94). ISBN 84-8021-590-9.
9. O'Leary, D. Kuokka, D. Plant, R. (1997). Artificial intelligence and virtual organizations. Communications of the ACM CACM Homepage archive. Volume 40 Issue 1, Jan. Pp. 52-59.
10. Poole, D. Mackworth A.(2010). Artificial Intelligence. Fundation of Computational Agents. Cambridge University Press. ISBN: 9780521519007
11. Salmen, D. et al.(2011). Integration of Intelligence Data through Semantic Enhancement, Proceedings of the Conference on Semantic Technology in Intelligence, Defense and Security (STIDS), George Mason University, Fairfax, VA, November 16-17, , CEUR, Vol. 808, 6-13.
12. Stadler C., Lehmann J., Höffner K., Auer S., (2011) LinkedGeoData: A Core for a Web of Spatial Open Data. Semantic Web -Interoperability, Usability, Applicability an IOS Press Journal, 1570-0844/0-1900
13. Trigo, P. Coelho H. (2009) Agent Inferencing Meets the Semantic Web. In Book Progress in Artificial Intelligence. Lecture Notes in Computer Science. DOI 10.1007/978-3-642-04686-5_41. Pp 497-507.