ONTOLOGY SERVER FOR THE INTEGRATION OF DEMOGRAPHIC SERVERS

I. Román*, L.M. Roa**, G. Madinabeitia* and A.M. Oria*

* Area of Telematic Engineering, University of Seville, Seville, Spain
** Biomedical Engineering Group, University of Seville, Seville, Spain

isabel@trajano.us.es, german@trajano.us.es, laura@esi.us.es

Abstract: One of the first steps to recover the clinical information about a patient in a distributed environment is his/her unequivocal identification in each of the involved systems and the query of his/her demographic information. These are the main tasks of a demographic server.

Systems that have been developed independently might have functional and semantic incompatibilities in their demographic servers [1,2]. The aim of this work is to offer support to the resolution of semantic conflicts in order to ease the integration of demographic servers through a unified interface. This would allow the identification of a person and the query of his/her demographic information from several heterogeneous systems.

Introduction

Clinical information systems are usually developed in a local and proprietary manner, created for the specific requirements of a healthcare provider. For this reason together with the present healthcare delivery model the clinical information about a patient is frequently stored in several independent and heterogeneous information systems. The compilation and visualization of the whole clinical information related to a patient, the complete Electronic Healthcare Record (EHR), can make easier the healthcare professional tasks and improve the quality of service.

In a conventional system, the first step to recover the clinical information about a patient is his/her identification. This is usually accomplished by sending some combination of identity parameters or person traits [3]. When a candidate is found the identification is finished and a unique identifier, or PersonID, is recovered and used in this particular system to reference this patient. Most of these systems assign and maintain PersonIDs autonomously. This management procedure is suitable for recording and retrieval of patient's records within the local organization.

After the identification of the person, some demographic data about the patient could be requested, because this information should be annexed to the EHR extract [4,5]. When the system has been developed like a modular architecture, the identification and the recovery of demographic information are tasks usually assigned to "Demographic Servers" [3].

In a distributed scenario with heterogeneous systems, the performance of these activities is more complex [1,2]. The aim is to identify a particular patient in every system and to obtain a unified view of his/her demographic information stored in different locations. The systems should maintain their autonomy but collaboration is needed in a federated approach. The interoperability issues could be summarised as follows:

Functional: the interfaces to manage identifiers and demographic data are different.

Semantic: the information models or database schemas are different.

Instance: the information about the same patient could differ.

A solution to interoperability issues is standardization. Following this philosophy a service for person identification functions has been normalized by CORBAmed group from OMG[3]. This standard solves the functional and part of the semantic problems. If all the systems of the federation offer the defined interfaces for managing identification and querying demographic data, the functional problem is solved. But although an information model has been defined for concepts like PersonId or Person Trait, more work has to be done for semantic conflicts because this standard does not give the specific semantic for person traits. The problem of instance reconciliation is not considered in the standard.

This work contributes to the resolution of semantic conflicts applying ontology techniques in the domain of demographic information. The instance reconciliation is addressed in other works of our research group [6].

Materials and Methods

Our work is supported in several standards, widely accepted by the healthcare community. As mentioned in the introduction, in order to solve functional incompatibilities we use the person identification service (PIDS) [3] developed by the CORBAmed group from the OMG.

The PID server defines several interfaces for person's location, demographic information query, management of identifiers and others. If some attributes about a person are known, a query over the system can be made based on this profile. The profile is specified as a set of elements, in the traditional Name/Value pair, known as traits. But the standard does not normalize the valid set of names or values for these traits. This is the main obstacle for semantic compatibility between demographic servers based on PID standard.

ISSN: 1727-1983

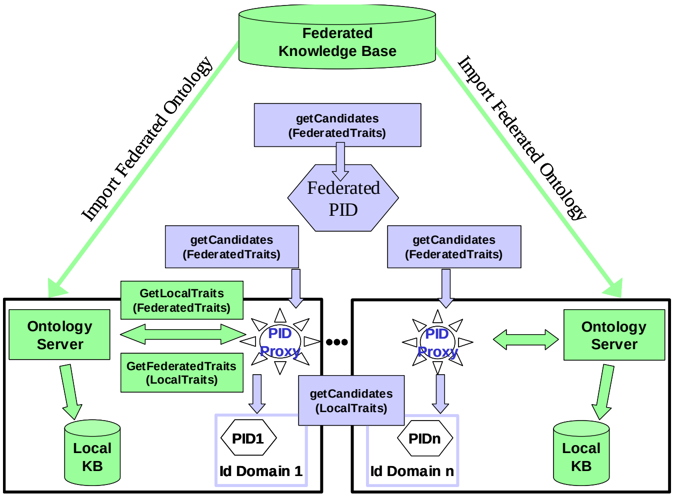

Figure 1: Architecture for the integration of PID servers

Another standard useful for interoperability and that could solve semantic conflicts is the Non-clinical General Purpose Information Components (GPIC) [7], from CEN. This standard addresses the definition and structure of information related to entities that are commonly encountered in communications with and between clinical information computer systems.

Among others, GPIC defines components related to subjects of care (including persons and animals), subject of care related parties and healthcare agents (including healthcare professionals, organizations and devices). In our work this standard is used to give a common specific semantic of traits in the federation, a federated ontology. But the local ontology of each system could differ from this common one. We work with ontology techniques for the reconciliation, as is explained in the next section.

OWL [8] has been chosen as the language for the specification of our ontologies for two reasons. First, because it is the standard defined by W3C and second, we are studying the possibilities of Semantic Web techniques for solve some healthcare services integration problems. But the main intention of the work is that the developed methodology could be applied with different technologies.

As far as software is concerned Protégé[9] has been used for ontology edition and Jena API [10] for ontology management. Both are open-source tools.

Results

Figure 1 shows the integration architecture for demographic servers. The federated knowledge base stores the federated ontology. This common ontology describes the demographic concepts and their relations. The design of this ontology is supported in the Non-clinical GPIC from CEN that normalized demographic information components. All the relations and restrictions specified in the standard have been included in this ontology. Figure 2 only shows the class hierarchy of the federated ontology.

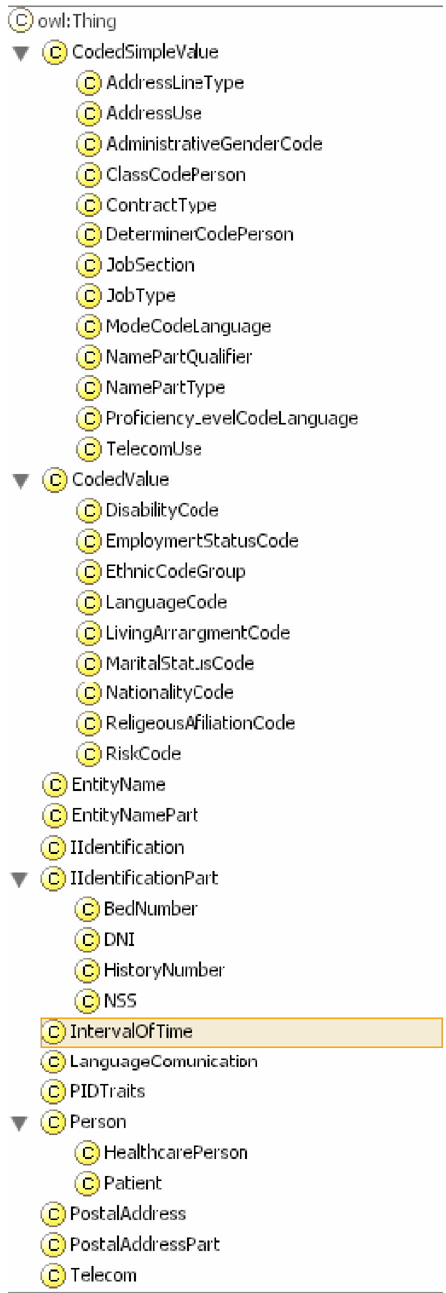

Figure 1: The federated ontology for demographic information

In each of the domains, managed by a specific demographic server, the relations between federated and local concepts are settled down in the local knowledge base. The integration method is very easy. When a new system is going to be added to the federation, the knowledge engineer has to import the federated ontology and make the correspondences between local and federated concepts. All the traits that can be locally queried as traits are specified as PIDTraits, a concept included in the federated ontology, as figures 2 and 3 show. The current development of ontologies merging could improve this task making it more automatic. The result is a local knowledge base in which the equivalences and restrictions between local and federated ontologies have been established, as detailed in figure 3.

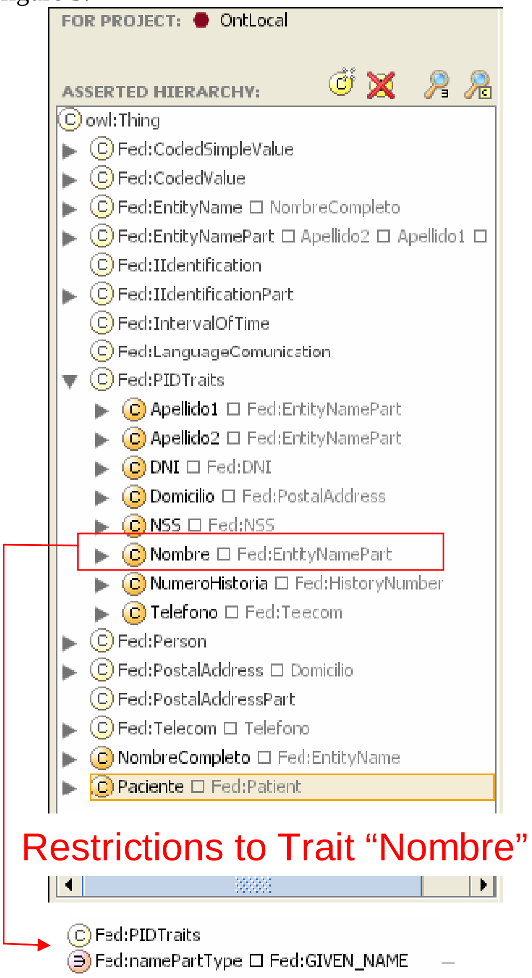

Figure 2: Mapping local ontology with federated.

The use of an extended and open standard like GPICS for the definition of the federated ontology guaranties that every concept in the local system is covered in the federated ontology. Therefore a correspondence is always easy to establish and changes to the federated ontology are expected to be minimum.

Once the ontology support has been introduced, the agents for functional interoperability need to be described. The federated PID offers CORBAmed interfaces but it does not act over a domain but over the whole federation. In fact this is a complex component because in order to solve instance problems it interacts with some other components not represented in figure 1. This simplification has been made because this work is focused on semantic but not on instance reconciliation. This federated PID interacts with each domain belonging to the federation.

In each local system three agents are identified:

PID: It is the local demographic server. This is used by local components to manage identification and demographic data recovery. In figure 1 the interface of this server is compliant with CORBAmed one, but this is not necessary.

PID proxy: It is the bridge between the federation and the local domain. The PID proxy interacts with the local PID. The methods, protocols and the semantic of the person traits in these interactions are the local one. It offers a CORBAmed-compliant interface to the federated PID, in which the trait name and data types comply with the federated ontology. To make the translation from local to federated concepts, and vice versa, the ontology server is used. This component can act as a technology proxy too, because the local technology could differ from the federated one. This allows, for example, CORBA to be used locally and web services in the federation.

Ontology server: It is the main agent presented in this work. It has a very simple interface used by PID Proxy for the translation of traits semantic. This is not a trivial problem because the equivalences are not always one to one, and more complex relations and restrictions could have been established between local and federated concepts. The ontology server manages all these relations and restrictions and shows them through an easy interface. The main methods of this server are:

- *GetLocalTraits*: The input parameter is a federated concept and the output is a vector with all the local traits needed for the definition of this federated concept. For example, if the federated concept is EntityName the local traits could be “*Nombre*”, “*Apellido1*” and “*Apellido2*”. Is evident that language is not a problem once the equivalences are well established in the local knowledgebase.
- *GetFedTraits:* It has the complementary functionality. The input is a local concept and the output a vector with the corresponding federated traits.
- *GetRestrictionsList:* The input parameter is a concept and the output is a vector with elements representing the restrictions imposed to a property of the class.

Some other methods with more complex functionality have been defined and implemented. For example several traits could be translated in only one call or a local concept can be translated to federated and federated to local in the same call.

Discussion

The architecture and methodology presented in this paper for semantic conflict resolution in the integration of demographic servers could be easily extended to other components in a healthcare organisation. At present the integration of healthcare information systems is based on the use of standards, and most of these are focused on the definition of information models that give a common semantic to the different systems. But usually the systems that need to be integrated are not been developed to comply with a standard, or different standards are used for different systems. The methodology that we have presented represents an approach to the semantic integration without the modification of the local information model.

All the design is inspired on an open architecture philosophy, so the software could be reusable too for different domains inside the healthcare organisation. Although the ontology server in this case is used to translate person traits, it could be used to translate every other kind of concepts, like clinical observations or security parameters. The only difference is that the ontology will represent a domain that is not demographic information, like in this case, and that clients are supposed to query for concepts that are not person traits.

The decision to use local ontology servers instead of a centralized one was made looking for the scalability and reliability of the system. The aggregation of a new domain does not imply any change in the federated knowledge base and the failure of an ontology server does not make the whole system to fail.

The decomposition of tasks and the assignation to specific components developing a distributed scenario are fundamental for scalability, and represent a powerful tool to facilitate the integration in a federated environment [11]. But the situation of the systems to integrate could be very different and the design of these systems could not follow this philosophy when they where developed. Therefore integration can be guided by a methodology but it has to be studied for each particular case. An important lesson learned from the difficulties during integration activities is that integration has to be in mind during the design and development of each particular system and it cannot be considered a secondary task.

Conclusions

The main contribution of this work is the introduction of an architecture for the resolution of the semantic conflict in person traits, which arise during the integration of several demographic servers.

We have chosen a common ontology for demographic information, based on GPICS, which is a CEN well accepted standard, and we have expressed it in OWL.

We have developed an ontology server that acts over a local knowledge base that contains the relations between local and federated ontology. This is used by the proxy PID, a component that allows making queries to the local system using a federated concept.

We are developing the proxyPID adding some interesting functionalities that extend the interfaces specified in the PID service. For example it can construct standard objects for demographic information like EX_PARTY from prEN13606-1 [5] or a PARTY from demographic information model of openEHR [4], objects that are usually annexed to the EHR. Some improvements for instance reconciliation are in progress too [6].

Acknowledgements

This work has been supported by the Spanish Ministry of Science and Technology TIC2003-09300-C04.

References

[1] FORSLUND D.W., SMITH R.K., CULPEPPER T.C.: "Federation of the Person Identification Service between enterprises," Proc. AMIA Symp. 2000, pp. 240-244, 2000.

[2] FACCHINETTI S., BERRY R.K., PARDON J., GRIMSON J.: "Distributed Patient Records require distributed patient ID Management: enter PIDS?," Notes from PICNIC conference.

[3] CORBAMED FROM OMG. Person Identification Service (PIDS) Specification. April 2001,Version 1.1.

[4] The openEHR Demographic Information Model. Revision 1.4.4. © Copyright openEHR Foundation 2001-2003. All rights reserved. www.openEHR.org

[5] prEN 13606-1: CEN/TC 251 Health informatics — Electronic health record communication — Part 1: Reference model.

[6] ROMÁN I., ROA L.M., REINA L.J., MADINABEITIA G: "Standardization of demographic service for a federated healthcare environment". Medical and Care compunetics 2 L. Bos et al. (Eds.) IOS Press, 2005-06-08, ISBN I-58603-520-7, ISSN 0926-9630

[7] prEN 14822-1:2003 (E)CEN/TC 251 WI. Health informatics — General purpose components for messages — Part 2: Non clinical

[8] Web Ontology Language, by W3C. http://www.w3.org/2004/OWL/. Last access September 2005

[9] Protégé Project. http://protege.stanford.edu/. Last access September 2005

[10] Jena website: http://jena.sourceforge.net/ Last access September 2005.

[11] CEN/TC 251, prEN 12967-1: Health Informatics — Service architecture

[12] DOGAC A., LALECI G., KIRBAS S., KABAK Y., SINIR S., YILDIZ A., GURCAN Y. "ARTEMIS: Deploying semantically enriched Web services in healthcare domain" Information Systems, Article in press, available online at www.sciencedirect.com