Nomenclature Case Study

Authors: Tomás Pariente-Lobo, German Herrero

Introduction and motivation

In the last few years we have witnessed the increasing attention of the health experts towards what is called Semantic Interoperability in eHealth. One of the pillars to achieve semantic interoperability within the health domain is precisely to have a common or interoperable description of drugs. Several efforts have been made in this regard as, for example, the widely usage of knowledge resources such as Snomed-CT, HL7 in certain countries. More important to us is the extended idea that ontologies are a very useful way to describe drug models. Initiatives such as BioPortal, where access to the most widely used ontologies in the biological community can be found; or the OBO Foundry, which states a language and a set of principles for creating biomedical ontologies; or the recommendations from roadmaps, such as the one delivered by the SemanticHEALTH project, are all clear examples that ontologies are becoming an undeniably way to describe the entities of the health domain and to provide the description of the drugs. However, it is not foreseen that a new standard for drugs description will appear in the near future. Nevertheless, it is also clear that isolated formal descriptions are not useful when talking about semantic interoperability. Mappings between different descriptions should be made to achieve interoperability. It is here where NeOn, and particularly the approach developed within this case study, could be helpful.

There is also a different angle to consider, which is related to the Open Linked Data initiative. Linked Data is a term 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. Among the current Linked Data datasets there are several health-related open resources such as Drugbank, Dailymed or Diseasome, besides other generic-purpose resources containing information about drugs, such as DBPedia .

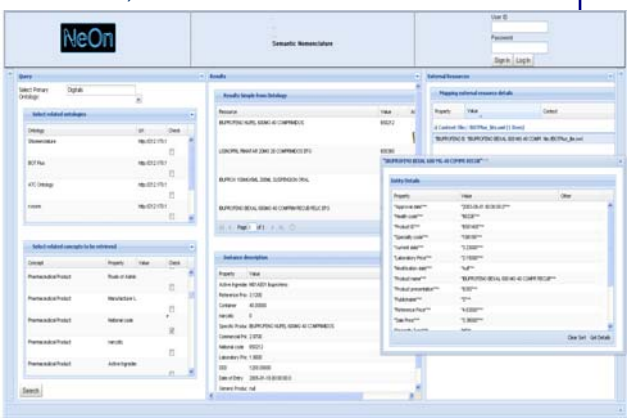
The Semantic Nomenclature case study is developing a network of interconnected pharma ontologies, whose aim is to help to bridge the gap between different drug terminologies. The case study is reusing and engineering ontologies based on pharmaceutical resources and using NeOn technology

Application scenario

The Semantic Nomenclature is focused on providing the infrastructure and guidelines for interconnecting ontologies in the pharma domain within a growing and open environment. While creating the Nomenclature ontology network, we faced different scenarios:

- Reusing and reengineering non-ontological resources. An example of this scenario is the engineering of the Digitalis ontology from the Nomenclator Digitalis database provided by the Ministry of Heath, or the ATC ontology that comes from the classification provided by the WHO. The SPC ontology is also the result of engineering the Summary of Product Characteristics specification produced by the EU.
- Reusing ontological resources. In the eHealth domain we can already find several ontologies representing drugs descriptions. Examples of this are Galen or the UMLS semantic concepts.
- Reuse and engineer ontological resources. In order to develop a suitable classification and mapping schema, several ontological resources (SNOMED-CT, RxNorm, etc.) have been studied and the result is the new Nomenclature ontology.
- Reusing Ontology Design Patterns (ODPs).

The case study also provides a prototype Web 2.0 application on top of the ontology network. Its aim is to ease the access to the underlying knowledge base and it demonstrates the added value of the solution envisaged by NeOn. This application is an example of how to build applications on top of networked ontologies [3] .



Nomenclature Case Study

Ontology development life cycle

The ontology network life cycle model chosen in the Semantic Nomenclature case study is the iterative-incremental model. On the one hand, the usage of this life-cycle model was imperative in a project such as NeOn, where the methodology and tools to model and map ontologies was being created at the same time as the case studies were running. On the other hand, being a long lasting case study (4 years long), and given the expected changes in the ontology landscape on the pharma domain, it was also advisable to follow several iterations in order to come up with a good ontology network.

In the Semantic Nomenclature we had three iterations. In each iteration we planned the ontology development for that phase as a waterfall model. The last iteration was planned following the methodological support provided by the gOntt plug-in of the NeOn Toolkit, which consists basically in a six-phase waterfall model that allows the reuse of both ontological and non-ontological resources plus reengineering.

Ontology requirements

a. Non-Functional Requirements

- The Semantic Nomenclature should target international terminologies
- The result would be an easy to extend network of ontologies
- The resultant network should be evaluated by domain experts

b. Functional Requirements (groups of competency questions)

- CQG1. Competency questions about pharmaceutical products
- CQG2. Competency questions regarding Laboratories
- CQG3. Competency questions regarding active ingredients and substances
- CQG4. Composite competency questions

Intended End-users. The requirements below reflect the views and features expected for the different set of users of the Semantic Nomenclature:

- U1: Ontology Engineers on the eHealth domain: maintenance and enhancement of the ontology network
- U2: Domain experts with limited knowledge about ontologies: navigation, search, rating and minimal enhancements
- U3. Pharmacists (users with no ontological knowledge): navigation and search
- U4. Application developers: maintenance and generation of Web-based applications on top of the ontology network

Ontology processes and activities performed (1)

The processes and activities carried out for building the Semantic Nomenclature ontology network are the following:

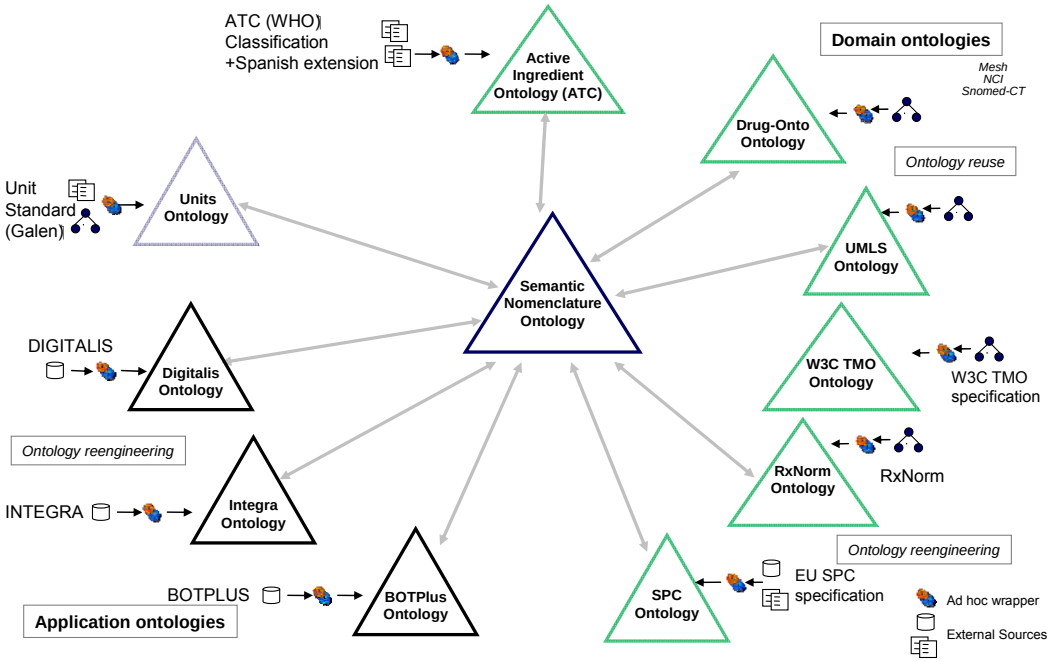
1. **Knowledge Acquisition and Ontology Elicitation**: The pharmaceutical domain was specified and studied in [2] . For each iteration, we included new and different interviews with domain experts from the domain (hospital staff and doctors), who added new requirements for the problems of integration between the different actors when exchanging information about drugs; they also told us about the difficulties of finding a classification of clinical drugs (not branded drugs) in hospitals. From these interviews, we identified a problem of ambiguity: how to difference between clinical drugs and branded drugs.

Nomenclature Case Study

Ontology processes and activities performed (2)

2. **Selection of the standards that cover most of the identified needs**. Some pharmaceutical classification systems and different thesaurus, taxonomies and vocabularies were identified in [2] . After the second iteration, the focus was on extending the ontology network while beginning to create mappings between drug models and some international terminologies adopted by the international bodies in this domain. Medical vocabularies, terminologies and ontologies in the eHealth domain, like Snomed-CT, MeSH, RxNorm, DM+D, all suggested by domain experts, were analyzed. Also, new standards used for describing pharmaceutical products, like the SPC template recommended by the WHO, were added to the knowledge base.
3. **Semantic enrichment of the standards**. Some of the new terminologies, medical vocabularies or documentation selected to be reused in the scenario of the case study were non ontological resources. So, these resources needed Ontology Reengineering and Ontology Enrichment activities to obtain OWL ontologies before reuse them in the model.
4. **Evaluation of Ontology Content, Ontology Conceptualization, Ontology Formalization, Ontology Integration and Ontology Evaluation of the Nomenclature ontology network**. Along with domain experts we performed a two-iteration evaluation of the network of ontologies.
5. **Ontology Implementation of the Nomenclature ontology network**. The language selected to describe the resources is OWL. New ontologies are added from scratch (SPC Ontology); others are connected to the network or have evolved from previous versions (Nomenclature Ontology).
6. **Maintenance Activities**. As part of the iteration model, the ontologies evolve according to the changes and suggestions given by domain experts from hospitals. Following the NeOn Methodology, the activities involved in this stage are Ontology Documentation, Ontology Configuration Management, Ontology Assessment and Ontology Verification & Validation
7. Other activity related to the ontology lifecycle and development of the Nomenclature ontology network is the Ontology Localization. However, in this regard little work has been done in the case study.

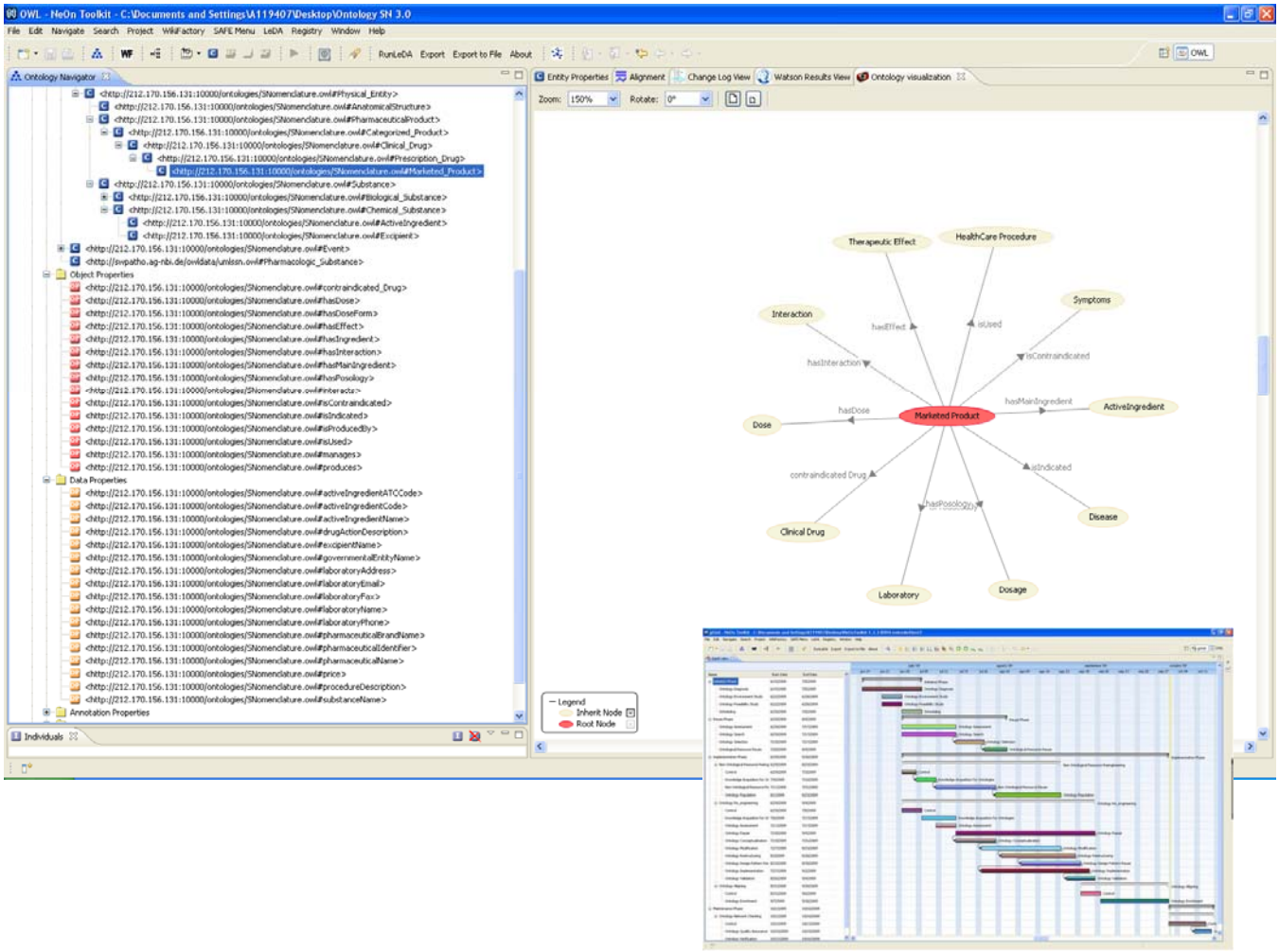
As a result of applying these processes and activities, the following networked ontology was produced for the Semantic Nomenclature case:



Nomenclature Case Study

Technology used

During the different iterations, the development of the ontology network has been carried out using, on the one hand, the NeOn Toolkit, and plug-ins such as the gOntt plug-in for managing the ontology life-cycle; the R2O-ODEMapster plug-in for non ontological reuse; the Alignment plug-in for ontology alignment; RaDON for ontology repair; Cicero for ontology argumentation; ODP to enhance the ontology network; and CupBoard as ontology store; and on the other hand, using infrastructure for domain experts and ontology engineers and the Web application infrastructure.



Additional information:

Contact person: tomas.parientelobo@atosresearch.eu

- Gómez-Pérez J.M, Pariente T, Buil-Aranda C, Herrero G. “Ontologies for the Pharmaceutical Case Studies”. NeOn Deliverable D8.3.1, 2007. [1]
- Gómez-Pérez J.M, Pariente T, Daviaud C, Herrero G. “Analysis of the pharma domain and requirement’s”. NeOn Deliverable D8.1.1, 2006. [2]
- Herrero G., Pariente T, Jamin E, “Second Prototype of the Semantic Nomenclature. NeOn Deliverable D8.5.2, 2009. [3]