A PROTÉGÉ PLUG-IN-BASED SYSTEM TO MANAGE AND QUERY LARGE DOMAIN ONTOLOGIES

Ernesto Jiménez-Ruiz
Victoria Nebot
Rafael Berlanga
Ismael Sanz
Alfonso Rios

Jaume I University of Castellon (Spain)

Maat Gknowledge, Valencia, (Spain)
Outline

- Context and Motivation
- Ontology Management System
- Conclusions and Future Work
Bioinformatics Domain

Health-e-Child Project

PhD Thesis

Developed System
Health-e-Child Project

- **General Objective**: Aims to develop an integrated healthcare platform for European pediatrics, achieving a comprehensive view of children’s health
  - Grid Architecture
  - Main Upper Level Applications: KDS, DSS

- **Our tasks**: Integration of biomedical data, information, and knowledge.

- **Web**: [http://health-e-child.org](http://health-e-child.org)
Health-e-Child Project

- The biomedical information sources will cover six distinct levels (vertical levels):
  - Molecular
  - Cellular
  - Tissue
  - Organ
  - Individual
  - Population

- And will focus on three representative diseases (inside paediatrics):
  - Heart diseases
  - Inflammatory diseases
  - Brain tumours.
Application of current Ontologies in HeC

- HeC *vertical abstraction levels expressed by* Ontologies
- Available several large biomedical ontologies and taxonomies, e.g.: GO, GALEN, FMA, NCI-Thesaurus, Tambis, BioPax, etc.
- Difficult too apply in concrete applications like HeC:
  - Scalability in reasoning.
  - Specificity: local view of the domain
  - Visualization and treatment
From Domain Ontologies to Applications (PhD Topic)

Thesis Proposal Available at: http://krono.act.uji.es/people/Ernesto
System Architecture
OWL Parser and Constructor

- Greater flexibility in the **OWL treatment** and storage capabilities (e.g. indexes)

- The OWL parser creates from the OWL file a set of structures for classes, properties, nominal and individuals.

- These structures will be stored in the graph-based database G.
G Semi-structured Database

- Backend to store, index and retrieve the OWL ontologies as graphs.

- Four database object types are needed: *ontology, property, concept, and enumeration (nominals)*
OntoPath Query Language

Disease / related_to / Rheumatoid_Factor
Ontology Editor Protégé

- Selected as the front-end to visualize and manipulate ontologies
- Flexibility to be extended
- Some plug-ins and GUIs has been created
Protégé Extensions

- Storing Ontologies
- Retrieving full ontologies or fragments
- Representation in a definition hierarchy
- Connection with Python codes
Storing Ontologies

SAVE OWL ONTOLOGY INTO THE ONTOLOGY-VIEW HIERARCHY REPRESENTATION

OWL File Selection

Biomedical (HeC) Coverage

References to other Ontologies (Views)
Retrieving full ontologies or fragments

Ontology Management System

Several Fragments
Source Ontology
Set of OntoPath Queries
Metadata

VIEW FRAGMENT DEFINITION OVER AN ONTOLOGY WITH ONTOPATH (FRAGMENT 2)

Ontologia
2-ncitthesaurus.owl

Query
Rheumatoid_Arthritis / has_Gene_Associated /*
Rheumatoid_Arthritis / ? /Interleukin-17
Rheumatoid_Factor / ? /

Name of the View
Simple_Rheumatoid_Arthritis_View

NameSpace for the View
http://protege4.thbg.uji.es/SRA/view#

Comments for the View
Simple View about the Rheumatoid Arthritis disease

Ok
Cancel
New Fragment
Representation in a definition hierarchy

Classification by Biomedical Level

Organization of Views in a Definition Hierarchy
Conclusions

- The system is work in progress

- Some improvements needed
  - Formalizations of connections between fragments and source knowledge.
    - e-connections? → Manchester
  - Enrichment by text mining techniques
    - Work at EBI: form text to ontologies
    - Draft: http://krono.act.uji.es/people/Ernesto
  - Apply the ontology: evaluation and validation
Questions and Feedback

- Ernesto Jiménez-Ruiz
  - [http://www3.uji.es/~ejimenez](http://www3.uji.es/~ejimenez), [ejimenez@uji.es](mailto:e jimenez@uji.es)

- Resources:
  - Thesis proposal: [http://krono.act.uji.es/people/E rnesto](http://krono.act.uji.es/people/E rnesto)

- Main Contacts:
  - TKBG: [http://krono.act.uji.es](http://krono .act.uji.es) ➔ Rafael Berlanga ([berlanga@uji.es](mailto:berlanga@uji.es))
  - Maat GKnowledge: [http://maat-g.com](http://maat-g.com) ➔ Alfonso Rios ([arios@maat-g.com](mailto:arios@maat-g.com))