Tenth International Protégé Conference, Budapest, Hungary

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


**Ernesto Jiménez-Ruiz**

**Victoria Nebot**

**Rafael Berlanga**

**Ismael Sanz**

**Alfonso Rios**

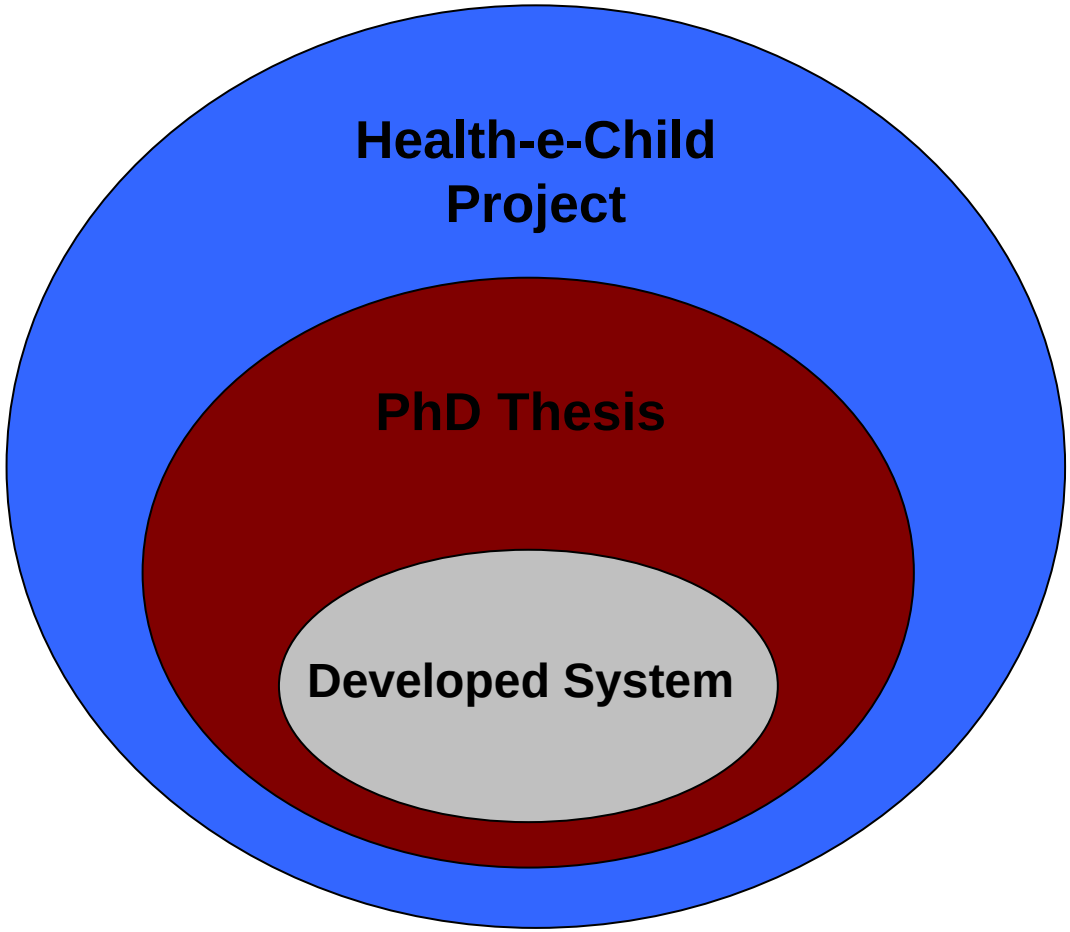
**Jaume I**
**University of Castellon**
**(Spain)**

**Maat Gknowledge,**
**Valencia, (Spain)**

# Outline

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

# Bioinformatics Domain



# Health-e-Child Project

- 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

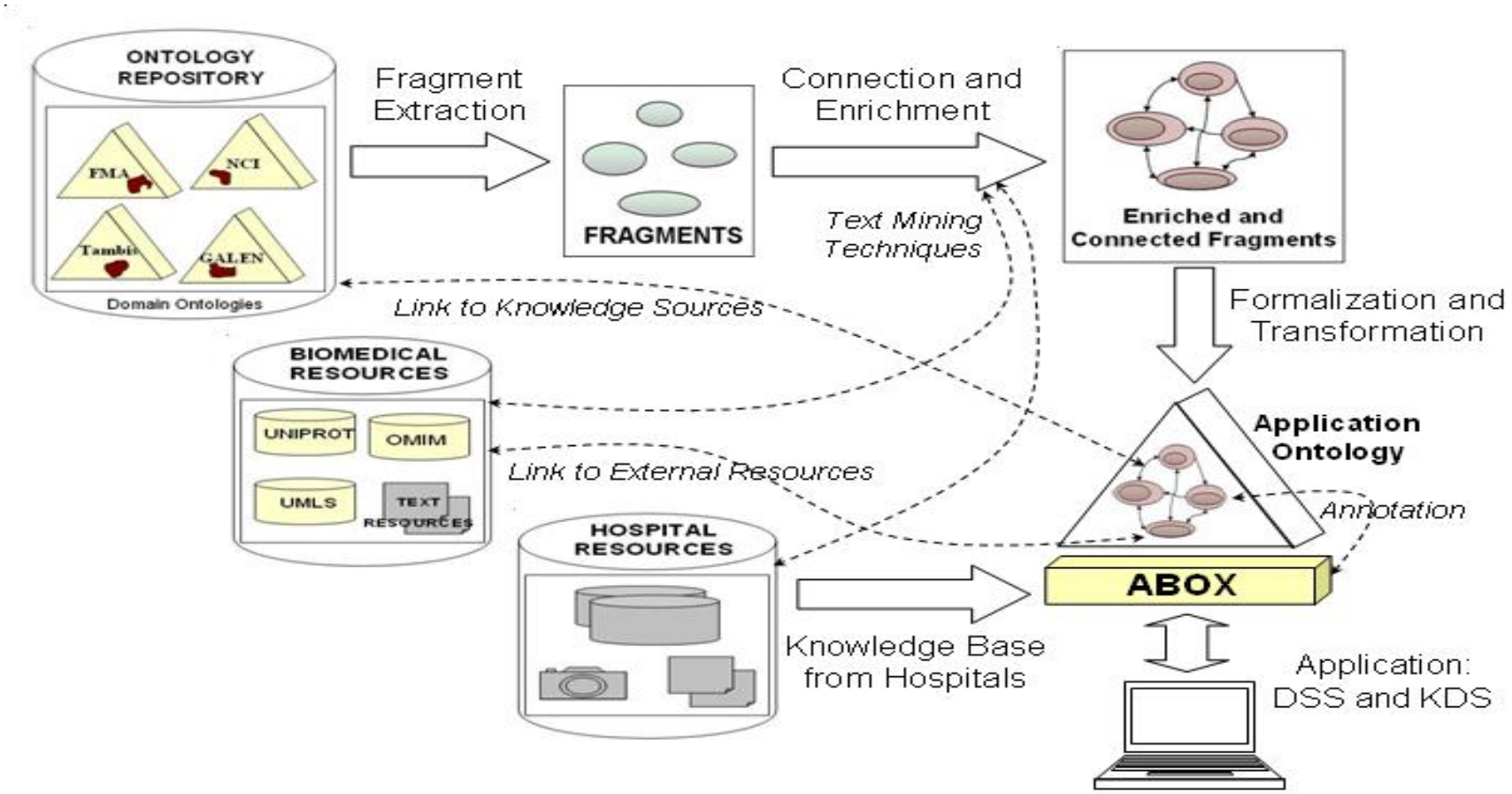
# Health-e-Child Project

- 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

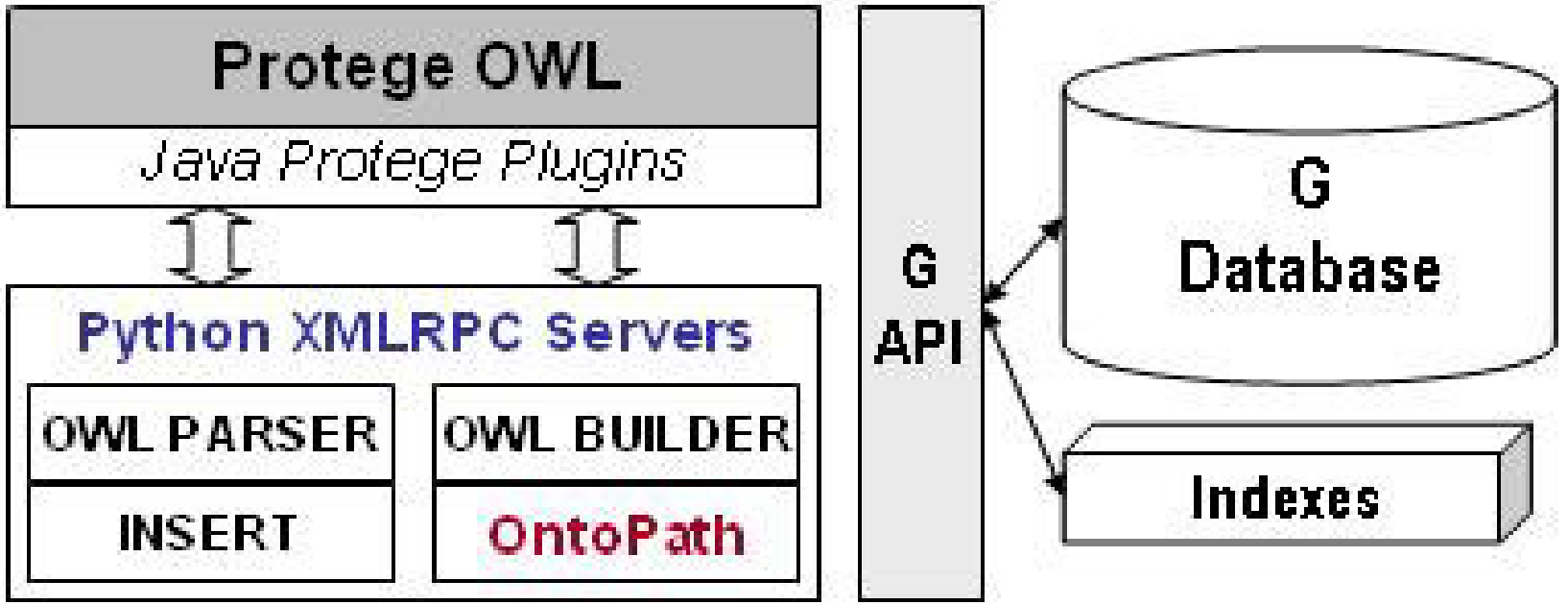
- HeC *vertical abstraction levels expressed* by Ontologies
- Available several large biomedical ontologies and taxonomies, e.g: GO, GALEN, FMA], NCI-Thesurus, Tambis, BioPax[, etc.
- Difficult too apply in concrete applications like HeC:
  - Scalability in reasoning.
  - Specificity: local view of the domain
  - Visualization and treatment

# From Domain Ontologies to Applications (PhD Topic)




Thesis Proposal Available at: http://krono.act.uji.es/people/Ernesto

# System Architecture

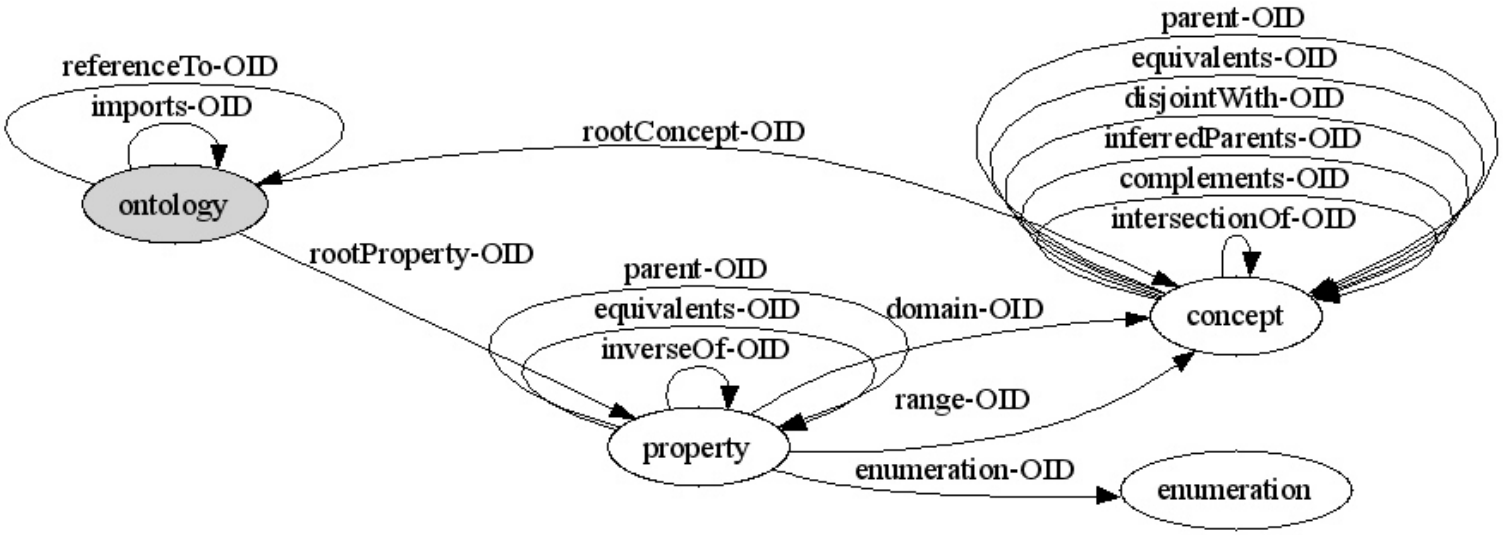


# OWL Parser and Constructor

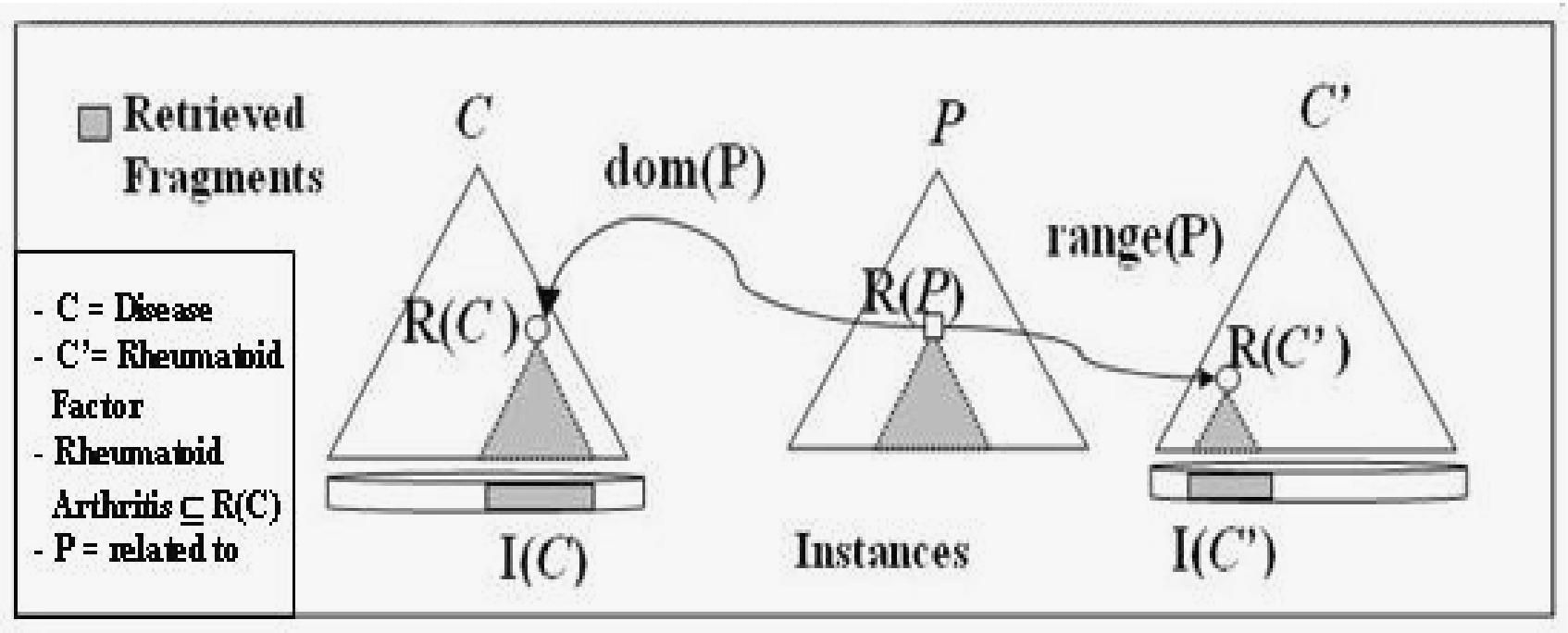
- Greater flexibility in the ***OWL treatment*** and storage capabilities (e.g. indexes)
  - "OntoPath: a Language for Retrieving Ontology Fragments." Submitted to OTM-ODBASE 2007
- The OWL parser creates from the OWL file a set of structures for classes, properties, nominal and individuals.
- These 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**

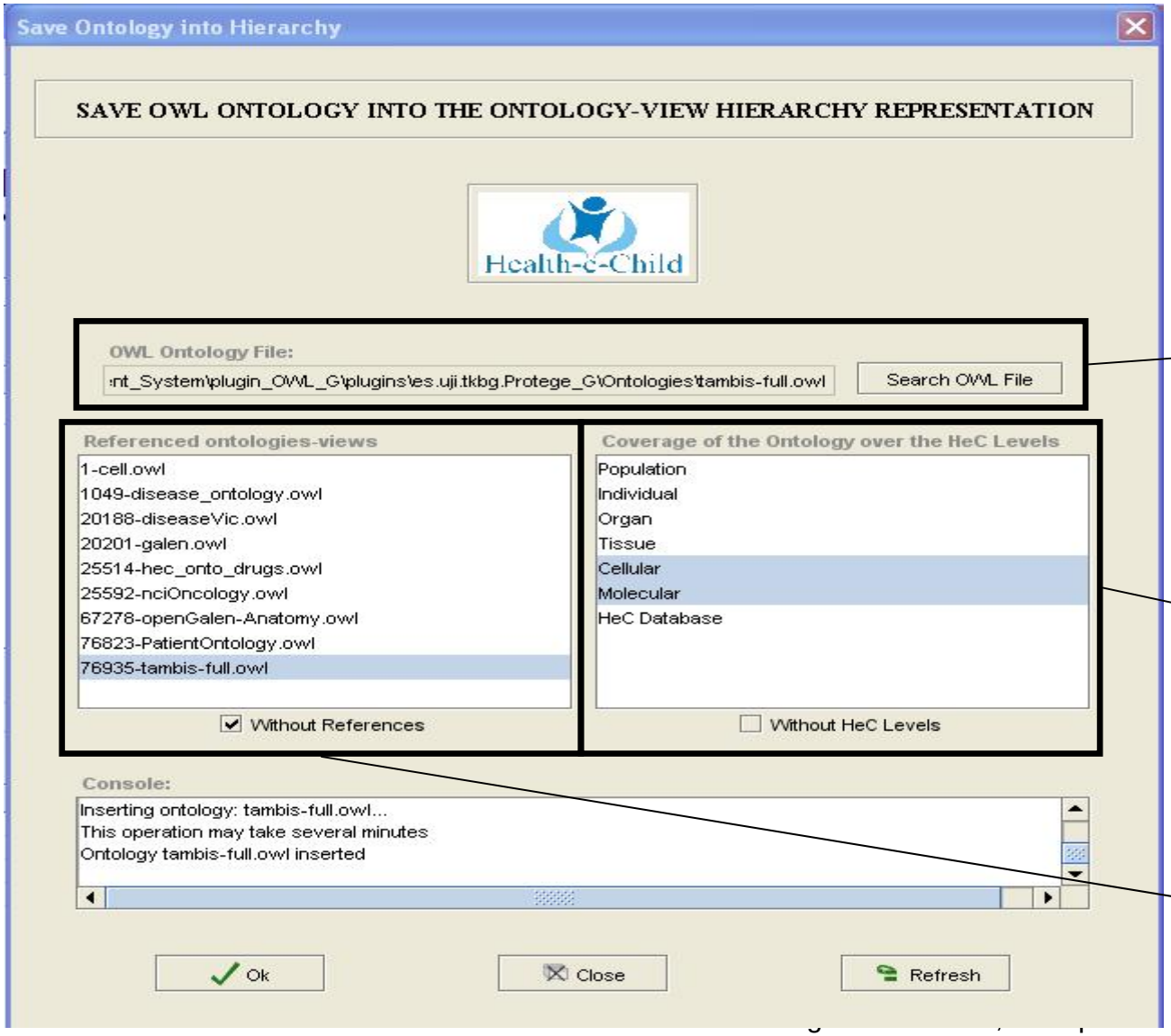
# Ontology Editor Protégé

- 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 defintion hierarchy
- Connection with Python codes

# Storing Ontologies

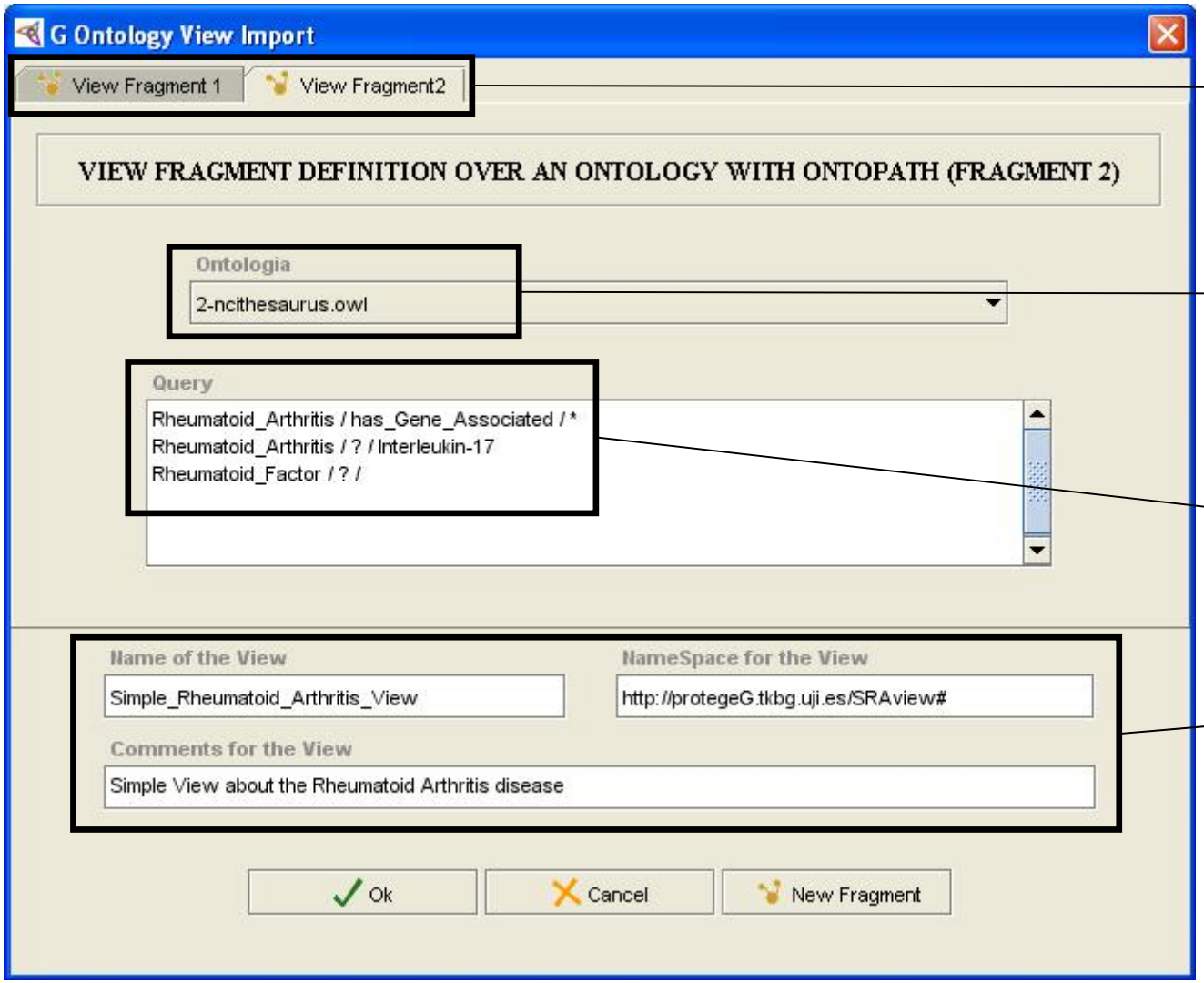



**OWL File Selection**

**Biomedical (HeC) Coverage**

**References to other Ontologies (Views)**

# Retrieving full ontologies or fragments



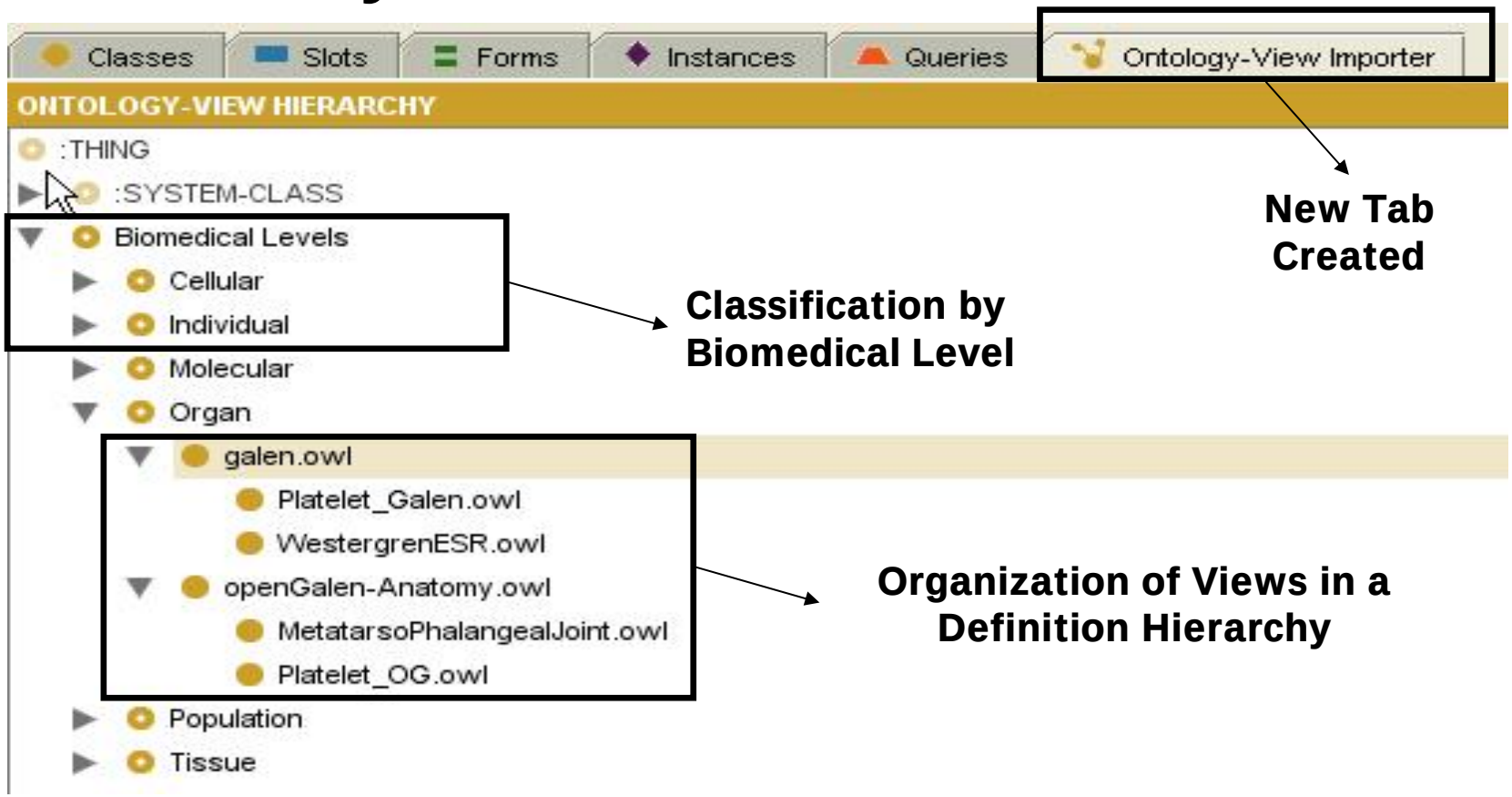

**Several Fragments**

**Source Ontology**

**Set of OntoPath Queries**

**Metadata**

# Representation 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 , ejimenez@uji.es
- Reources:
  - Plug-in (beta): http://krono.act.uji.es/people/Ernesto/G_Protege_Plugin
  - Thesis proposal: http://krono.act.uji.es/people/Ernesto
  - “OntoPath: a Language for Retrieving Ontology Fragments.” Submitted to OTM-ODBASE 2007
- Main Contacts:
  - TKBG: http://krono.act.uji.es → Rafael Berlanga (berlanga@uji.es)
  - Maat GKnowledge: http://maat-g.com → Alfonso Rios (arios@maat-g.com)