Ángel Esteban-Gil, Jesualdo Tomás Fernández-Breis and Martin Boeker. Analysis and visualization of disease courses in a semantic enabled cancer registry
Index

• Cancer registries
• Objectives
• Methods
• Results
• Discussion
• Conclusions
Cancer registries

- What is a cancer registry?
- What information can we find in a cancer registry?
- Cancer registry software.
- Limitations.
Cancer registries
Standards and classification systems

- ICD-10 (International Classification of Diseases)
- SNOMED-CT (Systematized Nomenclature in Medicine – Clinical Terms)
- ICD-O (International Classification of Diseases of Oncology)
  - Morphology
  - Topography
- ICD-10-PCS (Procedure Coding System)
- TNM (Classification of Malignant Tumors)
Cancer registries
Semantic web technologies

• Ontologies
• OWL (Domain level)
• RDF (Data level)
• SPARQL
Objectives

- Design a semantic model for local cancer registry.
- Implement a web platform with semantic technology in a feasibility study.
Methods

• Data transformation and exploitation
  – Semantic transformation engine
  – Ontology-driven searcher (ODS)

• Semantic profiles
  – Disease timeline of a cancer patient
  – Aggregated disease timelines of a group of patients
Methods
Disease timeline of a cancer patient
Methods
Aggregated disease timelines of a group of patients

- **Ontology Driven Search** → **Patient Group**
  - Calculate semantic profile of every patient
  - Aggregated disease timeline of a patient group
    - Calculate
    - Matrix disease evolution
    - Recalculate
    - Matrix therapy
      - Select a concrete therapy
Results
Cancer registry ontology
Results
The semantic cancer registry system
## Results

### Simulated use case

<table>
<thead>
<tr>
<th>Query</th>
<th>SQL count result</th>
<th>SQL time</th>
<th>SPARQL count result</th>
<th>SPARQL time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery all Patients</td>
<td>207.190</td>
<td>0.060s</td>
<td>207.190</td>
<td>0.189s</td>
</tr>
<tr>
<td>Recovery all therapies</td>
<td>400.290</td>
<td>0.132s</td>
<td>400.290</td>
<td>0.317s</td>
</tr>
<tr>
<td>Recovery all diagnosis</td>
<td>240.088</td>
<td>0.070s</td>
<td>240.088</td>
<td>0.220s</td>
</tr>
<tr>
<td>Recovery all courses</td>
<td>108.297</td>
<td>0.030s</td>
<td>108.297</td>
<td>0.155s</td>
</tr>
<tr>
<td>Recovery patients with diagnosis, therapies and courses</td>
<td>207.190</td>
<td>1.048s</td>
<td>207.190</td>
<td>0.204s</td>
</tr>
<tr>
<td>Recovery all female patients</td>
<td>105.714</td>
<td>0.231s</td>
<td>105.714</td>
<td>0.189s</td>
</tr>
<tr>
<td>Recovery all female patients with more of 60 years old</td>
<td>62.603</td>
<td>0.245s</td>
<td>62.603</td>
<td>0.192s</td>
</tr>
</tbody>
</table>
Results
Semantic web platform

• Video demo
  – http://youtu.be/QwL-eI8ilc0
Discussion

Result

• Development of a semantic web platform
  – Representation of the disease course of a patient
  – Representation of the aggregated disease courses of a group of patients
  – Definition of customizable dashboards
Discussion
Relational database vs Semantic datastore

• Our approach provide powerful and precise search capabilities.
• The query editor has been developed guided by OWL.
• Sharing information and comparison of clinical cases and processes.
Discussion
Limitations

• Preliminary version of an ontology of epidemiological cancer registry.
• Lack of real data to test the semantic platform.
• Lack of clinical validation.
Discussion
Related work

- Rule-based systems and logic-based models have been approaches to cancer registries:
  - Analysis of cancer registry processes.
  - Quality assurance
  - Decision support
Discussion
Future work

• We plan to perform a real study with data from a large local cancer registry, which might also include a clinical validation.

• Use our methodology to generate rules that serve to generate patient groups automatically or for quality assurance of the data.
Conclusions

• This work demonstrates that the semantic web can be used for exploiting a local cancer registry.

• The presented platform is an example of the parallel development of ontologies and applications that take advantage of semantic web technologies in the medical field.
This project has been possible thanks to the Spanish Ministry of Science and Innovation and the FEDER program through grant TIN2010-21388-C02-02 and the Fundación Séneca through grant 15295/PI/10.
Q&A time