Recommendation of Standardized Health Learning Contents using Archetypes and Semantic Web Technologies

María del Carmen Legaz-García, Catalina Martínez-Costa, Marcos Menárguez-Tortosa, Jesualdo Tomás Fernández-Breis
Motivation

- The web as the most prominent information source

Problems
- Reliability of the information of such websites?
- More trust on Google than on the physicians?

Our contribution
- Personalized contents for patients and professionals
SemanticHEALTH recommendations

- Adopt a standardized approach for representing and sharing of clinical data structure specifications: agree to use archetypes

- **Linking of tools** to developments in Web 2.0, Social Computing, the **Semantic Web**, Text Mining, and related disciplines in order to implement collaborative web-based workflows

- Linking **EHR data to educational materials** and clinical evidence, to enable consumer engagement and support health professional training

The Semantic Web

- Evolution of the traditional web in which the content can be understood by the machine
- Ontology as domain model that consists of classes, properties, constraints, individuals.
- Bioportal contains more than 200 biomedical ontologies and controlled vocabularies
- OWL is the W3C recommendation for expressing ontologies
Archetypes

- Formal models that represent shareable clinical information in EHR
- Dual-model EHR architectures (openEHR, ISO 13606)
  - Reference Model + Archetypes
  - Archetypes are built by constraining the entities of the reference model
- ADL is the standard language for expressing archetypes
Structure of ADL archetypes

```
archetype (adl_version=1.4)
  archetype_id
[specialise]
  archetype_id
concept
  concept_id
language
  dADL: language details
[description]
  dADL: descriptive meta-data
definition
  cADL: formal constraint definition
ontology
  dADL: terminology and language definitions
[revision_history]
  dADL: history of change audits
```
Archetype Management System (ArchMS)

Archetype

VALIDATION → TRANSFORMATION

ARCHETYPE REPOSITORY

SEARCH METHODS

ANNOTATION

BIOPORTAL

OWL Resources

María del Carmen Legaz-García, Catalina Martínez-Costa, Marcos Menárguez-Tortosa, Jesualdo Tomás Fernández-Breis. Exploitation of ontologies for the management of clinical archetypes in ArchMS, ICBO 2012
Learning Contents

- SCORM learning objects
- Scientific papers from Pubmed
- SICARA
Archetype-based learning content recommendation

- Archetype Repository
- Classification Resources: ontologies, terminologies
- Learning Contents Repository
- Archetype Annotations Repository
- Recommendation Algorithm
- Relevant Learning Contents

The diagram illustrates a process where archetype-based learning content recommendation is facilitated through various resources and repositories, culminating in relevant learning contents.
Semantic similarity of annotations

- Weighted mean of:
  - Linguistic similarity: The similarity between the terms associated with the annotated concepts using a string-based calculation.
  - Taxonomic similarity: The distance between the annotated concepts in the taxonomic structure of the OWL annotation resource.
  - Properties similarity: The similarity between the set of properties associated with the annotated concepts.
Taxonomic similarity

\( TS(A,B) = \frac{\text{common}(A,B)}{\text{sum}(A,B)} \)

Common = 2
Sum = 5
TS = 0.4
Find the contents for the non-TN staging scores for the colorectal cancer archetype
<table>
<thead>
<tr>
<th>Type</th>
<th>Archetype Annotations</th>
<th>Clinical Guideline Annotations</th>
</tr>
</thead>
</table>
| Body structure | • Colon and rectum (combined site)                         | • Appendix structure  
|              |                                                             | • Colon structure  
|              |                                                             | • Rectum part |
| Disorder      | • **Multiple malignancy**                                 | • Adenoma of rectum  
|              |                                                             | • **Adenosquamous carcinoma**  
|              |                                                             | • **Carcinoid tumor**  
|              |                                                             | • Carcinoma in situ of colon  
|              |                                                             | • Carcinoma in situ of rectum  
|              |                                                             | • Squamous cell carcinoma |
Why the resource is selected

Disorder
Neoplastic disease
Malignant neoplastic disease
  - Multiple malignancy
Carcinoid tumor
Squamous cell carcinoma
  - Adenosquamous carcinoma
Conclusions and further work

- Bringing archetypes, semantic web technologies and learning resources together
- Extension in progress to data extracts
- Social computing based recommendation method
- Need for experiments with a larger base of contents
Thank you for your attention!!

Contact:
http://webs.um.es/jfernand
jfernand@um.es

Acknowledgements: Grants TIN2010-21388-C02-02, 15295/PI/10