HealthFinland (Tervesuomi)

Finnish Health Information on the Semantic Web

30 November 2007

Osma Suominen, Kim Viljanen, Eero Hyvönen
Semantic Computing Research Group (SeCo)
Helsinki University of Technology (TKK),
Laboratory of Media Technology
University of Helsinki, Department of Computer Science

http://www.seco.tkk.fi/

Outline of Talk

1. Goals of HealthFinland
2. Challenges
3. Content creation
4. KOS in HealthFinland
5. User interface and navigation
   - demonstration of portal
6. Evaluation of results

HealthFinland portal

- national health promotion portal for citizens [1]
- SW technologies: ontologies, metadata...
  - content aggregated from many websites into a single portal
- diet, exercise, (non)smoking, healthy living...


Goals of HealthFinland: Citizens

1. Global view to health information from different organizations
2. Aggregated view to information from different organizations
3. Semantic search and browsing
   - faceted browsing using user-centric categorizations
   - string search based on ontologies
   - semantic recommendations of related information

Goals of HealthFinland: Publishers

1. Rationalizing content creation by eliminating redundancy
2. Enriching content with other providers' content
3. Automatic and dynamic content linking
4. Reusing global services cost-efficiently as Web 2.0 mash-ups
5. Using centralized ontology services for indexing

Challenges

1. create a compelling user experience for the general public
   - usability is not just a surface feature!
   - solve actual problems users have – in an intuitive way
2. tri-lingual portal and KOS
   - Finnish, Swedish, English
3. interoperable
   - with existing KOS
   - legacy metadata, other document repositories and portals
   - current and future content management systems
4. need to gain a critical mass of information, publishers, users
5. need processes for content creation, quality control, dealing with problems, KOS updates and maintenance
Content Creation Tools

- Content harvester collects metadata from HTML pages and RDF sources
- Metadata validator with feedback reports
- SAHA Annotation editor
- ONKI Ontology Server

Metadata Schema in HealthFinland

<table>
<thead>
<tr>
<th>Source KOS</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>YSO</td>
<td>&gt;20000 concepts in fi, en, sv</td>
</tr>
<tr>
<td>MeSH</td>
<td>&gt;20000 concepts in en, fi, sv</td>
</tr>
<tr>
<td>HPMULTI</td>
<td>1200 concepts in en, fi, sv</td>
</tr>
<tr>
<td>SKOS thesaurus</td>
<td>created by EU project</td>
</tr>
<tr>
<td>SKOS thesaurus</td>
<td>created by US NLM</td>
</tr>
<tr>
<td>General knowledge</td>
<td></td>
</tr>
<tr>
<td>MeSH</td>
<td></td>
</tr>
<tr>
<td>HPMULTI</td>
<td></td>
</tr>
<tr>
<td>SKOS thesaurus</td>
<td></td>
</tr>
<tr>
<td>SKOS thesaurus</td>
<td></td>
</tr>
<tr>
<td>HealthFinland Health Promotion Ontology</td>
<td></td>
</tr>
<tr>
<td>(all stored and maintained on Onki ontology server)</td>
<td></td>
</tr>
</tbody>
</table>

User interface and navigation

- Problem: complex ontologies not suitable for human consumption
- made for a different purpose (e.g. indexing scientific articles)
- expert terminology
- unintuitive hierarchies and groupings
- everything in a big hierarchy – not facet-based

need to build navigation structures that users of the site will understand [2]
solved using an approach [3] based on card sorting


Example of problem

My 70 year old aunt is depressed. How can I find information about mental illnesses in elderly people?

MeSH Tree Structures - 2007
1. Mental disorders [MeSH] Genealogical information [Biological Sciences (J)]
2. Depression [D] Neurology [N]
3. Chemical and Drugs [C]
4. Alcohol and Alcoholism [F]
5. Substance Abuse [I]
6. Psychiatry and Psychology [P]
7. Biomedical Sciences (B)
8. Mental Health [H]
9. Psychology [M]
10. Sociology, Education, Sociology and Social Phenomena [T]
11. Psychology, Education, Sociology and Social Phenomena [T]
12. Information Sources [I]
13. Medical Specialties [M]
14. Health Care [N]
15. Bibliographic Classification [Y]
16. Geographic [Z]

Demonstration of HealthFinland
Solution approach

- Faceted browsing very useful and usable in earlier portals, e.g. Flamenco [4], SWED [5], MuseumFinland [6]...
- So, need to build intuitive facets and categories for the site
- To find out what works for users, do real user research [7]
- When you know what works for users, handle the technicalities
  - Mapping user facets to ontologies

Card sorting

- Method to find out how users conceptualize the information space (i.e. how they group things in their head)
- Often used to build website navigation (information architecture)
  - What should the main sections of the site be
  - How should they be named
  - What things go where
  - What things belong together
- Idea: Print a stack of cards with names of documents, let users sort them into piles and give names to the piles
  - Make notes
- Repeat with several users, try to find common patterns
- Raw output: Sets of labeled piles of cards
  - Easy to do, very effective, enlightening for a designer

Facets and categories

- Each category contains a set of concepts
- Category taxonomy represented using SKOS vocabulary
  - Labels: skos:prefLabel
  - Hierarchy: skos:broader
- Mappings to ontology using SKOS Mapping
  - skosmap:exactMatch, skosmap:narrowMatch
- Result: Categories contain documents
  - Annotated with concepts contained in the category
  - Or subconcepts (rdfs:subClassOf or skos:broader)
  - From all subcategories, recursively
  - Creating a proper subsumption hierarchy

Example mappings

- Topic: Weight Control
  - Losing Weight
  - Nutrition & Food
  - Diet

- Group of People: Age Group
  - Senior
  - Baby, Infant

- Life Event: Goals (e.g., pregnancy, each pregnancy)
  - Concept 1: Pregnancy

- Body Part: Bone
  - Bone

- Ontological concepts: Body Weight
  - Weight Loss
  - Energy Intake
**Evaluation**

- Closed card sorting session to test intuitiveness of facets
  - Given the final facets, do test subjects place a set of concepts in the intended categories?
  - Promising results (but only 2 subjects)
- Review of facets by domain experts: some problems corrected
  - Too much lumping gives wrong message to users
  - Omission of important topics
- Prototype portal to test whether the approach can actually be implemented
- Currently performing 3rd phase user testing with prototype
  - So far the problems found have not concerned the categorizations
  - More user tests under way (in the following 2 weeks)
  - Results will be published in academic conferences, journals etc.
- Prototype will go live in early 2008, production use in 2008-2009

---

**Thanks!**

Osma Suominen

http://www.seco.tkk.fi

Thanks!

Hyvönen et al: HealthFinland - Finnish Health Information on the Semantic Web

Suominen et al: User-centric faceted search for semantic portals
Proc. ESWC 2007, Innsbruck, Austria, Jun 2007