Ontology Engineering
Seminar 2010 / 2011
Oct 12, 2010
Katharina Siorpaes

Agenda

• Seminar organization
• Seminar topic assignment
• Folksonomies as lightweight ontologies
• FOLCOM cost estimation overview
• FOLCOM experiment

Next session: Dec 10, 2010
Detailed agenda TBA.
Course overview

- "An ontology is a formal representation of a shared conceptualization of a domain." Gruber, 1999.
- This seminar covers topics related to ontology engineering and the process of semantic content creation.
  - ontology building methodology
  - ontology matching
  - ontology evaluation
  - ontology learning
  - incentives for ontology building
  - technologies and tools for building ontologies
  - collaborative ontology building
  - cost estimation of ontologies

Overall learning goal

- Get an overview of ontology engineering.
- Dive into a specific topic in detail.
- Small, hands-on ontology engineering project.
  Also:
  - Presentation skills
  - Working in teams
  - Scientific writing
Administrative issues

• The course is structured in 3 blocks.
  – Di 12.10.2010 14.00 - 16.00 3W03
  – Fr 10.12.2010 08.00 - 18.00 3W03
  – Fr 28.01.2011 08.00 - 18.00 3W03
• First session: experiment, topic assignment.
• Second session: presentation of theoretical topics, submission of short papers, ontology project assignment.
• Third session: ontology project assignment final presentations (teams).
• Attendance mandatory.
• Grades: participation in three sessions, papers, presentations, and final projects.

About the lecturer

Dr. Katharina Siorpaes
University of Innsbruck
Technikerstr. 21a
6020 Innsbruck, Austria

e: katharina.siorpaes@sti2.at

Office hours: only by appointment. Please contact me per e-mail at katharina.siorpaes@sti2.at
Resources

• Course home page
  http://www.sti-innsbruck.at/teaching

• Please subscribe to the mailinglist! (see website)

First assignment

• Language: English
• Topic + state of the art survey about the topic.
• Paper
• 20 minutes presentation
• Grade requirements: successful participation assessment
  (a presentation and a written seminar paper)
First assignment

• Presentation
  – Students need to submit their presentation one week in advance to the final seminar presentation
  – Send per e-mail

• Written Paper
  – Students need to prepare 5 pages paper
  – Papers have to formatted in Springer LNCS format:
    http://www.springer.com/computer/lncs?SGWID=0-164-7-72376-0
    Hard deadline: 5.12.2010 – no extensions

First assignment

• Work in teams of 2
• Start to work immediately.
• Do not go to the presentation without having understood what you are presenting.
• This is a seminar where you should do scientific work.
• If you have questions / problems – do not hesitate to ask me.
• DO NOT copy! PLAGIARISM IS NOT ACCEPTABLE!
• Whenever you cite somebody, cite properly.
• References will be part of your work!
• When you use work of somebody else, re-phrase and summarize in your own words and cite properly.
• Scientific writing:
## First assignment - topics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Collaborative ontology building methodologies (Klotz)</td>
</tr>
<tr>
<td>2</td>
<td>Ontology matching methods (Tuymaniuk)</td>
</tr>
<tr>
<td>3</td>
<td>Ontology evaluation methods (Häser)</td>
</tr>
<tr>
<td>4</td>
<td>Ontology learning methods (Thaler)</td>
</tr>
<tr>
<td>5</td>
<td>Ontology building tools (Fried)</td>
</tr>
<tr>
<td>6</td>
<td>Ontology building languages (Neuruhrer)</td>
</tr>
<tr>
<td>7</td>
<td>Cost estimation of ontologies (Pudenz)</td>
</tr>
<tr>
<td>8</td>
<td>The role of ontologies in Linked Data and interlinking methods (Wach)</td>
</tr>
<tr>
<td>9</td>
<td>Access to Linked data</td>
</tr>
<tr>
<td>10</td>
<td>Multimedia ontologies (Scheiber)</td>
</tr>
<tr>
<td>11</td>
<td>Ontology visualization methods (Wieser)</td>
</tr>
<tr>
<td>12</td>
<td>Social Semantic Web: using Semantic Wikis, games, etc. for building ontologies (Dobler)</td>
</tr>
</tbody>
</table>

## First assignment

- Form a team of 2
- Send me an email with
  - The topic assigned
  - The members of your team (Names + email addresses)
- Literature pointers will be provided but you should independently do literature research yourself!
- Once you have a list of sources, please drop me an email to get feedback on the sources.
On ontologies

- Slides on „Ontologies“ of Semantic Web lecture (see STI website).
- Fabien Gandon: Ontologies in a nutshell (http://www.slideshare.net/fabien_gandon/presentations) - more interesting presentations.

Tagging

- Idea: Enrich contents by user chosen keywords
- Replace folder based structure by a organisation using tags
- New: Simple user interfaces for tagging and tag based search
- First steps to Semantic Web?
- Technically: user interfaces
- **Social: collaborative** (own contents, shared tags)
Tagging: Flickr.com

Collaborative Tagging
Collaborative Tagging: Delicious

- Browser plug-ins available from http://del.icio.us
- Allows the tagging of bookmarks
- Community aspect:
  - Suggestion of tags that were used by other users
  - Availability of tag clouds for bookmarks of the whole community
  - Possibility to browse related bookmarks based on tags

Folksonomies

Data created by tagging, knowledge structures

Mary tags www.wikipedia.org with wiki wikipedia encyclopedia
Folksonomies: Taxonomy  Marlow et al. (2006)

- Rights for Tagging
  - Self-tagging: Contents only tagged by owner (Technorati)
  - Free-for-all tagging: Tagging by all users (Yahoo!)

- Support of Tagging
  - Blind Tagging: Existing Tags are not displayed (Flickr)
  - Viewable Tagging: Existing Tags are displayed (Del.icio.us)
  - Suggestive Tagging: Suggestions for Tags (MyWeb 2.0)

- Aggregation of Tags
  - Bag-model: Multiple entries (Del.icio.us)
  - Set-model: Only single entries (YouTube)

Tag Clouds

- Size of Tags: count of usage
- Browsing replaces Searching
- Different meaning for different users
- Orientation in Information Set
**FOLCOM: Folksonomy Cost Estimation Model**

- **Challenges:**
  - Folksonomy has significant different characteristics from other knowledge structures.
  - The life cycle of folksonomy does not fall in any of the classical engineering process models.
  - Tailor-made folksonomy cost estimation methods are not available.

- **Goal:**
  - FOLCOM: a reliable cost estimation model for folksonomy.

- **Approach:**
  - **FOLCOM:** Folksonomy Cost Estimation Model

**FOLCOM Process**

- **Requirement Analysis**
  - Number of objects, types of the objects.
  - Knowledge of how often a single object can be tagged.
  - Possibility to measure tagging times from users.
  - Stable tagging interface during project.

- **Size Estimation**
  - Story points $SP_{est}$: the estimated size of tagging work expected from Collection:=${O_1, O_2, \ldots, O_n}$
  - $SP_{est} = \sum_{o \in O} (\text{complexity}(o))$ or $SP_{est} = \sum_{o \in O} (\text{complexity}(o) \times |o|)$

- **Velocity Determination**
  - Tagging Time $\text{taggingTime}$: the measured time a user required for tagging object $o$.
  - Total tagging effort $\text{totalEffort}_t$: sum of all tagging time $\text{totalEffort}_t = \sum \text{taggingTime}$.
  - Velocity $\text{velocity}$: the effort value to tag a story point $\text{velocity} = \frac{\text{totalEffort}_t}{|SPs|}$.

- **Effort Estimation**
  - Effort estimation $\text{effortEstimation}$: the cost of the tagging project $\text{effortEstimation} = \text{multiTagFactor} \times (SP_{est} - SPs) \times \text{velocity}$.
FOLCOM: Evaluation and Experiments

- Evaluation
  - Evaluation of the feasibility of FOLCOM with 10 criteria by a team of experts.
  - Reliable and feasible model for folksonomy.

- Experiments
  - Test and evaluate FOLCOM model with 200 different images
  - Online test using a customized folksonomy tool

The experiments show that the estimation accuracy of FOLCOM is sufficient.

### Experiment Results

<table>
<thead>
<tr>
<th>Exp ID</th>
<th>Actual total effort</th>
<th>Estimate (50 samples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta</td>
<td>39.05 min.</td>
<td>38.28 min.</td>
</tr>
<tr>
<td>g1</td>
<td>71.27 min.</td>
<td>71.33 min.</td>
</tr>
<tr>
<td>ocl</td>
<td>60.77 min.</td>
<td>70.11 min.</td>
</tr>
<tr>
<td>mm1</td>
<td>63.33 min.</td>
<td>92.11 min.</td>
</tr>
<tr>
<td>ps2</td>
<td>52.22 min.</td>
<td>57.00 min.</td>
</tr>
<tr>
<td>ps1</td>
<td>40.92 min.</td>
<td>48.00 min.</td>
</tr>
</tbody>
</table>

### Error in %

- [Image of error graph]
FOLCOM tagging efficiency monitor

Thank You!

Questions?