

HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

#### **Opas - An Ontology-based Library Help Desk Service**

Olli Alm, Antti Vehviläinen Semantic Computing Research Group (SeCo) Helsinki University of Technology (TKK), Laboratory of Media Technology and University of Helsinki, Department of Computer Science

http://www.seco.tkk.fi





helsinki university of technology Media Technology

## An Ontology-based Library Help Desk Service

- Developed by Antti Vehviläinen
- Based on an existing the national Ask a Librarian service of the Kirjastot.fi portal
- 20 000 textual question-answer (QA) pairs
- A common Finnish thesaurus (YSA) is used to annotate QA pairs with index terms





# Why Opas?



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

### To ease the indexing task:

- Choosing appropriate index terms can be difficult and time-consuming
- Different librarians have different conventions in choosing index terms.

Solution:

**Semi-Automatic Semantic Indexing** 



# Why Opas?



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

## To find similar Question-Answer pairs:

- Questions tend to repeat themselves
- Librarians don't have a systematic
- way to search the QA-archive

Solution:

**Retrieval of similar QA-cases** 





# Why Opas?



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

## To support background information:

- Background information is needed
  - for the answering
- Integrated sources may be suggested automatically based on the content of a question.

Solution:

**Integration of Information Sources** 





## Solutions



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

### Semi-Automatic Semantic Indexing:

- OPAS uses an ontology-based information extraction tool to suggest possible indexing terms
- Suggested terms are ordered on likely relevance:
  Semantic cluster --> more weight
  High term frequency --> more weight



## Solutions



#### **Retrieval of similar QA-cases**

- Opas applies case-based reasoning techniques to find QA pairs similar to a new question
- The search is based on the list of suggested index terms for the question



# Solutions



HELSINKI UNIVERSITY OF TECHNOLOGY Media Technology

## **Integration of Information Sources**

Library system classification ontology (HCLCS) is utilized to

- find classifications
- find books from the library collection
- find essential links from the library's link collection



