Towards Models for Judging the Maturity of Enterprises for Semantics

Marek Nekvasil, Vojtěch Svátek
Dpt. of Knowledge Engineering
University of Economics Prague

Agenda

- Motivation
- Categorization of Semantic Applications
  - Criteria and results
  - Possible Archetypes
- Critical Success Factors
- Future Work – Maturity models

Agenda

- Motivation
- Categorization of Semantic Applications
  - Criteria and results
  - Possible Archetypes
- Critical Success Factors
- Future Work – Maturity models

Motivation

- Semantic technologies have been included in broader and broader areas of application deployment
- Differences amongst them are vast
- The investments are uncertain
- -> The business segment is sceptical

What are Semantic Applications?

- There is no common definition
  - e.g.: *any application that stores data separately from the meaning and content files can be called* ***semantic application***
    - *S. Staab*
- Semantic applications can not be considered all at once
- The general scepticism is unjustified

Agenda

- Motivation
- Categorization of Semantic Applications
  - Criteria and results
  - Possible Archetypes
- Critical Success Factors
- Future Work – Maturity models

Categorization Criteria

- Before judging a particular semantic application it has to be categorized so that the evaluation process can be standardized
- The categorization has to be multidimensional
- We came up with a set of dimensions - criteria

General Criteria

- Domain-specificity and reusability
  - (domain independent, domain specific, case specific)
- Number and kind of users
  - (1-10,10-100, >100; experts/managers/public/...)
- User × provider relationship
  - (matters of financing and outsourcing)
- Frequency of access to the application and its availability
  - (once, regularly, irregularly, 24/7)

Specific criteria

- Information sources
  - (structured knowledge, structured data, unstructured data)
- Data source provenance
  - (data are created manually/ automatically / arise elsewhere)
- Accuracy of inputs and outputs
  - (full, partial, concerning uncertainty)
- Subject of operation
  - (data indexing, data integration, inference)

Considered Applications

- Catalogue of semantic applications by W3C SWEO interest group
  - www.w3.org/2001/sw/sweo/public/UseCases/
- Contains description of 25 Case Studies and 12 Use Cases
  - -> constantly growing

SWEO catalogue

Case study (25)

1. A Digital Music Archive (DMA) for the Norwegian National Broadcaster (NRK) using Semantic Web techniques (Case study), by Robert Engels and Jon Roar Tønnesen, ESIS and NRK (Norway)
   Activity area: broadcasting
   Application area of SW technologies: improved search, content discovery, and data integration
   SW technologies used: RDF(S), OWL, SPARQL, and in-house vocabularies
   SW technology benefits: improved search, identify new relationships, and share and re-use data
2. A Linked Open Data Resource List Management Tool for Undergraduate Students (Case study), by Chris Clarke, Talis Information Limited and University of Plymouth (United Kingdom)
   Activity area: education, learning technology, and publishing
   Application area of SW technologies: content discovery, content management, data integration, and semantic annotation
   SW technologies used: RDF, RDFa, SPARQL, RDF(S), SKOS, public datasets, and public vocabularies
   SW technology benefits: explicit content relationships, personalization, reduced time to market, and share and re-use data
3. A Semantic Web Content Repository for Clinical Research (Case study), by Chimezie Ogbuji, Eugene Blackstone, and Chris Pierce, Cleveland Clinic (United States)
   Activity area: health care and public institution
   Application area of SW technologies: data integration
   SW technologies used: RDF(S), OWL, GRDDL, Rules, Rules (N3), and public vocabularies
   SW technology benefits: automation, incremental modeling, and improved search

2 publishing
1 search
1 semantic desktop
4 telecommunications
1 utilities

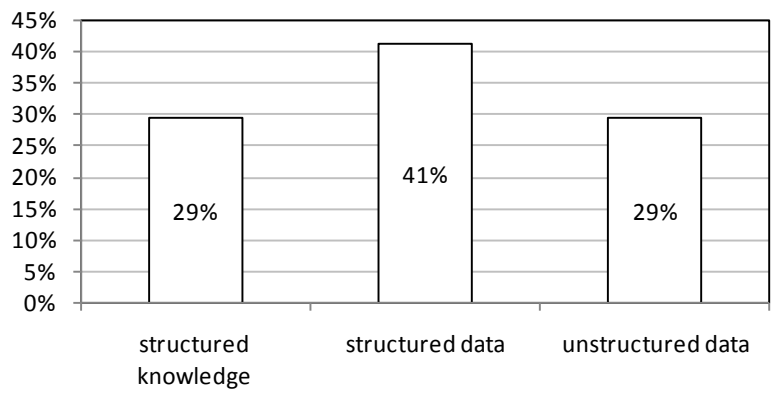
Application area of SW technologies

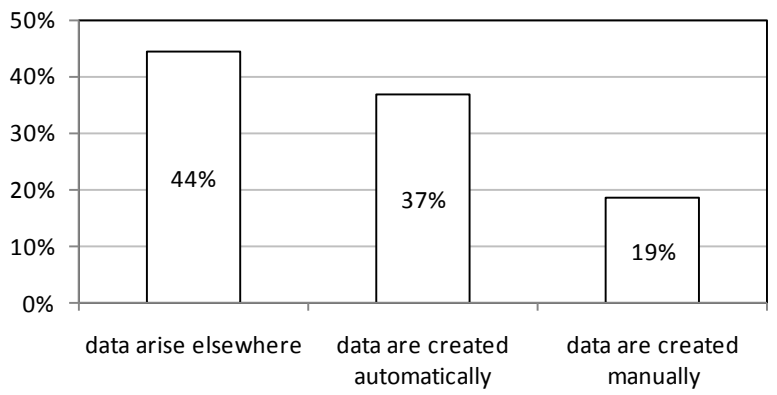
10 content discovery
2 content management
2 customization
24 data integration
1 domain modeling
14 improved search

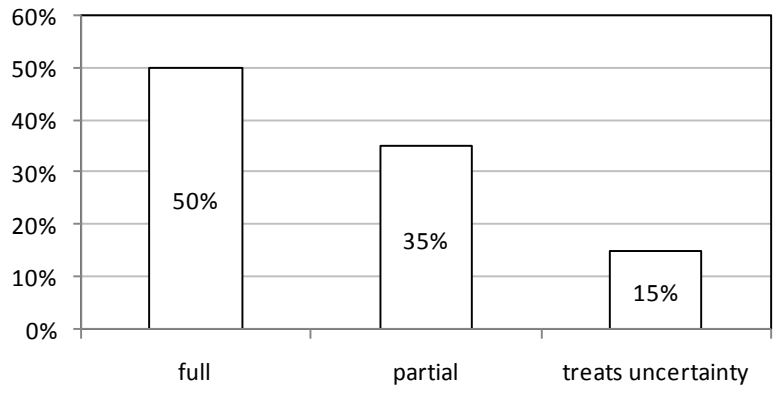
SW technologies used

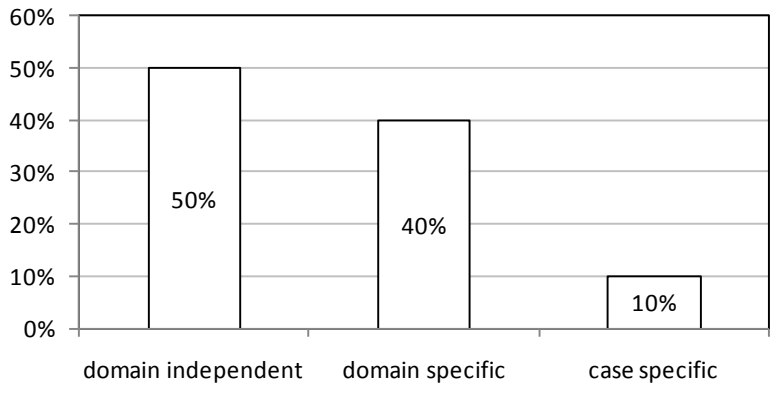
2 GRDDL
18 OWL
1 OWL DL
1 OWL-S
2 RDF

Results of Categorization pt. 1

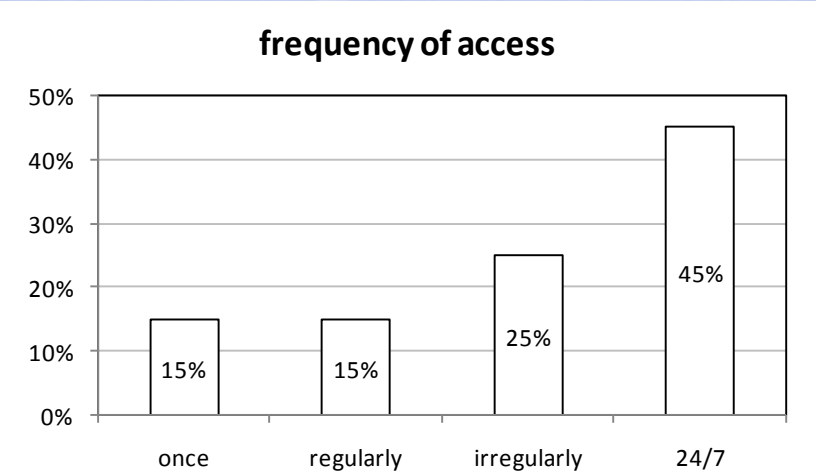


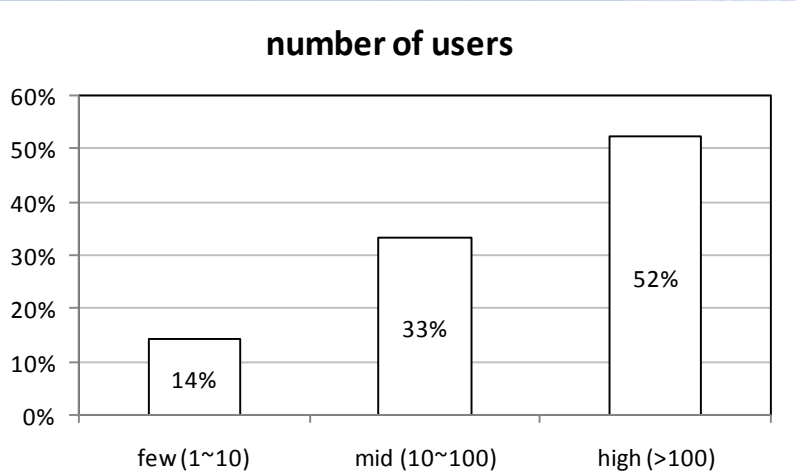


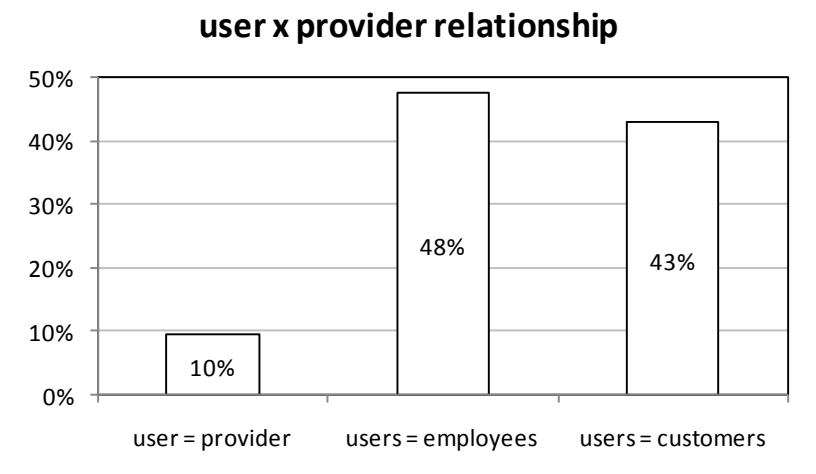


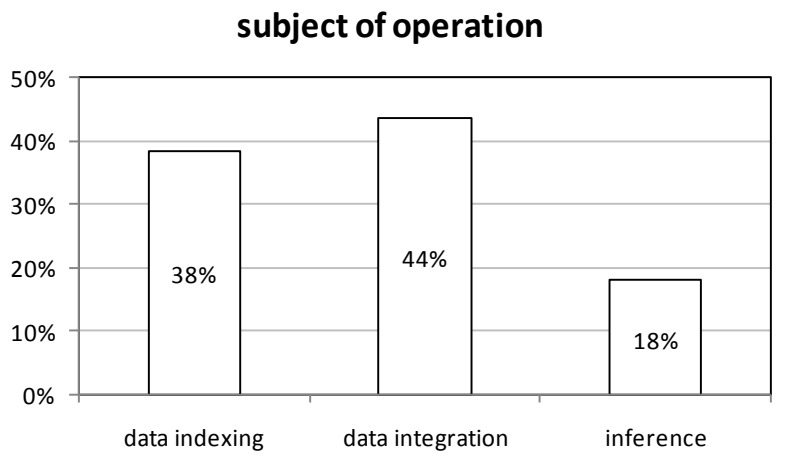


Results of Categorization pt. 2









Agenda

- Motivation
- Categorization of Semantic Applications
  - Criteria and results
  - Possible Archetypes
- Critical Success Factors
- Future Work – Maturity models

Typical applications?

- Once we have a database of classified applications, the next logical step is performing a cluster analysis
  - -> future work
- However, we have a first guess

Possible Archetypes

- “Improved search engine”
- “Data-browsing interface”
- “Recommending system”
- “Data interchange framework”

Agenda

- Motivation
- Categorization of Semantic Applications
  - Criteria and results
  - Possible Archetypes
- Critical Success Factors
- Future Work – Maturity models

Critical Success Factors

- By the synthesis of the risks mentioned in individual SWEO case studies, we identified many CSFs
- 1) General CSFs, including:
  - CSFs of strategy and planning
  - CSFs of disponible resources
  - CSFs of deployment quality
  - CSFs of implementation
  - CSFs of maintenance
- 2) CSFs according to categorization

Specific CSFs

- Correctness of the core ontology/taxonomy
- Sufficiently steep learning curve of end-users
- The potential of possible benefits to compensate the temporary reduction in productivity during implementation and learning
- Will and discipline of all parties to use the same knowledge model

Specific CSFs (cont'd)

- Synchronized distribution of central ontology
- Sufficient number of users
- Users' motivation
- Sufficient supply of data
- Diversity of sources and forms of data
- Maintaining at least the same accuracy of results as the sub-systems
- Reliability of parsers and wrappers

Agenda

- Motivation
- Categorization of Semantic Applications
  - Criteria and results
  - Possible Archetypes
- Critical Success Factors
- Future Work – Maturity models

Future work

- 1) Identify application archetypes based on sofisticated cluster analysis
- 2) Evaluate the Critical Success Factors (via questioning the management of existing successful applications)
- 3) Establish enterprise maturity models for the deployment of a certain type of semantic technologies (starting from the archetypes) based on the aspects of categorization and the associated critical success factors

Maturity model example

- The Maturity Models will be exemplar, without a linked content
- Example for the Archetype of „Semantic search engine“:
  - *If an enterprise uses a single source of data and a proprietary data structure, then it is unprepared for the introduction of this kind of system. If it is using multiple systems with heterogeneous data structure, the introduction of search engines with semantic indexing can bring some improvements to the search results. The enterprise achieves next level of readiness if it uses more systems with a standardized data structure; in such case it can start thinking about the integration of these systems with semantic data exchange, etc.*

Thank you for your attention

Any questions?