E-D11

The structure of the Semantic Web Topics Progress Report and Applications

Project title:	Reasoning on the Web with Rules and Semantics
Project acronym:	REWERSE
Project number:	IST-2004-506779
Project instrument:	EU FP6 Network of Excellence (NoE)
Project thematic priority:	Priority 2: Information Society Technologies (IST)
Document type:	D (deliverable)
Nature of document:	R (report)
Dissemination level:	PU (public)
Document number:	IST506779/Linköping/E-D11/D/PU/a0
Responsible editors:	Joerg Diederich, Jose Alferes, Jan Maluszynski
Reviewers:	Uta Schwertel, Tim Geisler
Contributing participants:	Linkoeping, Hannover, Lisbon, Munich
Contributing workpackages:	ET, TTA
Contractual date of deliverable:	31 August 2007
Actual submission date:	15 September 2007

Abstract
This document presents a revised proposal for structuring the body of knowledge in the field of Semantic Web and reports on its use. The presented taxonomy emerged from the analysis of existing courses and from discussions within REWERSE and within Knowledge Web. The taxonomy was already used to develop a joint curriculum for the European Academy for Semantic-Web Education. This Academy was initiated by Knowledge Web as an international joint masters program. This document reports on the courses offered within this program by the REWERSE participant Universidade Nova de Lisboa. It also reports on the use of the taxonomy for classification of e-learning resources in REASE (the repository of the European Association for Semantic Web Education).

Keyword List
Semantic Web, education and training

Project co-funded by the European Commission and the Swiss Federal Office for Education and Science within the Sixth Framework Programme.

The structure of the Semantic Web Topics Progress Report and Applications

Jan Maluszynski[1] **Jörg Diederich**[2] **Jose Alferes**[3]
Norbert Eisinger[4]

[1] Department of Computer Science, Linköping University, Sweden
{janma@ida.liu.se
[2] L3S Research Center, Hannover, Germany
diederich@l3s.de
[3] Departmento de Informática, Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa, Portugal
jja@di.fct.unl.pt
[4] Ludwig-Maximilians-Universität München, Germany
Norbert.Eisinger@ifi.lmu.de

15 September 2007

Abstract
This document presents a revised proposal for structuring the body of knowledge in the field of Semantic Web and reports on its use. The presented taxonomy emerged from the analysis of existing courses and from discussions within REWERSE and within Knowledge Web. The taxonomy was already used to develop a joint curriculum for the European Academy for Semantic-Web Education. This Academy was initiated by Knowledge Web as an international joint masters program. This document reports on the courses offered within this program by the REWERSE participant Universidade Nova de Lisboa. It also reports on the use of the taxonomy for classification of e-learning resources in REASE (the repository of the European Association for Semantic Web Education).

Keyword List
Semantic Web, education and training

Contents

1 Introduction

This document reports on the actual state of the REWERSE work aiming at defining and implementing Semantic Web graduate curricula.

It presents a revised proposal for structuring the body of knowledge in the field of Semantic Web (known as the Semantic Web topic hierachy and available in the latest version at `http://ontoworld.org/wiki/Semantic_Web_Topic_Hierarchy`) and reports on its use

- for development of the joint curriculum for a shared masters program in the European Academy for Semantic-Web Education. `http://www.semantic-web-academy.eu/`
- for development of the program of Semantic Web courses at Universidade Nova de Lisboa as their contribution to the
- for development of a catalogue of e-learning resources in the joint Knowledge Web and REWERSE educational infrastructure REASE `http://rease.semanticweb.org`

To ensure that our work has a sustaining impact on future developments of the topic hierarchy, we have documented all our efforts regarding the development and the history of the topic hierarchy in a separate wiki page, available at `http://ontoworld.org/wiki/History_of_the_Semantic_Web_Topic_Hierarchy`. The following section is a copy of that wiki page.

2 History of the Semantic Web Topic Hierarchy

2.1 General motivation

The Semantic Web Topic Hierarchy was created to merge several different collections of topics into a single structure. In summary, it was started from the following activities:

- The European Network of Excellence REWERSE started to develop a graduate curriculum for Semantic Web education in September 2004.
- The European Network of Excellence Knowledge Web continued the discussion about the topic hierarchy in June 2005. This was related to close cooperation with REWERSE in setting up REASE, a repository for Semantic Web learning resources. In REASE the topic hierarchy is used as a basis for the REASE catalogue.
- The International series of Semantic Web conferences (ISWC) and the European series (ESWC) was using a couple of topics for the classification of their papers into categories, which were used for the assignment of paper submissions to reviewers as well as for naming sessions during the conference.
- In May 2007, Knowledge Web decided to use this wiki as main dissemination channel for the project results.

Of course, the topic hierarchy is by no means perfect and reflects the subjective opinion of the involved researchers about the field 'Semantic Web'. It has undergone several changes and versions, as more people started to contribute to the discussion of the topic hierarchy and also because the 'Semantic Web' research area developed over time during the years 2004–2007, when the topic hierarchy was initiated.

This document is intended to capture the rationales for the specific changes made to the topic hierachy and provide pointers to the documents where the changes were introduced.

2.2 History of Deliverables in REWERSE and Knowledge Web relevant for the topic hierarchy

2.2.1 REWERSE E-D1: Collecting existing university courses about Semantic Web topics (Sept. 2004)

To initiate the discussion about developing a Semantic Web graduate curriculum, existing university courses about Semantic Web topics were collected jointly by Knowledge Web and REWERSE in the first months of 2004 (cf. REWERSE deliverable E-D1). As a result, four main categories were identified:

- Foundations (relevant to Semantic Web), such as databases, logics, knowledge representations, and logic programming
- Basic Web Information technology, such as XML, web data integration
- Semantic Web concepts and techniques, such as RDF, OWL, and rules
- Other related techniques, such as agents, security, business rules, or knowledge management

It became clear that a better structuring of the field of Semantic Web is required for a final Semantic Web curriculum.

2.2.2 REWERSE E-D5: REWERSE Graduate courses (Feb. 2005)

In REWERSE deliverable E-D5, the above list of main categories was aligned with the ACM CC2001 curriculum, which includes some foundational aspects relevant for a Semantic Web curriculum such as

- IS (intelligent systems)
  - IS3 Knowledge Representation and Reasoning (core)
  - IS6 Agents (elective)
  - IS7 Natural Language Processing (elective)
- IM (information management)
  - IM2 Database systems (core)
  - IM3 Data modelling (core)
  - IM12 Hypertext and Hypermedia (elective)

However, many important concepts such as basic web information technologies and the specific Semantic Web concepts were missing in the ACM curriculum. Hence, the following preliminary structure of the topic hierarchy was presented (listed here because it's not available in the history of the page for the Semantic Web Topic Hierarchy).

- IS Intelligent Systems
  - IS1 Knowledge Representation and Reasoning

        * Logics (Predicate Logic, Description Logics, Horn Logic, F-logic, Modal Logics)
        * Basics of automated reasoning
        * Logic Programming and Non-monotonic Reasoning
        * Reasoning on Action and Change
        * Temporal and Spatial Reasoning
    - IS2 Agents
    - IS3 Natural Language Processing
- IM Information Management
    - IM1 Data Modeling
        * Conceptual models; ontologies, UML
        * Relational data model
        * Semistructured data
        * Object-oriented model
    - IM2 Database systems
        * Relational databases
        * Object-oriented databases
        * Distributed databases
    - IM3 Hypertext and Hypermedia
- WT Basic Web information technologies
    - WT1 XML
        * Namespaces
        * Schema languages
        * XML query and transformation languages
        * XML programming techniques
    - WT2 Web data integration
    - WT3 Security
    - WT4 Web services
    - WT5 Personalization techniques
- SW Semantic Web
    - SW1 Basic ideas of the Semantic Web
    - SW2 Resource Description Framework
        * RDF basics
        * RDF Schema
        * RDF Query Languages
    - SW3 Ontologies
        * Web Ontology languages

    - Ontology Engineering
- SW4 Rules on the Semantic Web
- SW5 Semantic Web Applications
    - Web Services
    - e-Learning
    - Bioinformatics
    - Other

2.2.3 Knowledge Web deliverable D3.1.5: REASE

Starting from E-D5, a discussion was initiated on the Knowledge Web education area mailing list and the REWERSE education and training mailing list in June 2005 with the special focus to extend the original REASE catalogue from September 2004 as described in the Knowledge Web deliverable D3.3.2v2. As a result, Version 1.0 of the Semantic Web Topic Hierarchy was created extending the previous proposal with several new main topics and refining existing topics. Specifically, it included also the main session topics from previous International and European Semantic Web Conferences (ISWC and ESWC), the work packages of Knowledge Web and the topics of the Knowledge Web summer school (the topics of the REWERSE summer school were already included in E-D5).

Version 1.0 of the topic hierarchy was reported in the Knowledge Web deliverable D3.1.5 in December 2005 and announced on the Knowledge Web and REWERSE mailing lists in January 2006. It was also used to describe the current state of REASE in REWERSE deliverable E-D6 in August 2005 after a first draft of it has been used in REASE in August 2005.

The main changes were as follows:

- Copying Knowledge Engineering / Ontology Engineering to the foundations (from SW) as it has existed long before the Semantic Web came to life; Semantic Web specific issues around ontology engineering are supposed to be classified into the SW3 parts while the Semantic Web independent parts are now to be classified into the foundational category 'Knowledge Engineering / Ontology Engineering'.
- The reasoning part in the foundations was cleaned.
- Within Basic Web Information Technologies two categories were added: Web Data Extraction and Architecture of Web Information Systems.
- The Semantic Web main category was extended along the Semantic Web layer cake, adding three new categories:
    - Semantic Web Infrastructure
    - Semantic Web Proof
    - Security / Trust / Privacy
- The main category 'Semantic Web Special Topics' was introduced. It has been in the initial version of the REASE catalogue and was now merged with the 'other related areas' part. For duplicated categories like Natural Language Processing, the main idea again was to put the Semantic Web specific things into the 'special topic' category while the Semantic Web independent parts are classified under 'IM Natural Language Processing' in the foundational categories.

2.2.4 REWERSE E-D7: Graduate Education Curriculum

REWERSE deliverable E-D7 reported in February 2007 about Version 1.1 of the topic hierarchy (cf. Table 1 in p.9), which included some minor modifications as a result of the analysis of existing material in REASE and learning resources for the Semantic Web curriculum as well as including feedback from the first public announcement.

- Added Temporal Logics as it is very important for example for the Agents community with the Semantic Web community.
- Removing 'Information Management' (originally included from the ACM CCS) because it was too general.
- Removing 'Semantic Web Infrastructure' because it turned out that Semantic Web Services are actually similar to the many topics in 'Semantic Web Special Topics' and the remaining 'Architecture' did not really justify its own main topic (it was actually very broad itself and never used for any categorization of resources in REASE, for example).
- Replacing 'Reasoning Engines' (this category was too general under the super-topic 'Rules') by Ontology Reasoners and Rule Reasoners.
- Adding Integration of Rules and Ontologies.
- Adding eGovernment as an important application domain for Semantic Web technologies.
- Adding Design and Testbed Case Studies.

As the ontoworld wiki `http://ontoworld.org/wiki/`, the wiki for the Semantic Web community, supports 'subtopic' links, we integrated the topic hierarchy into the category system of the wiki. So there is one wiki page for each topic which provides a brief explanation of the topic plus an automatically generated overview about all other wiki pages which have been annotated with this topic. Furthermore, we provide links to find resources about this topic (e.g., to REASE).

As an example, for the topic 'ontology engineering' the wiki contains 21 subcategories and 9 plain wiki pages being tagged with this topic (cf. `http://ontoworld.org/wiki/Category:Topic_ontology_engineering`).

Version 1.1 was the first version where all topics are described by such separate 'Topic' pages in the wiki.

2.2.5 Knowledge Web D3.1.5v2 and REWERSE E-D9: Evaluation of the topic hierarchy

The Knowledge Web deliverable D3.1.5v2 and the REWERSE deliverable E-D9 reported about the evaluation of the topic hierarchy along three lines:

- manual inspection by an expert not yet being involved in the definition
- comparison to existing author keywords in the DBLP++[1] corpus

[1] `http://dblp.l3s.de/dblp++.php`

- comparison to automatically generated topic hierarchies using the Semantic GrowBag approach[2]

As a result, several changes were proposed for inclusion in Version 2 of the topic hierarchy:

- Add 'Information Retrieval / Search' (actually being IM11 in ACM CC)
- Add 'Machine learning' (actually being IS8 in ACM CC)
- Add 'document classification' (which was finally rejected because document classification is often used to generate Semantic Web content (=metadata), but there is no Semantic Web specific part in document classification).

2.3 Version 2 of the topic hierarchy

Version 2 of the topic hierarchy (cf. Table 2) was based upon the following information:

- Collecting all feedback about previous version of the topic hierarchy (e.g., from the development of the joint master curriculum or the classification of learning materials in REASE, or from the creation of the Knowledge Web technology roadmap).
- Integrating topics which were added by other people using the Ontoworld wiki (a true community effort).

Version 2 is available at `http://wiki.ontoworld.org/index.php/Semantic_Web_Topic_Hierarchy`.

For all topics, their occurrence in Google Scholar, Google, and FacetedDBLP were analyzed to ensure that they are somehow 'commonly used'. As a result, the main structure of Version 1.1 of the taxonomy remained unchanged and consists of Foundations, Semantic Web Core Topics and Semantic Web Special Topics. In general, we have tried to simplify the topic hierarchy and to clean duplicates as far as possible, as well as to add categories only if we had strong indications of their importance. Furthermore, we have opted to delete categories for which we could not see any strong indication that they are actually important for the field (initially, we included basically all proposed areas as we assumed that areas that are worked upon in Knowledge Web / REWERSE are also of a somewhat importance).

Hence, the changes compared to Version 1.1 are related to obsolete categories, changed categories, and additional categories as explained in the following sections.

2.3.1 Foundations

- Renamed all subcategories of ontology engineering to include 'ontology' for practical reasons (like searching for related resources on external sources like google scholar).
- Refining category 'Ontology Population / Generation' to 'Ontology Construction / Ontology Learning / Ontology Population' as ontology learning is by far the most prominent term and ontology generation is used rather rarely.
- Added 'Ontology Evolution' as this is a popular term to express dynamics similar to (but more popular as) 'Maintenance and Versioning' of ontologies (this term was introduced by other users of the ontoworld wiki and subsequently integrated).

[2] `http://www.l3s.de/growbag`

- Added 'Ontology Merging' as this also is a very popular term / synonym for 'Ontology Integration'.
- Added 'Ontology Visualization' as requested by several people.
- Removed 'Predicate Logic' as it is covered by 'First-Order Logic'.
- Removed 'Modal Logics' as it is not so important for the Semantic Web and the topic hierarchy should do fine with the other Logics categories.
- Added 'Fuzzy Logic'.
- 'Reasoning' refined by 'Reasoning Engines / Theorem Provers' and 'Fuzzy Reasoning'.
- Removed the subtopics of 'XML' since 'namespaces' and 'programming techniques' are too detailed, 'query languages' are now available in a different category, and 'schema languages' did not justify another hierarchical level and is also a bit too detailed here).
- Added 'Web service discovery' and 'Web service composition' as other people already used them in ontoworld and they are indeed important subtopics of 'web service'.
- Removed 'Architecture of Web Information Systems' as it was too general and not really used.
- Moved 'Security' as a general security category and moved 'Security / Trust / Privacy' to emerging topics (not too much Semantic Web specific material there for now even though it has been part of the Semantic Web layer cake from the beginning.).
- Added 'Rules' and their sub-topics 'Deductive Rules', 'Reactive Rules','Rule Visualization' to the foundations as suggested by Francois Bry (rules have been around before the Semantic Web just as ontology engineering has) and left only the SW-specific parts in 'Semantic Web rules'.
- Added main foundational category Information Access, comprising the subtopics Query languages, Browsing / Navigational Access, Query Algebra, Visual Querying, and Event queries as also suggested by Francois Bry to also separate the 'non-Semantic Web specific' query things from the Semantic Web specific ones.

2.3.2 Semantic Web Core

- Renamed 'Semantic Web Query and Update Languages' to 'Semantic Web Information Access' to make it consistent with 'Information Access' in general.
- Dropped 'Semantic Web Update Languages' as there is not much about it yet (moved it to Emerging Semantic Web Topics).
- Added 'Semantic Web Browsing' as it is a popular research area and several prototypes have been developed recently.
- Refined 'Ontologies in the Semantic Web':
  - Added a separate category 'OWL' because it's one of the main components of the Semantic Web.

  - Moved 'Resource Description Framework / RDFSchema' to here as it is also an ontology language.
  - Removed 'ontology engineering' since there is no Semantic Web specific ontology engineering for now and the rest is covered by the foundation category 'ontology engineering'.
  - Removed 'ontology reasoners' since it is not a commonly used term.
  - Added 'Legacy Ontology Languages (DAML, DAML+OIL)' to be able to refer to the foundations of the Semantic Web languages.
  - Added 'Ontology Repositories' as there is a lot of work going on here.
  - Added 'Ontology Instances' to be able to find examples of ontologies.
  - Added 'Upper-Level Ontologies / Top-Level Ontologies' as the former was already used in the ontoworld wiki and the latter is a similarly often used synonym.
  - Added 'Domain Ontologies' as this was already used in the ontoworld wiki.
- Renamed 'Rules + Logics' to 'Rules' as logics are already covered by the foundational part.
  - Removed 'Rule reasoners' as it is not a commonly used term.
  - Added 'Distributed rule processing' as it is an important topic within 'rules for the Semantic Web'.
- Removed 'Semantic Web proof' since the term is not commonly used and Semantic Web reasoning is covered by other categories.
- Removed 'Semantic Web security / privacy / trust' as apparently not many people are currently working on this topic.
- Renamed 'Applications' to 'Application domains', and
  - Moved 'Law and Engineering' to emerging topics, not included in Version 2.
  - Added several application areas: 'Information Retrieval / Search' (as an outcome of analyzing co-occurring keywords in SW publications),'e-culture', 'Human resources', 'Blogs', 'Business Rules' (already in E-D1, but not considered initially), 'Wikis', 'Digital Libraries' (IM14 in the ACM CC2001), 'Data Integration'.
- Added 'Reasoning in the Semantic Web' for the categorization of Semantic Web specific issues.

2.3.3 Semantic Web Special Topics

- Removed 'design and test bed case studies' again as it did not attract any attention in the previous version.
- Added 'Semantic Desktop'.

Table 1: The Semantic Web Topic Hierarchy, Version 1.1

1 Foundations
- 1.0 Knowledge Engineering/ Ontology Engineering
  - 1.0.1 Methodologies
  - 1.0.2 Ontology Population/Generation
  - 1.0.3 Maintenance and Versioning
  - 1.0.4 Mapping / Translation / Matching / Aligning (Heterogeneity)
  - 1.0.5 Validation
  - 1.0.6 Interoperability / Integration
  - 1.0.7 Modularization and Composition
  - 1.0.8 Tools
- 1.1 Knowledge Representation and Reasoning
  - 1.1.1 Logics
  - 1.1.1.1 Predicate Logic
    - 1.1.1.2 Description Logics
    - 1.1.1.3 F-Logic
    - 1.1.1.4 Modal Logics
    - 1.1.1.5 First-Order Logic
    - 1.1.1.6 Temporal Logic
  - 1.1.2 Logic Programming
    - 1.1.2.1 Horn Logic
    - 1.1.2.2 Datalog
    - 1.1.2.3 Prolog
    - 1.1.2.4 Hilog
  - 1.1.3 Reasoning
- 1.2 Information Management
  - 1.2.1 Data Modelling
    - 1.2.1.1 Conceptual Models; ontologies, UML
    - 1.2.1.2 Relational data model
    - 1.2.1.3 Semistructured data
    - 1.2.1.4 Object-oriented model
  - 1.2.2 Database systems
- 1.3 Basic Web Information Technologies
  - 1.3.1 XML
    - 1.3.1.1 Namespaces
    - 1.3.1.2 Schema Languages
    - 1.3.1.3 XML Query and Transformation Languages
    - 1.3.1.4 XML Programming Techniques
  - 1.3.2 Web Data Integration
  - 1.3.3 Security
  - 1.3.4 Web Services
  - 1.3.5 Personalization Techniques
  - 1.3.6 Web Data Extraction / Information Extraction
  - 1.3.7 Architecture of Web Information Systems
- 1.4 Agents
- 1.5 Natural Language Processing

2 Semantic Web Core topics
- 2.1 Resource Description Framework / RDFSchema
- 2.2 Query and Update Languages
  - 2.2.1 Query Languages
  - 2.2.2 Update Languages
- 2.3 Ontologies
  - 2.3.1 Ontology Representation / ontology languages / OWL
  - 2.3.2 Ontology Engineering
  - 2.3.3 Ontology Reasoners
- 2.4 Rules + Logic
    - 2.4.1 Rule Languages
    - 2.4.2 Rule Markup
    - 2.4.3 Reasoning Languages
    - 2.4.4 Rule Reasoners
    - 2.4.5 Integration of Rules and Ontologies
- 2.5 Proof
- 2.6 Security / Trust / Privacy
- 2.7 Applications
  - 2.7.1 Knowledge Management
  - 2.7.2 e-learning
  - 2.7.3 Bioinformatics
  - 2.7.4 Multimedia
  - 2.7.5 e-health
  - 2.7.6 e-business
  - 2.7.7 Law
  - 2.7.8 Engineering
  - 2.7.9 e-government

3 Semantic Web Special Topics
- 3.1 Natural Language Processing / Human Language Technologies
- 3.2 Social Impact of the Semantic Web
- 3.3 Social Networks and Semantic Web
- 3.4 Peer-to-Peer and Semantic Web
- 3.5 Agents and Semantic Web
- 3.6 Semantic Grid
- 3.7 Outreach to Industry
- 3.8 Benchmarking and Semantic Web Scalability
- 3.9 Design and Testbed Case Studies
- 3.10 Semantic Web Services

Table 2: The Semantic Web Topic Hierarchy, Version 2

1 Foundations
- 1.0 Knowledge Engineering/ Ontology Engineering
  - 1.0.1 Ontology Engineering Methodologies
  - 1.0.2 Ontology Construction / Ontology Learning / Ontology Population
  - 1.0.3 Ontology Evolution / Ontology Maintenance / Ontology Versioning (Dynamics)
  - 1.0.4 Ontology Mapping / Ontology Translation / Ontology Matching / Ontology Aligning (Heterogeneity)
  - 1.0.5 Ontology Evaluation / Ontology Validation
  - 1.0.6 Ontology Interoperability / Ontology Integration / Ontology Merging
  - 1.0.7 Ontology Modularization / Ontology Composition
  - 1.0.8 Ontology Engineering Tools
  - 1.0.9 Ontology Visualization
- 1.1 Knowledge Representation and Reasoning
  - 1.1.1 Logics
    - 1.1.1.2 Description Logics
    - 1.1.1.3 F-Logic
    - 1.1.1.5 First-Order Logic
    - 1.1.1.6 Temporal Logics
    - 1.1.1.7 Fuzzy Logic
  - 1.1.2 Logic Programming
    - 1.1.2.1 Horn Logic
    - 1.1.2.2 Datalog
    - 1.1.2.3 Prolog
    - 1.1.2.4 Hilog
    - 1.1.2.4 Monotonic and Nonmonotonic Negations
  - 1.1.3 Reasoning
    - 1.1.3.1 Reasoning Engines / Theorem Provers
    - 1.1.3.2 Fuzzy Reasoning
- 1.3 Basic Web Information Technologies
  - 1.3.1 XML
  - 1.3.2 Web Data Integration
  - 1.3.4 Web Services
    - 1.3.4.1 Web Service Discovery
    - 1.3.4.2 Web Service Composition
  - 1.3.5 Personalization Techniques
  - 1.3.6 Web Data Extraction / Information Extraction
- 1.4 Agents
- 1.5 Natural Language Processing
- 1.6 Security / Trust / Privacy
- 1.7 Machine learning
- 1.8 Rules
  - 1.8.1 Deductive Rules
  - 1.8.2 Reactive Rules
  - 1.8.1 Rule Visualization
- 1.9 Information Access
  - 1.9.1 Query Languages
  - 1.9.2 Browsing / Navigational Access
  - 1.9.3 Query Algebra
  - 1.9.4 Query Optimization
  - 1.9.5 Visual Querying
  - 1.9.6 Event Queries

2 Semantic Web Core topics
- 2.2 Semantic Web Information Access
  - 2.2.1 Semantic Web Query Languages
  - 2.2.3 Semantic Web Browsing
- 2.3 Ontology on the Semantic Web
  - 2.3.1 OWL
  - 2.3.4 Resource Description Framework / RDFSchema
  - 2.3.5 Legacy Ontology Languages (DAML, DAML+OIL)
  - 2.3.6 Ontology Repositories
  - 2.3.7 Ontology Instances
    - 2.3.7.1 Upper-Level Ontologies / Top-Level Ontologies
    - 2.3.7.2 Domain Ontologies
  - 2.3.7 Semantic Annotation / Microformats
- 2.4 Web and Semantic Web Rules
  - 2.4.1 Rule Languages
  - 2.4.2 Rule Markup
  - 2.4.5 Integration of Rules and Ontologies
  - 2.4.6 Distributed Rule Processing
- 2.7 Semantic Web Application Domains
  - 2.7.1 Knowledge Management
  - 2.7.2 e-learning
  - 2.7.3 Bioinformatics
  - 2.7.4 Multimedia
  - 2.7.5 e-health
  - 2.7.6 e-business
  - 2.7.9 e-government
  - 2.7.10 Information Retrieval / Search
  - 2.7.11 e-culture
  - 2.7.12 Human resources
  - 2.7.13 Blogs

3 The Use of the Topic Hierarchy in the European Academy for Semantic-Web Education

3.1 The European Academy for Semantic-Web Education

In 2006 the European Networks of Excellence Knowledge Web and REWERSE have launched the European Academy for Semantic-Web Education (`http://www.semantic-web-academy.eu/`), an international distributed Master of Science program, in cooperation with leading European universities, all known for excellence in the field of Semantic Web:

- Free University of Bozen-Bolzano, Italy (coordinator)
- Vrije Universiteit Amsterdam, The Netherlands
- Università di Trento, Italy
- Universidad Politécnica de Madrid, Spain
- Universidade Nova de Lisboa, Portugal
- University of Innsbruck, Austria

The European Academy for Semantic-Web Education, completely in English, involves studying one year at one of the above universities, and completing the second year with a stay in another of the partner universities. After this, the student may obtain, together with the European degree, a national Master of Science degree for each of the visited universities (e.g., in Italy this would be the Laurea Specialistica degree).

The Academy is meant for students with a bachelor degree from computer science, artificial intelligence, or related areas. The students will study up-to-date developments in the Semantic Web area; the program consists of Foundation units, selected Advanced units, a Project, and a Master thesis, for a total of 120 ECTS credits. All participating universities offer the foundation units; the advanced units will be different and depend on the participating university.

Goal of the European Academy for Semantic-Web Education is to build a new generation of Semantic Web experts, and up-skill the European workforce with ideas that point to the future. It is foreseen that students graduating in the European Academy for Semantic-Web Education are suitable candidates to follow-up PhD studies and, e.g., to visit later summer schools from Knowledge Web or from associated partners of the network such as the REWERSE summer school. On the other hand they have sufficient knowledge to work in companies, e.g., realizing Semantic Web applications.

For the European Academy for Semantic-Web Education a single coherent structure is prescribed. It is based on common and compulsory foundational modules comprising 36 ECTS credit points at minimum, which are taught at each partner institution, selected advanced modules comprising at minimum 36 ECTS credit points, which are based on the specific strengths in research and teaching of the partner institutions (thus, they can vary from partner to partner) and a research master thesis project with a minimum of 36 ECTS credit points.

The program of the Academy has been developed using the Semantic Web Topic Hierarchy of Table 1 (p.9) and the course descriptions have been classified by the topics of the hierarchy. For details concerning the modules offered by Knowledge Web participants see Knowledge Web deliverable D3.2.4. The following sections provide descriptions of the modules offered within the Academy by the REWERSE participant Universidade Nova de Lisboa.

3.2 Selection of the topics for UNL Profile

UNL combines expertise in logic and knowledge representation, with expertise in distributed and parallel systems and the Grid. As such, in addition to the specific topics offered in the basics modules, UNL offers advanced courses in multi-agent systems (with special emphasis on logic based approaches, with application to the Semantic Web), advanced logics and usage of logic for specification and verification, architectures for distributed and parallel systems including web information systems, the Grid, and mobile and pervasive computing. Also taking advantage of in-house expertise, there is a course on constraints and constraint programming, and also a course on application to bioinformatics, which is being organised in collaboration with the Chemistry department.

Regarding the basic modules, they are offered in the first year by all partner institutions with the common aim of bringing the students to an equivalent level of skills and knowledge. Five basic modules are defined: *Catch-up, Basic Semantic Web Technologies, Formal Foundations of Knowledge Representation, Web Infrastructure, and Information and Knowledge Systems*. As mentioned above, each student has to complete at least 36 ECTS credits in courses of the basic modules.

Courses in the Catch-up module are mainly to fresh up the knowledge of the student with material which are the base for the joint program. These include logics, databases, and artificial intelligence. Furthermore especially designed bridging courses integrated into the basic modules offered in the first semester shall help 3rd country students and students from other but related professions in their adaptation to a Masters Course of European educational level. At UNL we anticipate bridging courses at least in computer networks and logic and constraint programming. The courses in the Catch-up module ensure that the next courses can build upon the same knowledge for all students. The Catch-up module has no minimal ECTS credit points in order to be flexible with respect to the basic knowledge of the students from each institution.

The basic representation languages for the Semantic Web are covered by the courses in the Basic Semantic Web Technologies. Obviously this includes the introduction into XML, RDF, RDFS, and OWL. For the courses in this (and the following) modules a minimum of 8 ECTS credit points is required.

All representation languages in the Basic Semantic Web Technologies module are based on formal knowledge representation. A deep understanding of the underlying logics is essential for these representation languages and communicated in the courses of the Formal Foundations of Knowledge Representation module (8 ECTS min). Courses in this module cover the basic knowledge on logic, formal methods, knowledge representation and reasoning, and theory of computing.

The next two modules complete the basic modules with topics strongly related to the Semantic Web. The Web Infrastructure module (8 ECTS min) covers the basic knowledge on distributed systems, network technologies, Internet technologies, mobile services, and security, i.e. knowledge about the basic infrastructure for the Semantic Web.

The last module, Information and Knowledge Systems (8 ECTS min), focuses on the more classical approaches for information retrieval. Courses in this module may cover the basic knowledge on knowledge management, information systems, advanced database technologies, and semi-structured data.

The descriptions of all the courses in these modules can be found in the following sections. The descriptions include references to the Semantic Web Topic Hierarchy Version 2.

3.3 Basic Modules at UNL

3.3.1 Catch-up Courses

The following courses are offered at UNL and can be used as catch-up courses by students depending on each one's individual need. These include topics on logics, artificial intelligence, databases, logic and constraint programming and computer networks.

Course	Introduction to Logics
ID	UNL-IL
ECTS:	6
Classification	1.1.1
Description:	The course offers a comprehensive introduction to Logics covering the main sub-areas as well as main methods and techniques. It recalls basic notions from propositional and first order logic, deduction, and proof theory.

Course	Artificial Intelligence
ID	UNL-IA
ECTS:	6
Classification	—
Description:	This course provides a comprehensive view of Artificial Intelligence, by selecting for study some of the main topics of the discipline: search, knowledge representation and resoning, planning and learning. The several approaches and paradigms are illustrated in the several course projects. These subjects are treated in sufficient detail that allows students to solve toy problems as well as understand the difficulties of real problem instances. In the lab classes students will develop selected programming projects where the learnt techniques can be applied to solve some problems, both by implementing small prototypes and by modeling the problems in existing tools.

Course	Databases
ID	UNL-BD
ECTS:	6
Classification	— (1.1.2.2, 1.3.1)
Description:	The goal of the course is to provide the basis for the modelling, implementation, analysis and manipulation of relational databases. It covers: an introduction to database management systems; the entity-relationship data model; relational model and relational algebra; SQL query and manipulation language; Datalog query language; topics on database integrity and security; relation databases normalisation theory. It also discusses other database models, including object/relational model, deductive databases and a brief introduction to XML. The course includes a strong practical component, where the students have to implement a relational database in a commercial database management system.

Course	Foundations of Logic and Constraint Programming
ID	UNL-LCP
ECTS:	6
Classification	1.1.2
Description:	This course is concerned with the foundations of logic programming and constraint logic programming. The basic computation mechanisms of unification and SLD-resolution are introduced. The declarative and the operational semantics of logic programs are given and related to the procedural semantics. A logic programming language is introduced as an example of a declarative programming language. Logic programs with constraints are introduced and basic computation mechanisms given. The module concludes with examples of constraint logic programming languages. In addition basic knowledge of combinatorics and analysis of algorithms is taught.

Course	Computer Networks
ID	UNL-RC
ECTS:	6
Classification	—
Description:	The course covers topics of principles of computer networks including: Application level Transport level; Network level; Data link level. Local networks. It also includes an introduction to mobile networks.

3.3.2 Basic Semantic Web Technologies Courses

Courses in this module cover the basic representation languages for the Semantic Web including XML, RDF, RDFS, and OWL

Course	XML Technologies
ID	UNL-XMLT
ECTS:	6
Classification	1.3.1
Description:	The first part describes XML based technology and focuses on representing hierarchical and semi-structured data, and related W3C recommendations: XML Schemas, XML Namespaces, and XML Base. The text-centred and data-centred document views are discussed and compared. XML data model integrity supporting mechanisms are analysed, namely XLink, XPointer and XML Inclusions. Querying and transformation languages for XML documents are described and deeply studied, in particular XSL based-languages (XPath and XSLT) and XQuery. The course continues by relating the database relational model (DBMSs) with the hierarchical model (XML), and studying the mappings between them. The course continues by presenting client-server architectures integrating XML and relational databases, as well as XML support in the major DBMSs. The course concludes with DOM and SAX programming techniques, and construction of Web Services using SOAP.

Course	Reasoning in the Semantic Web
ID	UNL-RSW
ECTS:	6
Classification	2.3.4, 1.1.1.2, 2.4.2, 2.4.5
Description:	The course, starts by explaining and motivating the origins of the Semantic Web and its logical layered structure. Some basic concepts are overviewed, namely UNICODE, URIs and IRIs, XML Base, XML Namespaces, XSL, and XML Canonicalization. The Resource Description Framework (RDF) and RDF Schema languages are introduced for describing resources and basic vocabularies in the Semantic Web. RDF(S) model theory and inference mechanisms are also addressed, as well as practical applications and its limitations. Description Logics are then introduced as better knowledge representation formalism. Its constructs and semantics are introduced, as well the basic reasoning tasks and corresponding algorithms. The OWL language is presented and applications are provided. The course finishes, by studying the existing proposals for the integration of ontologies with rules in the Semantic Web, in particular the RuleML language proposal is discussed.

3.3.3 Formal Foundations of Knowledge Representation Courses

Courses in this module cover the basic knowledge on logic, formal methods, knowledge representation and reasoning, and theory of computing.

Course	Science of Computational Logics
ID	UNL-SCL
ECTS:	6
Classification	1.1.2, 1.1.3
Description:	Topics in the areas of equational reasoning, deduction, proof theory, abduction and induction, non-monotonic reasoning, logic-based program development, natural language processing and machine learning as well as logic and connectionism are covered.

Course	Knowledge Representation and Reasoning
ID	UNL-RSW
ECTS:	6
Classification	1.1.1, 1.1.2, 2.4.1, 2.4.5
Description:	The course intends to give to the students a current perspective on the logical languages for representing knowledge, and its applications, as well as supplying a coherent and rigorous approach to the different functionalities of reasoning in Artificial Intelligence, with support in Computational Logic. It covers aspects of common sense reasoning with non-monotonic languages, and of representation of ontologies using description languages. Up-to-date languages and tools will be used for modelling concrete problems.

Course	Theory of Computing
ID	UNL-TC
ECTS:	6
Classification	
— Description:	The objective of the course is to present an introduction to: Definition of the syntax of formal languages (e.g. programming languages); Syntactic analysis (parsing); Existence of problems without algorithmic solution; Formal models of computation; Structural operational semantics and program verification. At the end of the course the student should know the algorithms presented, apply them to concrete examples, and should also be able to use the mathematical language to specify automata.

3.3.4 Web Infrastructure Courses

Courses in this module may cover the basic knowledge on distributed systems, network technologies, internet technologies, mobile services, and security.

Course	Introduction to Distributed Systems
ID	UNL-SD
ECTS:	6
Classification	1.3.4
Description:	In the end of this course, students are expected to understand the fundamental characteristics of distributed systems and know the models, methods and techniques used to create distributed systems. In particular, the course focuses on the discussion of the aspects that influence simplicity, security, performance and availability. The course covers the basic characteristics of distributed systems, interprocess communication, authentication and confidentiality, naming systems, Web services. It also includes an introduction to data replication.

Course	Advanced Distributed Systems
ID	UNL-DS
ECTS:	6
Classification	—
Description:	The course aims at introducing techniques and algorithms used in the design and implementation of distributed systems with some complexity, involving multiple elements. In this context, special emphasis will be given to problems and solutions that influence scalability, availability, fault tolerance and efficiency. The course covers: Distributed System Architecture; Time in Distributed systems; Introduction to distributed co-ordination algorithms; Distributed transactions; Replication and fault tolerance; Introduction to peer-to-peer systems.

Course	Computer and Network Security
ID	UNL-CNS
ECTS:	6
Classification	1.6
Description:	The skills acquired in this course aim at providing students with fundamental knowledge on security in computer systems and networks, and its applications. The course covers: Security in computer and network systems; Cryptography, authentication and Key management; Secure protocols and applications; Security in web environments; Secure infrastructures. .

3.3.5 Information and Knowledge Systems Courses

Courses in this module may cover the basic knowledge on knowledge management, information systems, advanced database technologies, semistructured data.

Course	Database Systems
ID	UNL-DBS
ECTS:	6
Classification	—
Description:	Introduces the concepts for the construction of database management systems. It covers database systems architectures, including parallel and distributed databases, data storage in database systems, query processing and optimization, and transaction management in databases.

Course	Data Warehousing
ID	UNL-DW
ECTS:	6
Classification	—
Description:	This course presents the concepts, architectures and models of Data Warehouse. The two mainstream schools are presented and discussed. Contents: Multidimensional Modeling for OLAP and Deriving OLAP Multidimensional models from OLTP models. Conceptual modeling for multidimensional models. ETL. Metadata. OLAP querying: SQL, MDX, JAVA APIs. Geographic Information in OLAP. At the end of this course the students should know about: Data Warehouse architectures and components; What is OLAP and what are its main operations; Multidimensional models and its ROLAP implementations; Conceptual models for multidimensional. Moreover, the students should be able to perform: Analysis and design a multidimensional model; Analysis and design ETL processes; Analytical exploitation of a multidimensional model.

Course	Middleware Systems and Technologies
ID	UNL-MWST
ECTS:	6
Classification	— (1.3.4)
Description:	The course provides fundaments of programming with different paradigms and environments of distributed programming and design of software architectures for integration, distribution and composition of services. It covers: Architectures and characterization of middleware; Client- Server architectures; distributed object middleware; Other paradigms supported by distributed middleware; Entreprise Architecture Integration; Workflow architectures and systems; Web Services and Web Application Servers; Service oriented architecture.

3.4 Advanced Courses at UNL

Course	Advanced Logics
ID	UNL-AL
ECTS:	9
Classification	1.1.1, 1.1.2.4, 1.1.3
Description:	The aim of this course is to introduce basic concepts beyond first-order predicate logics. In Computer Science many different logics and deductive systems exist. First we introduce higher order logic (HOL) as a framework for specifying syntactic and deductive notions of different logics. HOL is used in several interactive proof tools, like PVS and Isabelle. In addition, specific families of logics aimed at different application areas are introduced: logics of time and computation (modal logics, temporal Logics), logics for reasoning about knowledge (epistemic logic). Finally we introduce the mu-calculus which allows to define recursive temporal properties and we present a tableau based deduction calculus for the mu-calculus. The mu-calculus and its deduction system can be used to define problem oriented systems of modal operators and corresponding deduction systems.

Course	Multi-Agent Systems
ID	UNL-MAS
ECTS:	9
Classification	1.4, 3.5
Description:	This course introduces the students to multi-agent systems, with a coherent, rigourous and formal approach, and provide students with the required knowlege and methodologies to research in the area of Multi-Agent Systems. It covers: Computational logic paradigms and formalisms for expressing agents and agent societies, with emphasis on a Logic Programming approach; Agent and agent society architecture and evolution; Combining rationality and reactivity; Centralized and distributed control. Communication and cooperation among agents; Illustrative applications; Implemented systems and tools.

Course	Machine Learning and Data Mining
ID	UNL-MLDM
ECTS:	6
Classification	1.3.6
Description:	The course starts by improving the basic knowledge on machine learning that students acquired in basic artificial intelligence courses, namely with: data analysis, clustering (unsupervised classification) and supervised learning. It then applies these concepts and methodologies in Data Mining, where the issues of data denormalisation, and extraction and description of knowledge from large volumes of data is tackled. Special focus will then be given to extraction of knowledge from Web data.

Course	Logics for Specification and Verification
ID	UNL-LSV
ECTS:	6
Classification	—
Description:	Software production for critical applications, in which errors have serious consequences, makes use of formal verification techniques based on reasoning about software models. The goal of this course is to introduce these techniques, by using logics as a framework for specification and verification, based on tools that automate the analysis of the specification and verification process. The course cover: Boolean Satisfaction; Specification in relational first order logic; Model checking, including transition systems, temporal logic; Binary Decision Diagrams; Symbolic Model checking.

Course	Topics in Finite Domain Constraints
ID	UNL-FDC
ECTS:	6
Classification	—
Description:	This course deals with consistency criteria for constraint networks: arc-, path-, and k-consistency. Algorithms to enforce these criteria, and study of their complexity. Implementation of constraint solvers: indexical constraints. Constructive approach for combining constraints. Constructive disjunction, and cardinality constraints. Global constraints: their specification and implementation. Redundant constraints: advantages and disadvantages of their use in scheduling, planning and other resource management applications.

Course	Semantic Web for Bioinformatics
ID	UNL-BIO
ECTS:	6
Classification	2.7.3
Description:	This course will focus on current bioinformatics services that support Semantic Web standards, the application of Semantic Web concepts in bioinformatics, and challenges for data integration in bioinformatics. The curriculum will include examples of services that provide XML, SOAP, or RDF interfaces (such as the Protein Data Bank or UniprotRDF) and practical applications of data retrieval.

Course	Grid Computing Systems and Applications
ID	UNL-GRID
ECTS:	6
Classification	3.6
Description:	The course aims at providing a global perspective of paradigms and large scale application development environments based on the Grid. Moreover it should motivate for a new class of parallel and distributed applications over Grids and their impact in various areas, with access to virtual collaborative spaces, distributed information systems, and data and services repositories. It covers: Characteristics of Grid computing systems; Arquitectures for Grid computing systems; Grid programming and development models, tools and environments; Computational grids, data grids and interaction grids; Grid architectures, middleware and standards for Open Grid Services Architectures; Application case studies.

Course	Mobile and Pervasive Computing
ID	UNL-MPC
ECTS:	6
Classification	—
Description:	The course provides a basic knowledge on the technologies used in mobile and pervasive computing, and their characteristics, as well as specific problems of these environments. It also provides fundamental knowledge on models and techniques used for developing application in mobile and pervasive environments. It covers: Characterisation of mobile computing environments; Wireless networks; Models and support middleware for mobile systems; Technological aspects of mobile systems; Data management in a mobile setting; Mobile application development; Introduction to sensor networks.

Course	Parallel and Distributed Computing
ID	UNL-PDC
ECTS:	6
Classification	—
Description:	The course aims at providing a global perspective on parallel and distributed computing, centred in models and algorithms for problem solving that involve exploring concurrency, distributions and parallelism. The multiple dimensions are presented in an integrated manner, that covers the models and the distributed and parallel algorithms, the approaches for their programming, and their implementations in distributed architectures. It promotes the articulation between theory and practice. The course covers: An unified vision of concurrency, distribution and parallelism; Models and fundamental algorithms in parallel and distributed computations; Parallel and distributed algorithms; Development environments for parallel and distributed applications.

4 The Use of the Topic Hierarchy in REASE

A flattened version of the taxonomy in Table 2 is used as a REASE catalogue. The catalogue is presented in Table 3 p.24 (from Version 2); the figures show number of REASE resources classified by a given topic.

4.1 Creating the Catalogue

While the Semantic Web Topic Hierarchy reflects, of course, a compromise among the different opinions within the Semantic Web community (e.g., some consider 'natural language processing' as a foundational topic while others treat it as special topic), we had to generate an even more simplified version for technical reasons: the REASE catalogue, though customizable, can only handle up to two hierarchical levels at maximum. This has also the advantage that the number of categories is more limited, so REASE users are not lost in too deep catalogue categories. As a result, we skipped the first-level hierarchy of 'foundations', 'Semantic Web core topics' as there sometimes also is no real distinction between them. There was, for example, quite some discussion during the creation of the topic hierarchy whether ontologies are foundational or belong to the core topics. Furthermore, we ignored the subcategories of 'Logics', 'Logic Programming', and 'XML', since it was not expected that learning material in REASE will deal specifically with one of the subtopics. Instead, it is expected that learning units in these topics give an overview, for example, on 'Logics' and discuss most of the sub-categories.

4.2 Annotating Resources in REASE

For the categorization of the resources in REASE, we adopted a provider-based approach to be able to ensure a high-quality of the categorization (assuming that the providers of the material know best what the material is all about). In principle, resources can be associated with any number of categories though we carefully monitored the annotations of new resources since too many categories are a good indicator that the resources is actually not sufficiently focused and should perhaps better be broken into smaller (more focused) chunks.

4.3 Usage of Categories in REASE

As can be seen from Table 3 p.24, resources are not distributed evenly among the categories. For several general topics, quite a lot of resources are available, including:

1. Application Domains (68)
2. Semantic Web Special Topics (64)
3. Knowledge Engineering / Ontology Engineering (55)
4. Ontologies on the Semantic Web (50)
5. Basic Web Information Technology (33)
6. Outreach to Industry (28)
7. Ontology Engineering Tools (26)
8. OWL (22)

Table 3: The REASE Catalogue

- Foundations
  - Knowledge Engineering/ Ontology Engineering (55)
    - Ontology Engineering Methodologies (17)
    - Ontology Construction/Learning/ Population (4)
    - Ontology Evolution / Maintenance / Versioning (4)
    - Ontology Mapping /Translation / Matching / Aligning (8)
    - Ontology Evaluation / Validation (2)
    - Ontology Interoperability / Integration / Merging (10)
    - Ontology Modularization and Composition (2)
    - Ontology Engineering Tools (26)
    - Ontology Visualization (0)
  - Knowledge Representation and Reasoning (22)
    - Logics (8)
    - Logic Programming (3)
    - Reasoning (8)
  - Basic Web Information Technologies (33)
    - XML (6)
    - Web Data Integration (2)
    - Web Services (1)
    - Personalization Techniques (7)
    - Information Extraction / Web Data Extraction (9)
- Semantic Web Core topics:
  - Information Access (10)
    - Query Languages (7)
    - Browsing (2)
    - Query Algebra (0)
    - Query Optimization (0)
    - Visual Querying (0)
    - Event Queries (0)
  - Ontologies on the Semantic Web (50)
    - OWL (22)
    - RDF / RDFSchema (17)
    - Legacy Languages (DAML,DAML+OIL)(0)
    - Ontology Repositories (2)
    - Ontology Instances (1)
    - Semantic Annotation (4)
  - Security / Trust / Privacy in the Semantic Web (2)
  - Rules(15)
    - Rule Languages (4)
    - Rule Markup (3)
    - Reasoning Languages
    - Integration of Rules and Ontologies (3)
    - Distributed Rule Processing (0)
    - Deductive Rules (0)
    - Reactive Rules (0)
    - Rule Visualization (0)
  - Application Domains (68)
    - Knowledge Management (11)
    - e-learning (2)
    - Bioinformatics (6)
    - Multimedia (6)
    - e-health (5)
    - e-business (14)
    - e-government (2)
    - Information Retrieval / Search (12)
    - eCulture (2)
    - Human resources (0)
    - Blogs (2)
    - Business Rules (1)
    - Wikis (4)
    - Digital Libraries (1)
    - Data Integration / Enterprise Information Integration (10)
- Semantic Web Special Topics (64)
  - Natural Language Processing / Human Language Technologies (11)
  - Social Impact of the Semantic Web (6)
  - Social Networks and Semantic Web (12)
  - Peer-to-Peer and Semantic Web (3)
  - Agents and Semantic Web (3)
  - Semantic Grid (2)
  - Outreach to Industry (28)
  - Benchmarking and Semantic Web Scalability (1)
  - Semantic Web Services (14)
  - Semantic Desktop (1)

Many topics in this list are very general high-level categories which match for many resources; some reflect main areas of Semantic Web research (like ontology engineering, which is also a special area of many partners in the involved partners of the Networks of Excellence Knowledge Web and REWERSE). The category 'Outreach to Industry' in this list of most popular categories is such popular because we especially tried to find material suited for industrial education and used this category initially to tag the material (this is now superseded by another orthogonal mechanism).As some resources have been tagged with the 'high-level' tag only, in Table 3 the number of resources in a category may not equal the sum of the numbers of the resources in its subcategories.

4.4 Usage of the Catalogue for Browsing

In principle, we have three major ways of finding resources on REASE (apart from a search engine directly pointing to a page describing a particular resources):

1. Basic keyword search (as offered on the start page)
2. Advanced search (combining keyword + filters, where one filter uses the topic hierarchy to restrict search results to a set of categories)
3. Browsing the catalogue.

the latter two actually using the catalogue.

From analyzing our web log files for the three different versions of REASE we deployed during May 2006 to April 2007, we found that advanced search is used only rarely, and that browsing is the preferred option to search the platform. Table 4 shows which of the search functionalities (i.e., keyword search, advanced search, and browsing) is used how often, in relation to the overall number of browse/search events.

Table 4: Usage of search functionality

Search functionality	**Version 1** (07-10/2006)	**Version 2** (11/2006-01/2007)	**Version 3** (02/2007-04/2007)
Keyword search	221 = 9.9%	895 = 17.2%	448 = 22.1%
Advanced search	78 = 3.5%	631 = 12.1%	80 = 3.9%
Of which using catalogue filter	11 = 14.1%	336 = 53.2%	21 = 26%
Browse	1931 = 86.6%	3682 = 70.7%	1500 = 74.0%
Of which 'Browse most popular'	0	62	301
Browse with filters	60 (3%)	364 (10%)	111 (7%)
...of which using title filter	31	317	55
...of which using industrial filter	0	0	62
Total	2230	5208	2028

It can be seen that REASE users actually do use the browse functionality (about 75% of all search functionality) even though the keyword search functionality has gained more attention in the latest version (from 10% to 20%) because of the more prominent position of the search box on the entry page. The filtering options, however, which allow to restrict browsing only industrial resources or a specific kind of resources such as recorded lectures, are used rather rarely (less than 7%, not considering version 2 here, which included a user study where we

Table 5: Going to the resource details page from...

Search functionality	**Version 1** (07-10/2006)	**Version 2** (11/2006-01/2007)	**Version 3** (02/2007-04/2007)
...keyword search on the main page	883	1270	554
...the catalogue	1489	940	509

explicitly requested people to use the filter), the filter to show resources for an industrial audience only was used by less than 4% of all users. This is either because there are still too few resources to let this filter become useful or not many users in REASE actually come from industry. For those users who actually do use this special filter to support professional users, however, it seems to be as valuable as the title filter, which was previously used most often.

Furthermore, about 20% of the keyword queries were posed using advanced search (leaving aside the data from version 2, which included the log files from the user study). From the advanced queries, 14% used the catalogue filter in the first version and 26% in the third version, probably because of the improvements in the filter visualisation, though the absolute usage of the filter facility is still very low. This either means that REASE users can find what they are looking for even without the catalogue filtering (because there are not yet that many resources in REASE), or they don't really know how to use the catalogue filtering mechanisms or cannot find it (as opposed to the participants of the user study who were explicitly asked to use it).

Table 5 shows how many users actually looked at the details of a resource, either having found from keyword search or from browsing. While in version 1 most users (62%) found interesting results to look at more closely using the catalogue, using this facility has changed in version 2 and 3 towards a more equally balanced ratio between starting with a keyword search from the main page and browsing the catalogue. Since browsing is less dominant here than for the total number of browse/search events as shown in Table 4, we can deduce that users in general need more clicks to find an interesting resource while browsing compared with when doing a keyword search.

Hence, the catalogue is widely used not only for searching on REASE (as shown in Table 4), but also for successfully finding resources (as shown in Table 5).

In Table 6 we provide data about the most popular categories in the catalogue from the perspective of the REASE users, based on how often they actually clicked on each category (for version 1 and 3 only, we leave out version 2 since it was heavily biased because of the user study). In general, no category is far ahead of the others and actually the ranking changed not too much between the versions 1 and 3; only the category 'Outreach to Industry' was superseded by 'Ontology Engineering Methods' in this top-10 list and 'Basic Web Information Technology' lost quite some popularity (rank 2 down to rank 9). This should be related to the changes we made to the catalogue such as removing the resorting of the catalogue (clicking on this category actually put the catalogue back into it's initial sorting order, which some users might have used to circumvent the resorting).

To evaluate the usefulness of the catalogue in actually *f*inding resources we restricted the data to those accesses where users actually found a resource (i.e., clicked on one of the shown resources afterwards). The results are presented in Table 7. In general, it can be seen that 'Knowledge/Ontology engineering' and 'Ontologies for the Semantic Web' are the most popular topics for both clicks on the catalogue category and actually finding interesting resources, independent from the version of REASE. 'Basic Web Information Technologies' was very popular

Table 6: The top-10 most popular catalogue categories

Category	**Version 1**	Category	**Version 3**
Knowledge / Ontology Engineering	10.0%	Knowledge / Ontology Engineering	8.1%
Basic Web Inform. Techn.	9.7%	Ontologies for the SW	7.1%
Ontologies for the SW	8.6%	SW Applications	7.0%
SW Special Topics	6.3%	Knowledge Representation and Reasoning	4.9%
Knowledge Representation and Reasoning	5.8%	SW Special Topics	3.5%
RDF	5.2%	SW Rules	3.5%
SW Applications	5.1%	RDF	3.5%
Outreach to Indutry	4.9%	Ontology Engineering Methods	3.4%
SW Rules	4.4%	Basic Web Inform. Techn.	3.3%
SW Query and Update Languages	3.8%	SW Query and Update Languages	2.9%
...	...	...	...

Table 7: The top-10 most popular catalogue categories from which users found interesting resources

Category	**Version 1**	Category	**Version 3**
Knowledge/Ontology Engineering	13.7%	Knowledge/Ontology Engineering	14.1%
Basic Web Inform. Techn.	13.6%	Ontologies for the SW	13.1%
Ontologies for the SW	12.8%	Applications	10.2%
Outreach to Industry	10.2%	Knowledge Representation and Reasoning	7.8%
RDF	7.8%	Semantic Web Rules	5.3%
SW Special Topic	6.3%	Ontology Engineering	5.3%
Applications	5.8%	OWL	5.3%
Knowledge Representation and Reasoning	5.0%	Ontology Mapping	4.9%
SW Rules	4.3%	RDF	3.9%
OWL	2.2%	Ontology Engineering Methodology	3.4%
...	...	...	...

Table 8: Least popular categories to access resources from

Category	Percentage
...	...
Social networks and the SW	0.3%
(deleted category available shortly only)	0.3%
Web Services	0.1%
XML	0.1%
eLearning	0.1%
Ontology reasoners	0.1%
Modularization and Composition	0.1%
Benchmarking and Scalability	0.03%
Web Data Integration	0.03%
Ontology Validation	0.03%

in version 1 but is less popular in version 3 probably because of the changes made to the presentation of the catalogue as briefly mentioned above and described in more details in E/T-D9. Furthermore, for 'Outreach to Industry' the number changed significantly because we are using a different way to tag industrial resources in version 3, so this category will become obsolete in the future. In general, however, changes to the catalogue are also related to the resources being available in each category, which has changed over time while new resources were added.

Finally, we have examined the 10 least popular categories to access and find a resource (cf. Table 8). We have left out those categories where no resources are available on REASE and, hence, which are not shown in the REASE catalogue (in total 45 categories contain / contained resources in REASE).

Most of these resources are not accessed often just because they were populated with resources only recently, so it's difficult to judge the usefulness of these categories just by looking at how often they were used to find resources in REASE. There are few exceptions such as Web Services and XML, where people most likely know other sources in the web to find good material about. However, these categories are not only used for finding material in REASE, but also to inform people about the content when they look at the details of a resource and, hence, view the categories associated with the resource.

5 Conclusions

The initial REWERSE efforts in defining a Semantic Web topic hierarchy (deliverable E-D1) has led to establishing a broad cooperation with Knowledge Web in particular and contributed to building-up the Semantic Web community. The Semantic Web Topics Hierarchy resulting from this cooperation will have a durable impact on Semantic Web education and research. Some of the concrete durable effects are:

- A joint curriculum for a shared master program in the framework of the European Academy for Semantic-Web Education initiated by Knowledge Web (see Knowledge Web Deliverable D3.2.4) based on the Semantic Web Topics Hierarchy with REWERSE contribution described in 3.1
- A classification for Semantic Web learning resources, used as a catalogue in in the repos-

itory of the European Association for Semantic Web Education (Section 4). REASE created and maintained as a joint effort of Knowledge Web and REWERSE and is now also used as the central dissemination point for the results of the Knowledge Web industry area.

- A wiki presenting the Semantic Web topic hierarchy and facilitating discussion on and planning of future Semantic Web developments. In particular the wiki version of the Semantic Web Topic Hierarchy is now also used to include the Knowledge Web Technology Roadmap, which contains a summary of existing tools and applications and an outlook for the future for the most important topics in the Topic Hierarchy, which are related to the Semantic Web.