

SemanticWeb.org

- Content Analysis 2 -

Semantic Web Community Portal Project

Michael Stollberg

January 2004

Deliverable:
D4

Version:
2.1

Date:
29-Jan-04

DERI Ireland

University Road

Galway

IRELAND

www.deri.ie

DERI Austria

Technikerstrasse 13

A-6020 Innsbruck

AUSTRIA

www.deri.at

Table of Contents

1 EXECUTIVE SUMMARY	1
1.1 Project Members	1
2 SEMANTICWEB.ORG – CURRENT CONTENT.....	2
2.1 Screenshot / Areas.....	2
2.2 Left Column	3
2.2.1 <i>Mailing list Subscription</i>	3
2.2.2 <i>Projects / Products / Initiatives</i>	3
2.2.3 <i>SWWS (1st SW Working Symposium 2001)</i>	4
2.2.4 <i>Semantic Web Research Ontology</i>	5
2.2.5 <i>Links to Standardization Efforts</i>	5
2.3 Center Column.....	5
2.4 Right Column	5
2.4.1 <i>Welcome</i>	5
2.4.2 <i>Educational / Introductory Material</i>	5
2.4.3 <i>Resources</i>	6
2.4.4 <i>About SemanticWeb.org</i>	7
3 CONTENT STRUCTURE AND REPRESENTATION PROPOSAL	7
3.1 Information Types / Ontology Structure	8
3.1.1 <i>Static Content</i>	8
3.1.2 <i>Portal Information Items</i>	8
3.1.3 <i>Possible Future Extensions</i>	10
3.2 Information Coverage & Quality	11
3.3 Portal Functions and Services	14
3.3.1 <i>Functions</i>	14
3.3.2 <i>Services</i>	15
3.4 Information Representation / User Interface Design	16
4 CONCLUSION.....	17

1 Executive Summary

This document comprises the Content Analysis of Semantic Web.org – website (www.semanticweb.org). The analysis is a starting point to update the website into a decent Semantic Web driven web portal as a case study for the Semantic Web Community Portal Project. Therefore, the current content of the semanticweb.org - website is analyzed and ways to structure and represent it in a more sophisticated manner are discussed.

Section 2 presents an exhaustive description of the content currently available at the www.semanticweb.org, which has been already presented in the first version of the Deliverable D4. Section 3 takes this as a starting point for defining proposals for the update of www.semanticweb.org as a case study of the project, incl.

- verification of the ontology structure for the Portal Ontology (see D1)
- examination of problems arising, i.e. Coverage and Quality of Information
- proposal for portal functions and services

Section 3.2. and 3.3. are considered the most important parts of the paper, and can might serve as a requirements analysis for the update of SemanticWeb.org.

1.1 Project Members

Dieter Fensel

Michael Stollberg

Stefan Decker

Jos de Bruijn

Holger Lausen

Ying Ding

Ina M'Orchu

Michal Zaremba

Anna Zhdanova

Sung-Kook Han

Dumitru Roman



2 SemanticWeb.org – Current Content

2.1 Screenshot / Areas

The screenshot shows the SemanticWeb.org homepage in Microsoft Internet Explorer. The browser's address bar displays 'http://www.semanticweb.org'. The page title is 'SemanticWeb.org - Microsoft Internet Explorer'. The main content area is divided into three vertical columns:

- Left Column:** Contains sections for 'Join the SemanticWeb.org Mailing List!', 'Call for Semantic Web Initiatives!', 'Semantic Web Community Sites (provide Services for the Community)', and 'Semantic Web Project or Product Sites'. It lists various sub-projects like P2P, OpenCyc, DAML, TRIPLE, and others.
- Center Column:** Features the title 'SemanticWeb.org The Semantic Web Community Portal Together Towards a Web Of Knowledge...' and a 'News' section with a list of recent events and workshops, such as the 'Workshop on Semantic Web and Databases' and the '2nd International Semantic Web Conference'.
- Right Column:** Contains a 'Site Map' section with a 'Welcome' message and a list of navigation links including 'The Technology', 'Ontologies and Markup Languages', 'Inference Engines for the Semantic Web', and 'Resources'.

Three boxes at the bottom of the screenshot are connected by lines to their respective columns: 'Left Column' points to the left column, 'Center Column' points to the center column, and 'Right Column' points to the right column.

This structure description (which is based on the W3C-Homepage-style) will be followed throughout the analysis. In general it is to state that the website (it rather should be called a website and not a portal in its current state) contains a wide coverage of information for the Semantic Web Community. The structure of information has grown over time and appears to be a little bit chaotic. Currently, the website is not maintained soundly, thus information are mostly out of date.



2.2 Left Column

This column contains information about arbitrary aspects.

2.2.1 Mailing list Subscription

Information how to subscribe to the SemanticWeb – mailing list.

This is currently a YAHOO-egroups list which should be held at semanticweb.org

The list is frequently used, mostly for announcement related to the Semantic Web and as a exchange forum for novices / not-professionals.

2.2.2 Projects / Products / Initiatives

A great feature of the Website is that projects / products / initiatives can receive a web-domain NAME.semanticweb.org. This is a very good feature in order to raise importance of the portal.

Currently existing domains, incl. content check:

Table 1: Domain registered at NAME.semanticweb.org

URL (... .semanticweb.org)	Description	Remarks
deri.	...	Not even listed on the website !!
business.	SIG for business models for the Semantic Web	Not really used
challenges.	Meant for listing R%D-challenges to make the SW become reality	Not used, Relies on SemanticWeb.org – layout
Iswc	ISWC-homepage	Good / acceptable
p2p.	Short introduction to P2P as an infrastructure model for the SW	Short / static HTML-page, outdated
opencyc.	Link to Open Cyc page	
daml.	Link to DAML Project page	

triple.	Link to TRIPLE Homepage	
ontobroker.	Link to AIFB-page	Somehow cruel, outdated
annotation.	General information page about SW-Annotation	Good information resource
imorph.		DEAD LINK
Protégé.	Link to Protégé-Homepage	
SWAG.		DEAD LINK
COHSE.	Link to COHSE-project page (R&D-project)	
symontos.		DEAD LINK
healthcybermap.	Link to Healthcybermap.org- homepage	
networkinference.	Link to Network Inference Homepage	
swta.	Link to Dieter Homepage / CV	☺ what does it stand for ?
ontoweb.	Link to OntoWeb Project page / OntoWeb Portal	
kaon.	Link to KAON homepage	
ontoknowledge.	Link to OTK homepage	
sumo.	Home page of “Suggested Upper Merged Ontology”	OK
http://ccsw.dfki.de/	Link to CCSW homepage	Not: ccsw.semanticweb.org !!

2.2.3 SWWS (1st SW Working Symposium 2001)

Link to the very first international meeting, which now became ISWC

2.2.4 Semantic Web Research Ontology

Research Community Ontology in the style of (KA)2. Maintained by York Sure / AIFB. This should be investigated for the SW Portal Ontology !

2.2.5 Links to Standardization Efforts

<http://www.dmtf.org> develops Ontologies for the Telecommunication Industry

(Business Link)

<http://www.bpmi.org> develops a Process Ontology for Representing Business Processes

(BPMI – Business Process Management Initiative)

<http://www.papinet.org> develops global transaction standards for the paper supply chain.

(?Web Service Initiative)

<http://www.hr-xml.org> is dedicated to the development and promotion of standardized XML vocabularies for human resources (HR).

(Initiative for Human Resources Def. in XML)

2.3 Center Column

The center column contains news, i.e. what is new in the portal in every section. Mainly these are announcement of Events, but also from other sections of the portal.

This fills all the center area of the webpage. For update it should be implemented as a news-ticker, maybe in a separated area / on the edge. The center column should be left free for “real” content.

Contact information for the site-administration is given at the bottom.

2.4 Right Column

This contains the “About” – section and the “Information Repository“ of the website.

2.4.1 Welcome

Part of “About”

2.4.2 Educational / Introductory Material

- Ontologies and Markup Languages



- Inference Engines for the Semantic Web

This contains very short introductions to the Semantic Web idea and technologies and links to related resources. It is not sufficient for introducing novices into the Semantic Web area or to provide a basic overview.

In order to provide a reasonable introduction for novices / overview, there has to be some updating, as well as a framework of how to maintain the information.

2.4.3 Resources

This column contains the links to the “Information Repository” – which consists of static HTML-pages by now. The following content resources are existent, which are to be considered as ontology concepts in the SW portal ontology. In general this is the most important resources – they just have to be better structured and maintained / kept up to date. Most (nearly all) information are outdated.

Table 2: Information

Resource type	Description	Remarks
Event	Past & upcoming events	Short description + Link (usable format) Outdated and just enlisted for “kilometers”
Project	SW related projects	Not all projects enlisted
SW Interest Groups	R&D - / Standardization Groups	Mostly W3C groups
Call for Proposals	EU-IST call 2001	Outdated Only 1 entry
Research Programs	Self-explaining	Only DAML + TASK by now Only 1 entry
SW Companies	Businesses working in / with SW technologies	Not all listed, outdated
Tools and Knowledge Resources	List of tools / implementation initiatives	Not all listed, outdated

Ubiquitous Computing and the Semantic Web	Link to web-iq.com (Intelligent Wireless Web)	A initiative Only 1 entry
Emerging Standards	Links & short descriptions of research efforts	A bit ventured to call them “Emerging Standards” !?!
Industrial Deployments	Usage of SW technologies in industry	Only 1 entry
Related Publications and Publisher	W3C – activities and “famous” publications	Only a few authors mentioned, outdated
XSB	DDB-technology, short descriptions and links	
Library	Link to WordNet	...

2.4.4 About SemanticWeb.org

- website coordinators (SKC at Stanford, AIFB, Protégé)
 - => do we have to deal about rights / cooperation ?
- short “About”-text
- the Big Picture, showing (trying to show) the idea to the Semantic Web

3 Content Structure and Representation Proposal

In order to update the semanticweb.org – website into a sophisticated Semantic Web Community portal, its existing content has to be revealed, (re)structured, and possibly extended. Reasonable content structure and representation is determined by several aspects: what types of information shall be in the portal? What coverage and quality of information is aimed? What services shall be provided and shall the user interface look like?

The following addresses these issues, discusses the problems arising, and presents proposals. It is to note that these are mainly personally opinions of the authors and thus should not be understood as a specification but a proposal for further discussion for the SemanticWeb.org case study of the project. Moreover, there are different proposal in the project committee, especially in the deliverables “User Requirements Analysis” and the “Portal Ontology” which will be referred to.

The aim of this section is to determine the information structure and services for the SemanticWeb.org – portal (update) as a case study of our project. The intent is to point out the important aspects that have to be considered for the case study portal. It is to note that the following does not provide the final solution but rather a workplan.

3.1 Information Types / Ontology Structure

At first, we address the (re)structuring of informative content for SemanticWeb.org. The portal content consist of the information items that are semantically annotated and deposited in the portal as well as static, i.e. non-changing information.

3.1.1 Static Content

The latter (static information) comprises the following:

- Explanatory Information:
 - **“About”**: portal provider / sponsor / consortium (like the existing “About” – section)
 - **“Contact Information”**, for: web-interface (webmaster), administration (content), communities (SIG-leader)
 - **“Help”**: high-quality Help-Support for portal users (refer to SW Portal Survey for requirements on help-facility)
 - **“Documentation”**: explanation of the technologies used
- Access to:
 - **“Functions”**: a *Function* shall be a functionality offered by the portal, e.g. *Search* or *Information Submission*, etc. All functions have to be accessible
 - **“Services”**: a *Service* shall be a special function like a *News Ticker* or an *Event Calendar*, etc. – especially community features. All services have to be accessible

3.1.2 Portal Information Items

The more important content is the information items stored and retrievable in the portal. These have to be ontologically annotated in order to support semantically enhanced information processing. On basis of the current content analyzed above, Table 3 lists the main concepts of the domain ontology for the Semantic Web community, further definition and technical realization is provided in the “Portal Ontology”-deliverable. The current content structure of the SemanticWeb.org – website as identified above is

mapped to the new content structure. The ontology structure is inspired by previous work on ontologies for the Semantic Web research community, namely by the SWRC (Semantic Web Research Community Ontology)¹ and the OntoWeb Portal Ontology². A specific concept of the SemanticWeb.org Portal Ontology is the Educational Material – Concept which describes introductory material on the Semantic Web.

Table 3: Annotated Information Items for SemanticWeb.org (Proposal)

Information Type	Description	old semanticweb.org – Concepts
Event <u>Subconcepts:</u> Conference, Workshop, Tutorial	All Semantic Web – related events (announcement, CFP, etc.) Divided into upcoming / past events with automated maintenance	Event
Initiative <u>Subconcepts:</u> RnD (Project, Programme), Business, Standardization	Initiatives working in the Semantic Web This concept shall comprise all SW-activities : research programmes & projects, commercial / standardization / development efforts	Project, SW Interest Groups, Research Programs, SW Companies, Ubiquous Computing ... (as an instance), Emerging Standards, Industrial Deployments, XSB, NAME.semanticweb.org
Publication <u>Subconcepts:</u> Thesis (Master, PhD), Book, Article, Technical Report, Book Chapter, Proceedings	Publications enlisted / downloadable All types of scientific publications, scope to be defined (see below)	Related Publications & Publishers
Organisation <u>Subconcepts:</u> University, Research Institute, Company	Organizations working in Semantic Web	Not really separated now

¹ See: <http://ontobroker.semanticweb.org/ontos/swrc.html>

² see: www.ontoweb.org

Person	All members of SemanticWeb.org portal. Only members are allowed to submit new information items. For expert finder.	Not supported by now
News	News announced in the portal	Call for Proposal / papers (as an instance)
Tool	Semantic Web tools / products Description of Tools (free / license-ware, commercial)	Tools and Knowledge Resources
Educational Material	Introductory material for novices	Educational / Introductory Material (see2.4.2)

3.1.3 Possible Future Extensions

Apart from the information items listed in Table 3, the ontology can be extended in future versions of the portal. Possible extensions with respect to the Semantic Web Community – which would imply further portal functionalities – could be:

Ontology Library

Currently there are only proprietary, not user-friendly collections of ontologies – see DAML Ontology Library³. Following the semanticweb.org – mailing list, there are very often questions on existing, reusable ontologies.

As pointed out by [Ding and Fensel, 2001], a sophisticated ontology library has to provide appropriate means for ontology management (storage, versioning), adaption (search, editing) and supporting standards for import / export. As some of these functions are already in a SW portal, extending it to an Ontology Library System will be possible.

An Ontology Library for commonly used ontologies would be a great surplus for the Semantic Web Community. Portal Members would be enabled to publish their ontologies here.

³ see: www.daml.org/ontologies

Semantic Web Services Library / Register

Similar to an Ontology Library, a Semantic Web Service Register could be provided by the portal. This would have to support emerging standards for SWS description / access / etc., and the Web Services could be ontologically described (not only technically, but also for advertising), thereby dissemination of Web Services would be supported – similar to an Ontology Library for reusing ontologies.

Such a SWS Register / Library would also give keen surplus for the Semantic Web community.

3.2 Information Coverage & Quality

The second aspect to be investigated for the future content of SemanticWeb.org is the coverage, i.e. what breadth and what depth of information shall be covered, and quality of information in the portal, i.e. how to ensure a high quality regarding the content of information. Here, we point out important issues that need to be observed for creating the SemanticWeb.org – portal.

Clearly, the SemanticWeb.org – Portal is a vertical portal (see “Commercial Portal Products”) with a specialized field of interest. Nevertheless, it comprises a wide range of very different aspects – from ontology languages over ontology technologies to Semantic Web Services. Moreover, it is planned to have two general categories of content (see “User Requirements Analysis”): beginners or novices content which contains introductory and educational information for novices in the area and expert content for knowledgeable researchers.

This content will be ontologically annotated and be automatically processable as information items in the portal. Thus we will discuss the coverage and quality of information for the specific content ontology concepts identified in Table 3. In general, for the coverage of content it has to be decided whether all information related to the topic shall be covered in the portal or just selected issues and how this content shall be displayed (e.g. references to other web resources or explicitly holding everything in the portal). Regarding quality of information (QoI), the most reasonable recommendation is to have a quality assurance step during information item creation: a group of selected portal administrators have to verify a new information item before it will be displayed to the public. Apart from identifying these administrators, there also has to be QoI-guideline. Furthermore, a framework of maintenance of information items is required in order to keep information update and (re)move outdated or dead-linked information.

Table 4 summarizes recommendation for these aspects. It is to remark how users will be encouraged to use the portal – which is determined by the usefulness of available information for the user to a very high extend. Thus, it also has to be considered how to update the content in order to make SemanticWeb.org THE portal for the Semantic Web. Users can not be forced to enter all information manually, so there should be some solutions for interchanging information items with other portals / information resources and for automatically extracting information items from web resources. Required solutions – which actually should be Semantic Web tools – will be investigated in section 3.3 in more detail.

Table 4: Coverage and Quality of Information on SemanticWeb.org (Proposal)

Information Type / Ontology Concept	Information Coverage	Quality of Information
Event	<ul style="list-style-type: none"> - should comprise all events related to SW / SWS - <i>announcements should be a) distributed (over mailing list), b) automatically added when announced somewhere else</i> 	<ul style="list-style-type: none"> - administrator – QoI - automated maintenance (moving from “upcoming” to “past”)
Initiative	<ul style="list-style-type: none"> - related initiatives are referenced here 	<ul style="list-style-type: none"> - administrator – QoI - important: keep up to date & regularly check linking
Publication	<ul style="list-style-type: none"> - What kind of publications? We don't need a 2nd Citeseer! => possibilities: 1) only most important ones (BUT: who decides?!?) 2) provide ontology-based-literature-web-search 3) timed web crawler (see Michal / Anna's idea) 4) (semi)automated im- & export, interchange with other portals <i>... => to be discussed</i> 	<ul style="list-style-type: none"> - administrator – QoI - <i>most other portals lack in quality of information because users do not want to enter every publication manually.</i>

Organisation	- related organizations are referenced here	- administrator – QoI - important: keep up to date & regularly check linking
Person	- all members of the portal - usable for expert finder.	- registered members get assigned & maintained automatically
News	1) news in the portal (start page) 2) news in the community (news ticker) - <i>how to get (2)? Should be automated, e.g. automatically checking related websites (W3C, DERI, ...)</i>	1) retrieved automatically 2) also retrieved automatically? - news archive needed?
Tool	- Semantic Web tools / products - <i>aim should be to list (nearly) all related tools & products => how to do that??</i>	- administrator – QoI - important: keep up to date & regularly check linking
Educational Material	<u>Form “User Requirements Analysis”:</u> • A short description, saying what the Semantic Web is • A tutorial on the Semantic Web • A paper, explaining in more detail what the Semantic Web is about • A tutorial on ontologies • A tutorial on Semantic Web Services • A paper, explaining Semantic Web Services <u>further:</u> - links to most important resources - announcements of introductory tutorials / workshops - educational institutes (e.g. DERI Austria)	- needs to be defined by administrators - <i>how to update / maintain?</i>

In order to classify information items regarding the topic / area they are concerned with is to have a general ontology concept “Topic” which defines the area, for example “Ontology Edition” or “SWS Description”. This is referenced by related the information

item concepts, e.g. in publication as “topic”, by Tool as “for /about” or by Person as “Working field”. Topic-instances can be created during system runtime and are to be quality assured by the normal administrator-QoI.

3.3 Portal Functions and Services

After having determined the content structure, its coverage and required functionalities for creation and maintenance, this section investigates functions and services which should be provided by the SemanticWeb.org Portal. Thereby a function is understood as an information processing function for information items of the portal while a service is a user-service that provides an additional feature.

3.3.1 Functions

Portal Functions are the grounding information processing features of the portal. Thus the required functions will be exposed according to the portal information item workflow defined in [SW Portal Survey]. The following is extended from the “User Requirements Analysis”:

Creation

- Form-based creation support incl. immediate ontology assignment
- Import of existing information items from other portals
- Special features for specific Content Concepts:
 - Event: automated monitoring of other event-announcement resources
 - Publication: automated monitoring / crawling / search of other resources
 - Person: automated instance creation when new member registers
 - News: (1)-type: automatically deriving when new information item is added; (2)-type: automatically monitoring other news-resources

Publication

- administrator-QoI workflow: Submission by users, quality-assurance by portal administrators
- Tool Support for QA (similar for every Content Concept with QoI)

Organization

Storage & Indexing of Information items, important aspects:



- information items have to be retrievable with ontological assignments, also via other interfaces than the portal-UI
- storage & item management has to be secure & scalable (short-coming of ontology-repositories)

Access

- = Search Facilities on information items:
(*important: highly usable GUI, provided as SWS*)
 - **Ontology-Browsing:** hierarchy < advanced hierarchical (see OntoWeb) < visualized (see Mondeca-tool)
 - **Ontology Search:** full text search on ontological notions of information items
 - **Enhanced Ontology Search:** inference-powered / logical query support
 - **(Full) Text Search:** on documents
- Special features for specific Content Concepts:
 - Event: automated distributed over mailing list / other medium
 - Publication: automated monitoring / crawling / search of other resources

Maintenance

- Generally: Content Management (i.e. maintenance) for all information items and for the ontology schema needed
- Special features for specific Content Concepts:
 - Event: automatically moving to “past events” when over
 - News: news archive (needed?)
 - Initiative, Organization, Tool: automated link checking
 - Educational Material: manual update required (determine responsible)

3.3.2 Services

Services are additional functionalities offered by the portal that provide surplus value for the user, especially community features. The following enlists appropriate services for the SemanticWeb.org Portal, based to the results of the “Groupware Evaluation” – Deliverable. This list should be regarded as a proposal and not to be finite. It is to note that most of the services mentioned are not related to Semantic Web technologies and it has to be considered in what way semantic annotations can improve these facilities.

- Email facilities:
 - Mailing list: existing semanticweb – mailing list
 - Newsgroups (for SIGs)
- Wiki / Weblogs / Blogging: for expert exchange / SIGs to support distributed collaboration. This could be based on Bruno's tool.
- News Ticker (for news in the community)
- Chat / Discussion Forums (for SIGs)
- Event Calendar (for upcoming events, incl. links to the Event information item).

3.4 Information Representation / User Interface Design

The last aspect of this section is how information should be presented on the portal. As the author does not consider himself to be a good designer and further design issues would exceed the scope of this paper, Figure 1 only shows the abstract user interface structure. It is mainly divided into 4 sections: the Head with the function access points and user navigation information, the Left Frame with Login and Content Browser, the Main Area and the Right Frame which holds portal services.

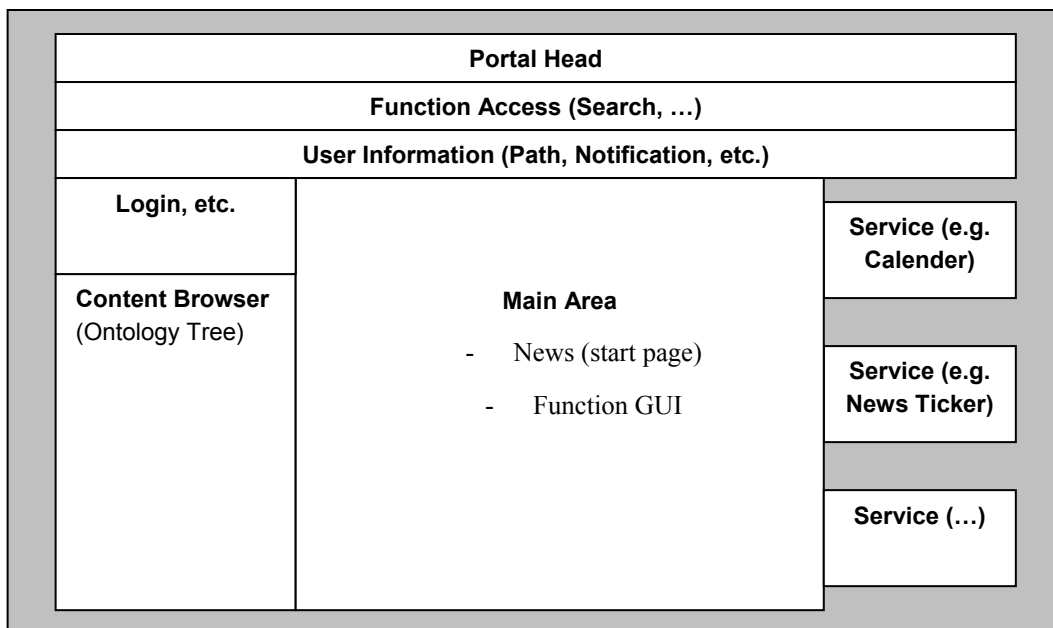


Figure 1: Information Presentation Structure (Web Interface)

The content of the Main Area displays the functional interface during user interaction. The start page when entering the portal can hold the type-(1)-news (i.e. new information items added to the portal), while during usage the presentation is determined by the ontology definition of the specific information item.

For the final web interface design, recommended portals are the OntoWeb portal⁴ and the Mondeca ITM⁵.

4 Conclusion

This deliverable has analyzed the content of the current semanticweb.org – website. It has been determined that this is outdated at this point of time, but it can be easily transferred into a reasonable ontology for the SemanticWeb.org Portal.

On this basis, the information structure of the SemanticWeb.org Portal Ontology has been derived and issues for the realization of this portal as a show case for the SW Portal Project have been discussed. It is to note that these aspects shall be understood as a proposal for discussion and further elaboration, but they do not state the final specification for the portal.

⁴ see: www.ontoweb.org

⁵ see: <http://www.mondeca.com/english/documents.htm>